Coupled Problems in Science and Engineering VI

San Servolo, Venice, Italy, 18 – 20 May 2015

B. Schrefler, E. Oñate and M. Papadrakakis (Eds.)

Coupled Problems in Science and Engineering VI

COUPLED PROBLEMS 2015

Proceedings of the VI International Conference on
Coupled Problems in Science and Engineering
San Servolo, Venice, Italy
May 18 – 20, 2015

Edited by:

Bernhard A. Schrefler
Department of Civil, Environmental and Architectural Engineering
Università degli Studi di Padova, Italy

Eugenio Oñate
International Center for Numerical Methods in Engineering (CIMNE), Spain
Universitat Politècnica de Catalunya (UPC)

Manolis Papadrakakis
Institute of Structural Analysis & Antiseismic Research
National Technical University of Athens, Greece

A publication of:
International Center for Numerical Methods in Engineering (CIMNE)
Barcelona, Spain

International Center for Numerical Methods in Engineering (CIMNE)
Gran Capitán s/n, 08034 Barcelona, Spain

COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis (Eds.)

First Edition: April 2015

Printed by: Artes Gráficas Torres S.L., Huelva 9, 08940 Cornellà de Llobregat, Spain

ISBN: 978-84-943928-3-2

TABLE OF CONTENTS

Preface 7

Supporting Organizations 9

Acknowledgements 10

SUMMARY 11

Contents 13

Invited Sessions 21

Contributed Sessions 635

Authors Index 1336

PREFACE

This volume contains the full papers accepted for presentation at the **VI International Conference on Computational Methods for Coupled Problems in Science and Engineering, *COUPLED PROBLEMS 2015* (18 - 20 May 2015, San Servolo, Venice, Italy).**

The previous five editions of this conference were held on the islands of *Santorini (Greece) on 25-28 May 2005, Ibiza (Spain) on 21-23 May 2007, Ischia (Italy) on 8-11 June 2009, Kos (Greece) on 20-22 June 2011 and Ibiza (Spain) on June 17 – 19 June 2013.*

The objectives of **COUPLED PROBLEMS 2015** are to present and discuss state of the art, mathematical models, numerical methods and computational techniques for solving coupling problems of multidisciplinary character in science and engineering. The conference goal is to make a step forward in the formulation and solution of real life problems with a multidisciplinary vision, accounting for all the complex couplings involved in the physical description of the problem.

The conference is one of the Thematic Conferences of the European Community on Computational Methods in Applied Sciences (**ECCOMAS**) and a Special Interest Conference of the International Association for Computational Mechanics (**IACM**).

The conference is jointly organized by the Department of Civil, Environmental and Architectural Engineering (DICEA), of the University of Padova (Italy), the International Centre for Numerical Methods in Engineering (CIMNE), of the Technical University of Catalonia (UPC) and the National Technical University of Athens (Greece).

The organizers would like to thank the authors for submitting their contributions and for their respect of the deadlines. Special thanks go to the colleagues who contributed to the organization of the **23 Invited Sessions** in the fields of the Conference, and to the colleagues of international prestige that accepted the invitation to address a **Plenary Lecture**.

Bernhard A. Schrefler

Department of Civil, Environmental and Architectural Engineering (DICEA)
Università di Padova
Padova, Italy

Eugenio Oñate

International Center for Numerical Methods in Engineering (CIMNE)
Univ. Politècnica de Catalunya (UPC),
Barcelona, Spain

Manolis Papadrakakis

Institute of Structural Analysis & Antiseismic Research
National Technical University Athens, Greece

SUPPORTING ORGANIZATIONS

Department of Civil, Environmental and Architectural Engineering (DICEA)

of the University of Padova, Italy

International Center for Numerical Methods in Engineering (CIMNE), Spain

National Technical University of Athens (NTUA), Greece

European Community on Computational Methods in Applied Sciences (ECCOMAS)

International Association for Computational Mechanics (IACM)

Technical University of Catalonia (UPC), Spain

ACKNOWLEDGEMENTS

The conference organizers acknowledge the support towards the organization of the COUPLED PROBLEMS 2015 Conference to the following **organizations**:

- Department of Civil, Environmental and Architectural Engineering (DICEA), Università di Padova, Italy
- European Community on Computational Methods in Applied Sciences (ECCOMAS)
- International Association for Computational Mechanics (IACM)
- International Center for Numerical Methods in Engineering (CIMNE), Barcelona, Spain
- National Technical University of Athens, Greece
- Universitat Politècnica de Catalunya, Spain

PLENARY SPEAKERS AND INVITED SESSION ORGANIZERS

We would also like to thank the Plenary Speakers and the Invited Session Organizers for their help in the setting up of a high standard Scientific Programme.

Plenary Speakers: *Santiago Badia, Yuri Bazilevs, Ronaldo Borja, Charbel Farhat, Dariusz Gawin, Christian Hellmich, Jacques M. Huyghe, Sergio Idelsohn, Kazuo Kashiyama, Elisabeth Massoni and Ole Sigmund.*

Invited Session Organizers: *Carlos Agelet de Saracibar, Piergiorgio Alotto, Gernot Beer, Pavel Bochev, Daniele Boffi, Norbert Böttcher, Oreste S. Bursi, Eric Cyr, Suvranu De, Paolo Decuzzi, Eduardo Divo, Carlos Felippa, Lucia Gastaldi, Massimo Guarnieri, Thomas J.R. Hughes, Alexander Karapetyan, Alain J. Kassab, Moon Ki Kim, Alexey Kireenkov, Manfred Krafczyk, Ju Liu, Wing Kam Liu, Bernd Markert, Miriam Mehl, Ralf-Peter Mundani, Roger Ohayon, Sebastià Olivella, K.C. Park, Jean-Philippe Ponthot, Florin Adrian Radu, Ernst Rank, Lorenzo Sanavia, Bozidar Šarler, Guglielmo Scovazzi, John Shadid, Lian Shen, Fotis Sotiropoulos, Dörte Carla Sternel, Bozidar Stojadinovic, Claudio Tamagnini, Haim Waisman and Zohar Yosibash*

SUMMARY

INVITED SESSIONS

IS - Advanced ALE (Arbitrary Lagrangian - Eulerian) Methods for Multiphysics Systems, *organized by John N. Shadid, Eric C. Cyr and Guglielmo Scovazzi* 23

IS - Computational Models and Methods for Multiphysics Processes in Multiphase Porous Media, *organized by Lorenzo Sanavia and Claudio Tamagnini* 35

IS - Coupled and Multi-Scale Bioengineering Applications, *organized by Eduardo Divo and Alain Kassab*.....116

IS - Coupled Problems in Biomechanics and Mechanobiology, *organized by Bernd Markert and Zohar Yosibash*.....125

IS - Coupled Problems in Free Surface Flow, *organized by Fotis Sotiropoulos and Lian Shen* 137

IS - Coupled Problems in Geo-environmental Engineering, *organized by Sebastià Olivella and Norbert Böttcher* 149

IS - Coupled Problems Involving Liquid-Solid Phase Transitions, *organized by Božidar Šarler*.....160

IS - Coupled Problems of the Dynamics of the Solid Systems with Friction, *organized by Alexander Karapetyan and Alexey Kireenkov* 191

IS - Coupled Problems with Numerical and Physical Subdomain Interactions, *organized by organized by Oreste Bursi and Bozidar Stojadinovic*.....222

IS - Coupled Reactive Transport and Multiphase Flow in Porous Media, *organized by Florin Adrian Radu*.....300

IS - Coupled Thermomechanical Modeling of Material Forming Processes, *organized by Carlos Agelet de Saracibar and Jean-Philippe Ponthot* 312

IS - Coupling Multiscale Infrastructure Models and Numerical Simulation, *organized by Ernst Rank, Manfred Krafczyk and Ralf Mundani* 334

IS - Coupling of Different Numerical Methods, *organized by Gernot Beer* 358

IS - Energy Related Coupled Problems,
organized by Massimo Guarnieri and Piergiorgio Alotto 382

IS - Iterative Methods for Coupled Field Problems,
organized by Suvranu De, John N. Shadid and Haim Waisman 452

IS - Multiphysics Simulations with Time Resolved Turbulent Flow Fields,
organized by Daniele Boffi and Lucia Gastaldi 477

IS - Multiple Scale Approach to Biological Problems,
organized by Wing Kam Liu and Moon Ki Kim 566

IS - Nanomedicine: From In Silico to In Vivo,
organized by Paolo Decuzzi 576

IS - Phase Field Modeling of Coupled Problems,
organized by Ju Liu and Thomas J.R. Hughes 605

CONTRIBUTED SESSIONS

Applications of Coupled Problems 635

Coupled Fluid Flows 753

Coupled Problems in Structural Mechanics 822

Coupled Solution Strategies 881

Design, Optimization and Control 909

Electromagnetics 976

Fluid Structure Interaction 1037

Geomechanics 1209

Numerical Methods for Coupled Problems 1261

Thermomechanics 1291

CONTENTS

INVITED SESSIONS

IS - Advanced ALE (Arbitrary Lagrangian - Eulerian) Methods for Multiphysics Systems

A conservative, optimization-based semi-lagrangian spectral element method for passive tracer transport
P. Bochev, S. Moe, K. Peterson, D. Ridzal* 23

IS - Computational Models and Methods for Multiphysics Processes in Multiphase Porous Media

Analysis of stress partitioning in biphasic mixtures based on a variational purely-macroscopic theory of compressible porous media: Recovery of Terzaghi‘s law
R. Serpieri, F. Travascio, S. Asfour, L. Rosati* 35

Finite strains fully coupled analysis of a horizontal wellbore drilled through a porous rock formation
N. Spiezia, V. Salomoni, C.E. Majorana* 47

Hygrothermal behaviour of hemp concrete; experimental evidences and modelling
A. Fabbri, P.A. Chabriac, D.T. Ngo, F. Sallet, E. Gourdon, H. Wong, J.C. Morel* 59

Numerical coupling between damage and gas permeability for concrete applied on a 3D splitting test
M. Ezzedine El Dandachy, M. Briffaut, F, Dufour, S, Dal Pont* 71

Simulation of consolidation in large strains: A comparison between finite element method and material point method
N. Spiezia, F. Ceccato, V. Salomoni, P. Simonini* 82

Thermo-mechanical properties of magnesia carbon foam composites
A. Jung, S. Diebels* 94

Two sides of a coin: A critical review, and mathematical and phenomenological study of what we call hydro-mechanical coupling
J.M. Huyghe, E. Nikooee, T. Sweijen, S.M. Hassanizadeh* 106

IS - Coupled and Multi-Scale Bioengineering Applications

A parallel fully-coupled fluid-structure interaction simulation of a cerebral aneurysm
*A. Eken, M. Sahin** 116

IS - Coupled Problems in Biomechanics and Mechanobiology

Advanced image processing methods for automatic liver segmentation
P. Strakos, M. Jaros, T. Karasek, T. Kozubek, P. Vavra, T. Jonszta* 125

IS - Coupled Problems in Free Surface Flow

Validation of high fidelity CFD/FE FSI for full-scale high-speed planing hull with composite bottom panels slamming
S. Volpi, H. Sadat-Hosseini, M. Diez, D-H. Kim, F. Stern, R. Thodal, J. Grenestedt* 137

IS - Coupled Problems in Geo-environmental Engineering

Keynote Lecture: On the term and concepts of numerical model validation in geoscientific applications
U-J. Görke, T. Nagel, O. Kolditz* 149

IS - Coupled Problems Involving Liquid-Solid Phase Transitions

A RBF-based local collocation method for modelling thermomechanical phenomena during DC casting of aluminium billets
B. Mavric, B. Šarler* 160

Modelling of electromagnetic breaking and electromagnetic stirring in the process of continuous casting of steel
K. Mramor, R. Vertnik, B. Šarler* 169

Simulation of multiscale industrial solidification problems under influence of electromagnetic field by a meshless method
B. Šarler, A. Gustin, V. Hatic, N. Kosnik, B. Mavric, R. Vertnik* 181

IS - Coupled Problems of the Dynamics of the Solid Systems with Friction

Contact types hierarchy and its object-oriented implementation
I.I. Kosenko, K.V. Gerasimov, M.E. Stavrovskiy* 191

Keynote Lecture: Further development of the theory of multicomponent dry friction
*A.A. Kireenkov** 203

Rotation of the body with movable internal masses around the center of mass on the rough plane
*A. Sakharov** 210

IS - Coupled Problems with Numerical and Physical Subdomain Interactions

Consolidated fire testing - A framework for thermo-mechanical modelling
*P. Schulthess, M. Neuenschwander, M. Knobloch, M. Fontana** 222

Design and development of ten-element hybrid simulator and generalized substructure element for coupled problems
S. Mojiri, X. Huang, O-H. Kwon, C. Christopoulos* 230

Heterogeneous Asynchronous Time Integrators built from the energy method for coupling Newmark and α-schemes
M. Brun, A. Gravouil, A. Combescure* 242

Hybrid simulation of heat transfer problems in structural applications
G. Abbiati, O.S. Bursi, B. Stojadinovic, N. Tondini, C. Whyte* 254

Keynote Lecture: Hybrid simulation of complex structural systems based on partitioned time integration schemes
*G. Abbiati, E. Cazzador, O.S. Bursi, P. Pegon, F.J. Molina, F. Paolacci** 266

Predominant periods of multi-degree-of-feedom system analysis and dynamic soil-structure interaction for building structures
M. Inukai, T. Kashima, T. Saito, T. Azuhata* 278

Validation of OpenFresco-based thermomechanical hybrid simulation to address an earthquake-fire coupled problem
C. Whyte, K.R. Mackie, G. Abbiati, B. Stojadinovic* 290

IS - Coupled Reactive Transport and Multiphase Flow in Porous Media

Numerical simulation of geochemical compaction with discontinuous reactions
*A. Agosti, L. Formaggia, B. Giovanardi, A. Scotti** *300*

IS - Coupled Thermomechanical Modeling of Material Forming Processes

A magneto-thermo-metallurgical finite element model applied to induction hardening processes
M. Spezzapria, M. Forzan, F. Dughiero* *312*

Numerical methods in simulation of resistance welding
C.V. Nielsen, P.A.F. Martins, W. Zhang, N. Bay* *322*

IS - Coupling Multiscale Infrastructure Models and Numerical Simulation

A computational multiscale approach to couple hygro-mechanical responses of large-scale masonry walls
G. Castellazzi, S. de Miranda, G. Formica, L. Molari, F. Ubertini* *334*

Modelling of air chamber supported floating platforms - Coupling free surface flow, compressible air, and flexible structures
F. Toth, M. Kaltenbacher, F. G. Rammerstorfer* *346*

IS - Coupling of Different Numerical Methods

Coupling of boundary element regions with the boundary element tearing and interconnecting method (BETI)
C. Duenser, B. Lindner, G. Beer* *358*

Implementation of an efficient coupled FEM-SBFEM approach for soil-structure-interaction analysis
M. Schauer, S. C. Langer* *370*

IS - Energy Related Coupled Problems

Boundary element solution of 2D coupled problem in anisotropic piezoelectric FGM plates
*M. Abdelsabour Fahmy** *382*

Coupled numerical multiphysics simulation methods in induction surface hardening
D. Schlesselmann, B. Nacke, A. Nikanorov, S. Galunin* *392*

Multi-physics and multi-objective design of a benchmark device: A problem of inverse induction heating
*P. Di Barba, F. Dughiero, M. Forzan, E. Sieni** *404*

Multiphysics finite-element modelling of an all-vanadium redox flow battery for stationary energy storage
*F. Moro, A. Bertucco, V. Fiorenzato, M. Giomo, M. Guarnieri** *416*

Multiphysics simulation tools for designing motors for traction applications in hybrid and electric vehicles
N.P. Sawant, C.P. Riley, A. Venskus, D.H. Vassilev, J.D. Wale, E. Wearing, A.M. Michaelides, P.J. Topping, H. Matharu* *428*

On the pseudo-incident wave technique for interacting inhomoheneities in electromechanical problems
X.D. Wang, H.C. Yu, S. Abdel-Gawad, C. Wang* *440*

IS - Iterative Methods for Coupled Field Problems

Convergence speed of coupling iterations for the unsteady transmission problem
A. Monge, P. Birken* *452*

Coupling requirements for well posed and stable multi-physics problems
J. Nordström, F. Ghasemi* *464*

IS - Multiphysics Simulations with Time Resolved Turbulent Flow Fields

A comparison of various quasi-Newton schemes for partitioned fluid-structure interaction
*F. Lindner, M. Mehl, K. Scheufele, B. Uekermann** 477

A dynamically adaptive Lattice Boltzmann Method for predicting wake phenomena in fully coupled wind engineering problems
R. Deiterding, S.L. Wood* 489

A hybrid LES/CAA method applied to a 3D shear flow simulation
O. Labbé, C. Peyret* 501

Comparison of advanced turbulence modeling approaches for fluid-structure interaction
A. Ali, T. Reimann, D. C. Sternel, M. Schäfer* 512

Coupled CFD-CAA approach for rotating systems
M. Kaltenbacher, A. Hüppe, A. Reppenhagen, W. Kuehnel* 524

Extension of the LS-STAG immersed boundary method for RANS-based turbulence models and its application for numerical simulation in coupled hydroelastic problems
V.V. Puzikova, I.K. Marchevsky* 532

On in-situ visualization for strongly coupled partitioned fluid-structure interaction
O. Fernandes, D.S. Blom, S. Frey, A.H. van Zuijlen, H. Bijl, T. Ertl* 544

On parallel scalability aspects of strongly coupled partitioned fluid-structure-acoustics interaction
D.S. Blom, V. Krupp, A.H. van Zuijlen, H. Klimach, S. Roller, H. Bijl* 556

IS - Multiple Scale Approach to Biological Problems

How the morphologies of progenitor and mature osteocytes contribute to their mechanotransduction
E. Budyn, M. Bensihoum, T. Marsan, F. Mannemare, S. Sasnouski, P. Tauc, E. Deprez, H. Petite* 566

IS - Nanomedicine: From In Silico to In Vivo

Coupled rings in RNA nanotubes and properties of biological nanoclusters
S. Badu, R. Melnik, S. Prabhakar* 576

Coupling fluid and solid domains in modeling drug transport within tumor
M. Kojic, M. Milosevic, V. Simic, A. Ziemys, M. Ferrari* 583

Mass transport phenomena at the solid-liquid nanoscale interface in biomedical applications
A. Cardellini, M. Fasano, E. Chiavazzo, P. Asinari* 593

IS - Phase Field Modeling of Coupled Problems

3D phase-field for pressurized fracture propagation in heterogeneous media
T. Wick, S. Lee, M.F. Wheeler* 605

An electro-chemo-mechanical analysis of solid oxide fuel cell considering evolution of microstructure in porous electrode using phase-field method
M. Muramatsu, S. Takase, K. Yashiro, T. Kawada, K. Terada* 614

CFD simulation of microscopic two-phase fluid motion on solid body with edges and heterogeneously-wetted surface using phase-field model
N. Takada, J. Matsumoto, S. Matsumoto, K. Kurihara* 625

CONTRIBUTED SESSIONS

Applications of Coupled Problems

3D numerical approximation of relativistic particle beams by asymptotic expansion
F. Assous, A. Le Blanc* 635

Adjoint harmonic balance method for forced response analysis in turbomachinery
A. Engels-Putzka, C. Frey* 645

Approaches for coupled numerical simulation of high frequency tube welding process
A. Nikanorov, E. Baake, H. Brauer, C. Weil* 657

Ballistic charge transport in a triple-gate Silicon Nanowire transistor
O. Muscato, T. Castiglione, C. Cavallaro* 666

Coupled Problems of localization of dust and gas
O.A. Averkova, K.I. Logachev, I.N. Logachev, A.E. Canar* 677

Drop coalescence simulations using level sets coupled with boundary integral methods
M. Garzon, L.J. Gray, J.A. Sethian* 689

Fluid-thermal-structure coupled analysis of grid fins for hypersonic flight vehicle
S. Li, Z. Jiang, W. Zhang, K. Peng* 701

Mesoscopic modelling of the RTM process for homogenization
C. Dammann, R. Mahnken* 713

Migration of a rigid disc in Couette flow subject to an external electric field simulated using ISPH
N. Tofighi, M. Ozbulut, A. Rahmat, J.J. Feng, M. Yildiz* 725

Modelling and numerical calculation of piston-like oil displacement for doubly-periodic systems of oil fields development
V.I. Astafev, A.E. Kasatkin* 734

Parameter identification of Chaboche material model using indantation test data and inverse approach
M. Cermak, R. Halama, T. Karasek, J. Rojicek* 743

Coupled Fluid Flows

A high-order fully coupled Electro-Fluid-Dynamics (EFD) solver for multi-phase flow simulations
N. Emamy, M. Karcher, R. Mousavi , M. Oberlack* 753

Analysis on flow around a sphere at high mach number, low Reynolds number and adiabatic condition for high accuracy analysis of gas particle flows
T. Nagata, T. Nonomura, S. Takahashi, Y. Mizuno, K. Fukuda* 760

Coupled simulation of shock waves in gas-particle mixtures introducing motion equations
Y. Mizuno, S. Takahashi, T. Nonomura, T. Nagata, K. Fukuda* 772

DO transport regimes across the water-sediment interface under low-energetic flows
E. Sánchez-Badorrey, F.J. García Anguita* 783

Numerical simulation of acoustics in heterogeneous media
D. Staab, S. Nowak, D. C. Sternel, M. Schäfer* 791

The efficiency comparison of the vortex element method and the immersed boundary method for numerical simulation of airfoil's hydroelastic oscillations
I.K. Marchevsky, V.S. Moreva, V.V. Puzikova* 800

The Navier-Stokes equations in two different formulations with moderate and high Reynolds numbers
B. Bermúdez, W.F. Guerrero* 812

Coupled Problems in Structural Mechanics

A numerical component to evaluate the iodine corrosion in a nuclear field accounting for chemisorption mechanisms Zirconia layer impact on SCC
N. Barnel, D. Palermo* 822

Complete formulation of the subloading surface model
*K. Hashiguchi** 837

Finite element analysis of damage-hearing behaviour in self-healing ceramic materials
S. Ozaki, Y. Horie, K. Hinata, T. Osada, W. Nakao* 849

Robust optimization approach for mixed numerical/experimental identification of elastic properties of orthotropic composite plates
J. Auzins, E. Skukis* 859

Simulation of self-healing processes in microcapsule based polymeric systems
S. Specht, J. Bluhm, J. Schröder* 871

Coupled Solution Strategies

A fast frequency sweep approach with a priori choice of Padé approximants and control of their interval of convergence
R. Rumpler, P. Göransson* 881

Propagation of acoustic-gravity waves in inhomogeneous ocean environment based on modal expansions and hp-FEM
*K.A. Belibassakis, G.A. Athanassoulis, A.E. Karperaki, T.K. Papathanasiou** 893

Design, Optimization and Control

A sensitivity based optimization approach for aeroacoustic problems
S. Nowak, D.C. Sternel, M. Schäfer* 909

Coupled inverse problems and visualization of atmosphere-ocean system
*S. Kabanikhin, O. Krivorotko** 921

Modelling pressure cycle and interaction with reed valves in a reciprocating compressor
B. Lemoine, L. Le Marrec, A. Hirschberg* 930

Multipoint global optimal shape design by morphing
*A. Nastase** 939

Resistance reduction of a military ship by variable-accuracy metamodel-based multidisciplinary robust design optimization
*C. Leotardi, E.F. Campana, M. Diez** 952

Structural design using a parallel sequential approximate optimization
D. Wang, Z. Wu, F. Hu, Z. Jiang* 964

Electromagnetics

3D-Simulation of electromagnetic and temperature fields in the continuous induction heaters
V.B. Demidovich, F.V. Chmilenko, V. V. Andrushkevich, I.I. Rastvorova* 976

An electroelastic problem of a semi-infinite body with D_∞ symmetry subjected to distributed surface loading
M. Ishihara, Y. Ootao, Y. Kameo* 985

Analysis of actuator system made of FGM using new electro-thermo-mechanical finite elements
J. Paulech, J. Hrabovský, J. Murín, V. Kutiš* 993

Comparing kinetic and hydrodynamical models for electron transport in monolayer graphene
*M. Coco, A. Majorana, G. Mascali, V. Romano** *1003*

Electro-thermo-mechanical analysis of SMA actuator
V. Kutiš, P. Drahoš, J. Murín, J. Paulech, J. Hrabovský* *1015*

Finite element analysis using a hierarchal decomposition for the interaction of structure, fluid and electrostatic field in MEMS
D. Ishihara, T. Horie, T. Niho, A. Baba* *1023*

The nanoCOPS project on algorithms for nanoelectronic coupled problems solutions
R. Janssen, J. ter Maten, C. Tischendorf, H.-G. Brachtendorf, K. Bittner, W. Schoenmaker, P. Benner, L. Feng, R. Pulch, F. Deleu, A. Wieers* *1029*

Fluid Structure Interaction

A Comparison of conforming and no-conforming mesh methods for a flow around a circular cylinder in non-inertial frame of references
S.H. Madani, J. Wissink, H. Bahai* *1037*

A New very large eddy simulation model in the context of fluid-structure interaction
A. Kondratyuk, M. Schäfer* *1047*

A numerical investigation of the added mass effect due to fluid-structure coupling in a rectangular tank
A. Brandely, E. Lefrancois* *1058*

Aerodynamic excitation forces in air conditioners with rotating fan-motor system
T. Sato, S. Nita, H. Ota, K. Nagahashi* *1072*

Comparison of coupled euler-lagrange and smoothed particle hydrodynamics in fluid-structure interaction
C. Zehetner, M. Schörgenhumer, F. Hammelmüller, A. Humer* *1080*

Efficient FSI codes coupling with possible large added mass effects: Applications to rigid and elongated flexible bodies in the maritime field
A. Leroyer, C. Yvin, E. Guilmineau, M. Visonneau, P. Queutey* *1089*

Flow passive control on the NACA airfoils. Experimental and numerical study
*H. Tebbiche, M. S. Boutoudj** *1101*

Higher-order FEM for nonlinear hydroelastic analysis of a floating elastic strip in shallow-water conditions
A.E. Karperaki, K.A. Belibassakis, T.K. Papathanasiou, S.I. Markolefas* *1110*

Hybrid analytical-numerical methodology for computationally efficient pre-design analysis of fluid-estructure interaction
*J. Rubio, S. Iakovlev, P. Diez** *1023*

Immersed boundary method for the simulation of lid-driven cavity flow with an embedded cylinder
S-G. Cai, A. Ouahsine, J. Favier, Y. Hoarau, H. Smaoui* *1130*

Numerical modelling of fluid-structure interactions
A. Fortin, A. Jendoubi, J. Deteix* *1138*

Numerical prediction of ship resistance and squat in confined waters
F. Linde, A. Ouahsine, N. Huybrechts, P. Sergent* *1147*

Numerical simulation of flow interaction between stationary and downstream elastically mounted cylinders at low Reynolds numbers
P.R.F. Teixeira, E. Didier* *1159*

Numerical simulation of tsunami force acting on a floating/submerged tsunami shelter
H. Mutsuda, S. Fujii, A. Ardianti, Y. Doi* *1171*

Three-node zero-thickness hydro-mechanical interface finite element for geotechnical applications
B. Cerfontaine, A-C. Dieudonné, J-P. Radu, F. Collin, R. Charlier* *1185*

Towards goal-oriented mesh adaptation for fluid-structure interaction
E. Gauci, F. Alauzet, A. Loseille, A. Dervieux* *1197*

Geomechanics

Application of higher-order elements for coupled analysis in geomechanics
M. Kardani, M. Nazem, J.P. Carter* *1209*

Coefficient of consolidation for soil – that elusive quantity
D. Wang, M.F. Randolph, S. Gourvenec* *1218*

Effective stress in unsaturated granular materials: Micro-mechanical insights
J. Duriez, R. Wan* *1232*

Seismic Response Analysis by Subloading Surface Model
K. Miyashita, K. Hashiguchi, S. Sato* *1243*

Thermo-poro-mechanical coupled processes during thermal pressurization around Nuclear Waste Reposity
M-N. Vu, D. Seyedi, G. Armand* *1251*

Numerical Methods for Coupled Problems

Analysis of time filters used with the leapfrog scheme
*Y. Li, C. Trenchea** *1261*

Mathematical modeling of reactionmechanism of formation of photochemical smog by applying the semi-implicit method
F.R. León, A.M. Aguilar, O.R. García, M.B. Pineda, J.G. Axotla* *1273*

Structural optimization of a wing body with uncertain aerodynamic loads
D.I. Papadimitriou, C. Papadimitriou* *1282*

Thermomechanics

An intercomponent heat transfer in a gravitational flow movements of particles in an inclined chute
I.N. Logachev, K.I. Logachev, O.A. Averkova* *1291*

Coupled thermo-aerodynamical problems in design of protection cloth
*I. Cherunova, M. Dhone, N. Kornev** *1303*

Response of RC columns with transient creep in a natural fire environment
A. Sadaoui, K. Lattari, A. Khennane* *1312*

Thermoelastic phenomena in the transversally graded laminates
J. Jędrysiak, E. Pazera* *1324*

AUTHORS INDEX 1336

INVITED SESSIONS

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis(Eds)

A CONSERVATIVE, OPTIMIZATION-BASED SEMI-LAGRANGIAN SPECTRAL ELEMENT METHOD FOR PASSIVE TRACER TRANSPORT

Pavel B. Bochev*, Scott A. Moe†, Kara J. Peterson* and Denis Ridzal**

*Computational Mathematics,
Sandia National Laboratories[1], Mail Stop 1320
Albuquerque, New Mexico, 87185-1320.

**Optimization and Uncertainty Quantification,
Sandia National Laboratories, Mail Stop 1320
Albuquerque, New Mexico, 87185-1320.

†University of Washington
Department of Applied Mathematics
Seattle, WA 98195

Key words: High order methods, transport, semi-Lagrangian methods, optimization

Abstract. We present a new optimization-based, conservative, and quasi-monotone method for passive tracer transport. The scheme combines high-order spectral element discretization in space with semi-Lagrangian time stepping. Solution of a singly linearly constrained quadratic program with simple bounds enforces conservation and physically motivated solution bounds. The scheme can handle efficiently a large number of passive tracers because the semi-Lagrangian time stepping only needs to evolve the grid points where the primitive variables are stored and allows for larger time steps than a conventional explicit spectral element method. Numerical examples show that the use of optimization to enforce physical properties does not affect significantly the spectral accuracy for smooth solutions. Performance studies reveal the benefits of high-order approximations, including for discontinuous solutions.

[1]Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

1 INTRODUCTION

In this paper we present a semi-Lagrangian spectral element method (SL-SEM) for the solution of the scalar transport equation for a positive density function ρ,

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\boldsymbol{v}\rho) = 0, \tag{1}$$

and the related equation

$$\frac{\partial \rho q}{\partial t} + \nabla \cdot \rho q \boldsymbol{v} = 0 \tag{2}$$

for transport of a passive tracer with mixing ratio q. Equations (1)–(2) imply that

$$\frac{\partial q}{\partial t} + \boldsymbol{v} \cdot \nabla q = 0\,, \tag{3}$$

that is, the passive tracer is constant along the characteristics. Although (3) is relatively simple, atmospheric models may involve large numbers of passive tracers, which makes their solution a major part of the computational cost.

The use of SEM for the spatial discretization of (3) offers important computational advantages such as a diagonal mass matrix and arbitrary order of accuracy [9]. However, in combination with an explicit time stepping scheme SEM, like other high-order methods, suffers from a severe, stability imposed, time step restriction.

On the other hand, schemes that adopt a Lagrangian viewpoint [3] and directly approximate the motion of "fluid particles" under a given velocity field have the potential to avoid restrictive CFL stability conditions. This makes such time stepping schemes particularly attractive for use with high-order spatial discretizations; see, e.g., [4, 11, 5] for some recent efforts to combine SEM with semi-Lagrangian time stepping.

An important complaint, though, about SL-SEM schemes is that they do not necessarily preserve physical properties such as conservation of total mass, or local solution bounds. Yet, such properties are often critical for accurate and physically consistent simulations of atmospheric models. This is especially true for schemes employing high-order spatial discretizations because in the presence of solution discontinuities such methods are prone to large unphysical oscillations, known as Runge or Gibbs phenomena.

In this work we combine the attractive traits of a *parent* SL-SEM with a novel, optimization-based strategy for the enforcement of the relevant physical properties [1]. Specifically, the raw high-order solution of the SL-SEM defines an *optimization* target, whereas mass conservation and physically motivated local solution bounds provide the optimization constraints. The actual solution is then determined by solving a singly linearly constrained quadratic program with simple bounds, which can be done very efficiently [2]. Similar ideas have been applied in an Eulerian spectral element method [6], however in the context of explicit time stepping. Our approach allows for the same high-order accuracy but avoids the stringent time stepping restriction experienced in that work.

2 PRELIMINARIES

2.1 Scalar advection

The basic problem considered in this work is the numerical solution of the multi-dimensional scalar advection equation (3). Atmospheric models require simultaneous advection of many tracers, which makes this problem a major source of computational cost. A semi-Lagrangian method would be ideal for this for two reasons. First it would allow much larger time steps than an Eulerian method. Second, to advect multiple tracer values located at a grid point, the method moves that point along a characteristic line and then computes the updated solution by interpolating the tracer values at the new point location. As a result, the cost to track the characteristics is amortized over a large number of passive tracer values.

This work focuses on divergence-free velocity fields. In this case equation (1) has the same form as (3), and $\int_\Omega q dV$ is a conserved quantity. Thus, we present the method for (3) alone. Formulations for general velocities will be considered in a forthcoming paper.

2.2 Spatial discretization

In this work we combine spatial discretization by spectral elements with semi-Lagrangian time stepping. For simplicity we describe the approach in two dimensions. Extension of the main ideas to three dimensions is straightforward. Thus, in what follows $K(\Omega)$ is a conforming partition of a bounded region $\Omega \in \mathbf{R}^2$ into quadrilateral cells κ_i, $i = 1, \ldots, K$.

In the spectral element method the basis functions on the reference quadrilateral $\hat{\kappa} = [-1, 1]^2$ are Lagrange polynomials corresponding to a tensor product grid of Gauss-Legendre-Lobatto (GLL) nodes. Specifically, let $\mathcal{N} = \{\xi_i\}_{i=1}^{r+1}$ denote a set of GLL nodes in $[-1, 1]$ and let

$$\hat{X} = \{(\xi_i, \xi_j) \,|\, \xi_i, \xi_j \in \mathcal{N}\}$$

be the corresponding GLL tensor product grid in $\hat{\kappa}$. The associated reference spectral element basis functions $\{\hat{\phi}_{ij}\}$ are rth degree Lagrange polynomials such that

$$\hat{\phi}_{ij}(\xi_l, \xi_k) = \delta_{li}\delta_{kj} \,. \tag{4}$$

Given an element $\kappa_m \in K(\Omega)$, let F_m be the isoparametric map defined by (4), which takes the reference element into κ_m, i.e., $F_m(\hat{\kappa}) = \kappa_m$. The image of $\hat{X}$ under F_m

$$X^m = \{\boldsymbol{x}_{ij} = F_m(\xi_i, \xi_j) \,|\, \xi_i, \xi_j \in \mathcal{N}\}$$

provides a GLL tensor product grid on κ_m. We define the spectral element basis $\{\phi_{kl}\}$ on $K(\Omega)$ by pullback, i.e., the local basis set $\{\phi_{ij}^m\}_{i=1}^{r+1}$ on an element κ_m comprises the functions

$$\phi_{ij}^m(\boldsymbol{x}) = \hat{\phi}_{ij} \circ F_m^{-1}(\boldsymbol{x}) \,.$$

The spectral element space $\mathbb{Q}^r = \text{span}\{\phi_{kl}\}$. Although the accuracy of the GLL nodes is slightly less than that of Gauss-Legendre nodes ($2r - 1$ vs. $2r + 1$ for $r + 1$ points), the use of the former brings about some very attractive computational properties.

- They display asymptotically optimal Lebesgue Constant growth [10].
- The associated GLL quadrature does not significantly degrade the accuracy of the discrete solution.
- The basis (4) is orthogonal with respect to the GLL points, which results in a diagonal mass matrix when using them as quadrature points for SEM.
- Inclusion of the endpoints of the interval allows to maintain inter-element continuity.

The ability to maintain inter-element continuity is important for the stability of the spectral element method. For instance, using Gauss-Legendre nodes to implement a semi-Lagrangian scheme may introduce large oscillations due to solution discontinuities across element interfaces. For brevity, we restrict attention to bi-cubic basis functions $\hat{\phi}_{ij}$. In this case the relevant set $\mathcal{N}$ contains the following four GLL nodes:

$$X = \left\{-1, -\sqrt{1/5}, \sqrt{1/5}, 1\right\} \tag{5}$$

2.3 A parent semi-Lagrangian spectral element method (SL-SEM)

Let $\tilde{\boldsymbol{x}}(t)$ be the solution of the characteristic equation

$$\frac{d\tilde{\boldsymbol{x}}}{dt} = \boldsymbol{v} \quad \text{and} \quad \tilde{\boldsymbol{x}}(0) = \tilde{\boldsymbol{x}}_0, \tag{6}$$

and let $q(\boldsymbol{x}, t)$ be a solution of (3). Then,

$$\frac{d}{dt} q(\tilde{\boldsymbol{x}}(t), t) = \frac{\partial q}{\partial t} + \frac{d\tilde{\boldsymbol{x}}}{dt} \cdot \nabla q = \frac{\partial q}{\partial t} + \boldsymbol{v} \cdot \nabla q = 0. \tag{7}$$

In other words, q is constant along the characteristics $\tilde{\boldsymbol{x}}(t)$. Thus, if q is known at $\tilde{\boldsymbol{x}}(0)$ its values along $\tilde{\boldsymbol{x}}(t)$ can be determined by tracking the characteristic line. This observation forms the basis of semi-Lagrangian methods.

2.3.1 Semi-Lagrangian time stepping scheme

The semi-Lagrangian approach solves equation (3) by breaking it up into a series of ODEs of the form (6) coupled with an interpolation. To describe the basic scheme let X be the union of all GLL points in the mesh $K(\Omega)$, i.e., $X = \cup_{\kappa_m} X^m$. Suppose that the values of q at the current time step t_n are known at all points in X. To find the approximation of q at the GLL points at $t_{n+1} = t_n + \Delta t$ we track these points back along the characteristic lines using X as initial data for (6). Succinctly, the two steps are

$$\begin{aligned} &\text{Trace back:} \quad \forall \boldsymbol{x}_{ij} \in X \ \text{ solve: } \ \frac{d\tilde{\boldsymbol{x}}_{ij}}{dt} = -\boldsymbol{v} \ \text{ in } \ [t_n, t_{n+1}] \quad \text{with} \quad \tilde{\boldsymbol{x}}_{ij}(t_n) = \boldsymbol{x}_{ij}, \\ &\text{Interpolate:} \quad q(\boldsymbol{x}_{ij}, t_{n+1}) := q(\tilde{\boldsymbol{x}}_{ij}(t_{n+1}), t_n). \end{aligned} \tag{8}$$

The point $\tilde{\boldsymbol{x}}_{ij}(t_{n+1})$ is generally referred to as the "trace back" of the GLL point $\boldsymbol{x}_{ij} \in X$. In a nutshell, the parent semi-Lagrangian scheme transports the function q back along the characteristics and then uses the spectral element basis to interpolate its value at $\tilde{\boldsymbol{x}}_{ij}(t_{n+1})$; see Figure 2.3.1.

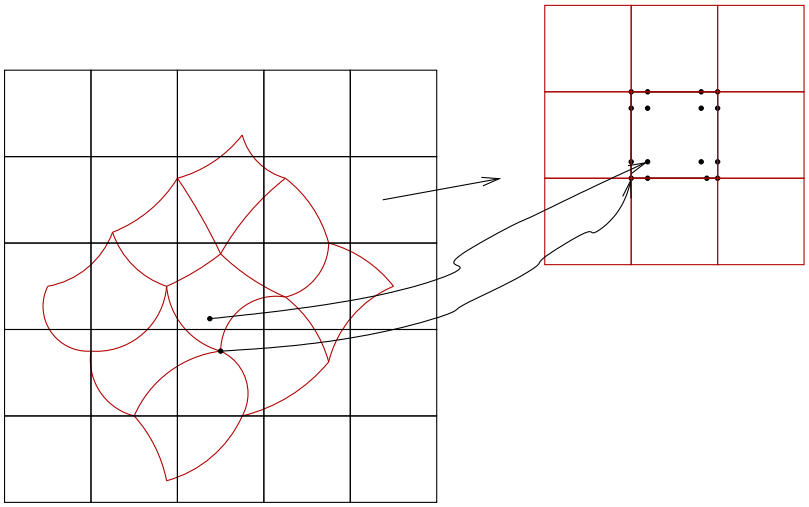


Figure 1: The SL-SEM traces back the GLL set along the characteristics $\tilde{\boldsymbol{x}}(t)$ and then uses the spectral basis to interpolate $q(\tilde{\boldsymbol{x}}(t_{n+1}), t_n)$. The interpolated values provide the approximation $q(\boldsymbol{x}_{ij}, t_{n+1})$ of q at the GLL points *at the next time step.*

Although formally of high-order, the parent SL-SEM does not preserve local solution bounds. A simple solid body rotation example illustrates this fact and motivates the need for an optimization-based solution to preserve relevant physical properties. Specifically, we apply the SL-SEM to advect a discontinuous profile given by the slotted Zalesak cylinder [7] using the rotational velocity field

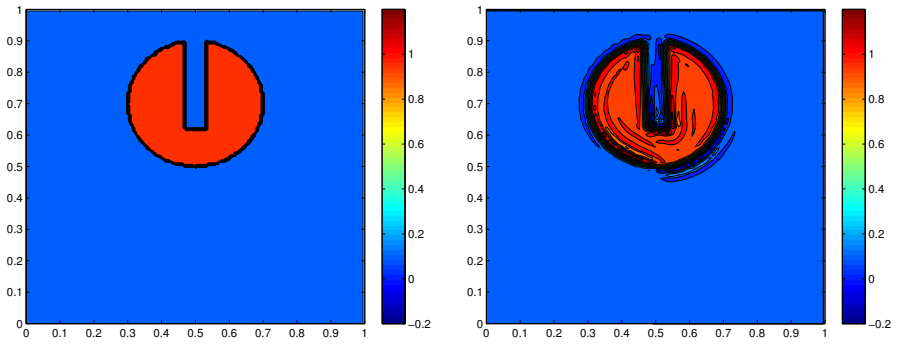
$$\boldsymbol{v} = \begin{bmatrix} (0.5 - y) & (0.5 - x) \end{bmatrix}^T. \tag{9}$$

The velocity field (9) rotates the initial profile around the center of the unit square with period 2π. There is no deformation and the velocity field is divergence free.

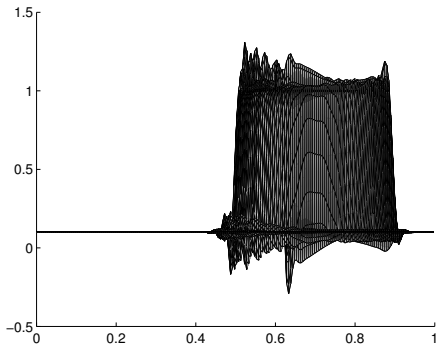
Figure 2(b) compares the initial profile with the SL-SEM solution after one full revolution. The exact solution satisfies global solution bounds and should remain between 0 and 1. The side view in Figure 2(c) clearly shows that the numerical solution develops spurious oscillations and significantly violates the global solution bounds.

3 OPTIMIZATION-BASED SL-SEM

In this section we combine the SL-SEM with an optimization-based approach to enforce conservation and local solution bounds. We note that in Eulerian methods preservation of



(a) $t = 0.0$ (b) $t = 2\pi$



(c) $t = 2\pi$ side view

Figure 2: SL-SEM solution of the solid body rotation problem with a slotted cylinder initial condition on an 80×80 uniform grid with $CFL = 0.7$. (a) Initial profile; (b) solution after one revolution; (c) side view after one revolution. The discontinuous initial profile results in spurious oscillations.

local bounds is accomplished through the use of limiters to obtain monotone reconstructions of the primitive variables. This prevents the numerical solution from developing new extrema.

The theory of monotonicity preserving limiters is well understood and developed for low-order finite volume methods [8]. However, because limiters modify the reconstruction process their application to high-order methods may degrade the accuracy. In fact, for such methods it is not clear how to preserve monotonicity without harming accuracy in smooth regions, or even if that is possible.

In contrast, following the ideas of [1] we separate reconstruction from the enforcement of physical properties such as solution bounds and conservation. To this end, we treat the solution of the parent SL-SEM as an optimization target, whereas the local solution bounds and mass conservation define the optimization constraints. The goal is to find the spectral element function that is the closest to the target field and simultaneously satisfies the constraints.

In particular, one can enforce bounds on the interpolation operator itself, giving completely decoupled bounds for each individual degree of freedom. Because the type of lower

and upper bounds introduced are not guaranteed to enforce exact monotonicity, we refer to this type of methods as *quasi-monotone* [6].

Optimization-based SL-SEM. Given SEM approximation $q_n(\boldsymbol{x}) \approx q(\boldsymbol{x}, t_n)$ the following algorithm computes the optimization-based SL-SEM solution $q_{n+1}(\boldsymbol{x}) \approx q(\boldsymbol{x}, t_{n+1})$ at the next time step:

1. Trace back: $\forall \boldsymbol{x}_{ij} \in X$ solve: $\dfrac{d\tilde{\boldsymbol{x}}_{ij}}{dt} = -\boldsymbol{v}$ in $[t_n, t_{n+1}]$ with $\tilde{\boldsymbol{x}}_{ij}(t_n) = \boldsymbol{x}_{ij}$.
2. Interpolate and set target: $\tilde{q}_{n+1}(\boldsymbol{x}_{ij}) := q_n(\tilde{\boldsymbol{x}}_{ij}(t_{n+1}))$.
3. Determine lower and upper bounds $\underline{q}_{ij}$ and $\overline{q}_{ij}$, respectively for $q_{n+1}(\boldsymbol{x}_{ij})$.
4. Determine $q_{n+1}(\boldsymbol{x})$ by solving the optimization problem:

$$q_{n+1} = \arg\min_{q \in \mathbb{Q}^r} \|q - \tilde{q}_{n+1}\|_0^2 \quad \text{subject to} \begin{cases} \int_\Omega q d\Omega = \int_\Omega q_n d\Omega & \text{(Conservation)} \\ \underline{q}_{ij} \le q(\boldsymbol{x}_{ij}) \le \overline{q}_{ij} & \text{(Local bounds)} \end{cases} \tag{10}$$

3.1 Determination of local solution bounds

Because the solution is constant along the characteristic lines, we can determine $\underline{q}_{ij}$ and $\overline{q}_{ij}$ by examining solution values in a neighborhood of the trace back point $\tilde{\boldsymbol{x}}_{ij}(t_{n+1})$. In this paper we adopt an approach where these bounds are set by computing the minimum and maximum solution values at the GLL points in the neighborhood of the element containing $\tilde{\boldsymbol{x}}_{ij}(t_{n+1})$. Given $\kappa_i \in K(\Omega)$ let $B(\kappa_i)$ be the set of all its neighbors and κ_i itself.

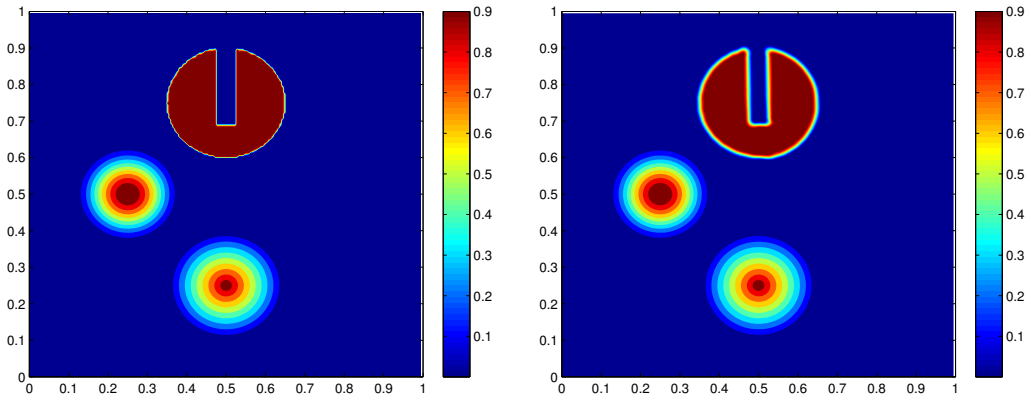
Local solution bounds. Given a trace back point $\tilde{\boldsymbol{x}}_{ij}(t_{n+1})$ the following procedure determines local solution bounds for the optimization-based SL-SEM:

1. Find element $\kappa_m \in K(\Omega)$ such that $\tilde{\boldsymbol{x}}_{ij}(t_{n+1}) \in \kappa_m$.
2. Define $\chi_m = \{\boldsymbol{x}_{kl} | \boldsymbol{x}_{kl} \in X \text{ and } \boldsymbol{x}_{kl} \in B(\kappa_m)\}$.
3. Set bounds $\underline{q}_{ij} = \min_{\boldsymbol{x} \in \chi_m} q_n(\boldsymbol{x})$ and $\overline{q}_{ij} = \max_{\boldsymbol{x} \in \chi_m} q_n(\boldsymbol{x})$.

3.2 Algebraic form of the optimization problem

For notational simplicity we drop the time step index from the target and simply write $\tilde{q}$. The coefficient vectors of $\tilde{q}$ and the state q are denoted by $\tilde{\mathbf{q}}$ and $\mathbf{q}$, respectively. Expanding the objective yields

$$\|q - \tilde{q}\|_0^2 = \sum_{i,j=1}^{r+1} \sum_{k,l=1}^{r+1} (q_{ij} q_{kl} - 2 q_{ij} \tilde{q}_{kl} + \tilde{q}_{ij} \tilde{q}_{kl}) \int_\Omega \phi_{ij} \phi_{kl} d\Omega, \tag{11}$$



(a) $t = 0.0$ (b) $t = 2\pi$

Figure 3: Optimization-based SL-SEM solution of the solid body rotation problem with initial profiles of varying degrees of smoothness; 100×100 uniform grid, bi-cubic elements, and CFL=0.7. (a) Initial time; (b) Solution after one full revolution.

where ϕ_{ij} are the spectral basis functions corresponding to the GLL nodes in X. Thus,

$$\|q - \tilde{q}\|_2^2 = \mathbf{q}^T M \mathbf{q} + \mathbf{c}^T \mathbf{q} + \mathbf{c}_0; \quad \mathbf{c} = -2M\tilde{\mathbf{q}}; \quad \mathbf{c}_0 = \tilde{\mathbf{q}}^T M \tilde{\mathbf{q}} \quad \text{and} \quad M_{ij,kl} = \int_\Omega \phi_{ij}\phi_{kl} d\Omega.$$

The SEM approximates the integrals above by using the GLL points in X along with suitable weights $\{w_{ij}\}$. Let $\mathbf{w}$ be the vector of these weights. Since the SEM basis functions are orthogonal with respect to the GLL nodes it follows that $M = diag(w_{ij}) = \mathbf{w}$. As a result, the SL-SEM optimization problem (10) assumes the following simple algebraic form

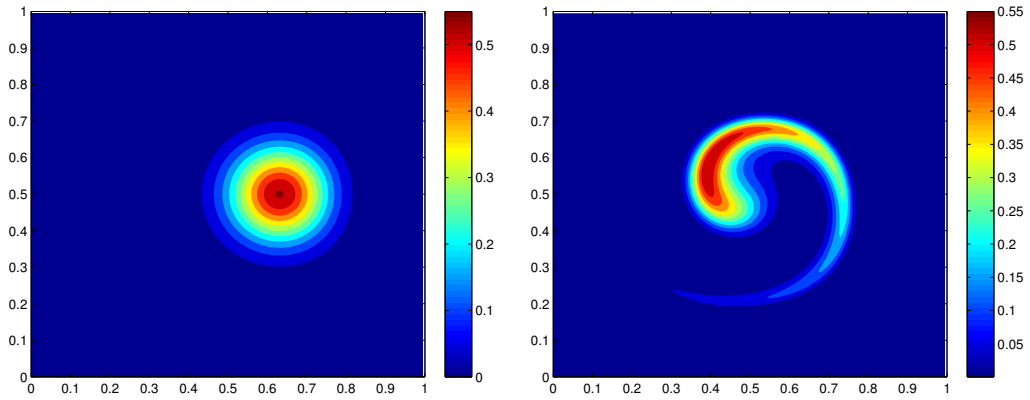
$$\mathbf{q}_{n+1} = \arg\min_{\mathbf{q}} \mathbf{q}^T M \mathbf{q} + \mathbf{c}^T \mathbf{q} + \mathbf{c}_0 \quad \text{subject to} \quad \begin{cases} \mathbf{w}^T \mathbf{q} = \mathbf{w}^T \mathbf{q}_n & \text{(Conservation)} \\ \underline{\mathbf{q}} \le \mathbf{q} \le \overline{\mathbf{q}} & \text{(Local bounds)} \end{cases} \tag{12}$$

Problem (12) is a singly linearly constrained quadratic program (QP) with simple bounds. The structure of this QP lends itself to an extremely efficient solution method; see [2].

4 NUMERICAL EXAMPLES

4.1 Solid body rotation test

To test the optimization-based SL-SEM we combine the rotational velocity field from Section 2.3.1 with an initial profile comprising a notched cylinder, a smooth hump and a cone. This example is a standard advection test introduced in [7]. Figure 3 shows the initial profile and the numerical solution after one full revolution. The plot in Fig.3(b) reveals minimal smearing of the discontinuous profiles and an essentially monotone solution.



(a) $t = 0.0$ (b) $t = 2.5$

Figure 4: Optimization-based SL-SEM solution of the modified deformational flow example using 80×80 uniform grid and bi-cubic basis functions. (a) Initial profile; (b) Solution at maximum deformation time.

4.2 Modified deformational flow test

This example uses a modified version of the deformational flow test on the sphere, which is standard in the climate modeling community [6]. The goal is to examine the convergence rates of the optimization-based SL-SEM. Specifically, we combine the divergence-free velocity field $\boldsymbol{v} = \begin{bmatrix} u & v \end{bmatrix}^T$ where

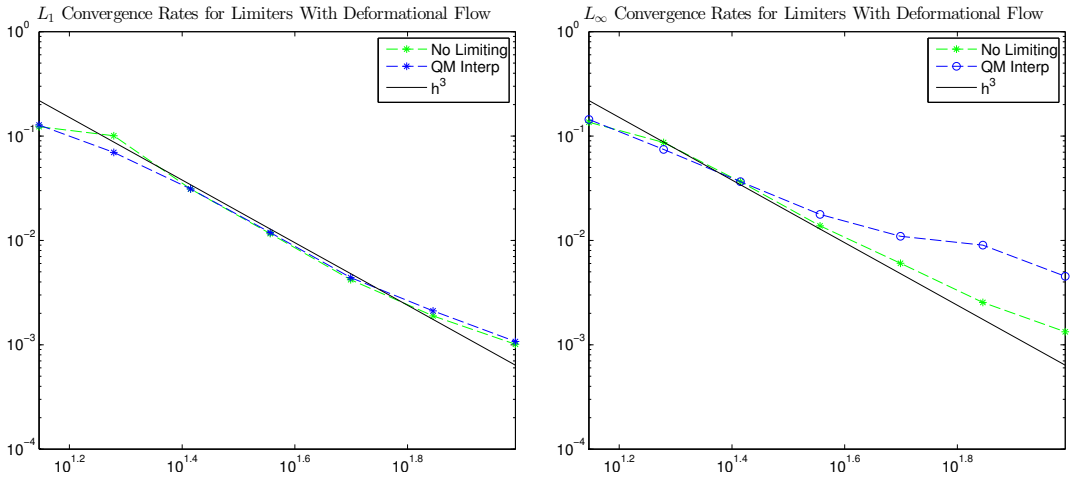
$$\begin{aligned} u &= \sin(\pi x)\sin(\pi x)\sin(2\pi y)\cos\left(\pi\frac{t}{T}\right) \\ v &= -\sin(\pi y)^2\sin(2\pi x)\cos\left(\pi\frac{t}{T}\right) \end{aligned} \tag{13}$$

with a Gaussian initial profile defined on $\Omega = [0,1] \times [0,1]$ by

$$q(x_1, x_2) = \sin(2\pi x)^4 \sin(2\pi y)^4 \exp\left(-\beta((x_1 - x_0)^2 + (x_2 - y_0)^2)\right) \tag{14}$$

where $\beta = -40.0$ and $(x_0, y_0) = (0.7, 0.5)$. The velocity field is designed so that after time $t = 2.5$ it reverses and the profile returns to the initial condition at 5.0. The initial profile is infinitely smooth except for the boundary where it is scaled so that it is zero on $\partial\Omega$. This allows us to impose periodic boundary conditions.

Figure 4 shows the initial profile and the solution at the maximum deformation instant $t = 2.5$. Figure 5 shows the convergence results for this example. We observe a slight degradation in the accuracy of the L_∞ norm errors, while the L_1-norm errors are not significantly affected. This is a very good result for a method that also eliminates oscillations. A method that is exactly monotonicity preserving would typically truncate the convergence rate to order 2 even in the L_1 norm.



(a) The L_1 norm convergence rate (b) The L_∞ norm convergence rate

Figure 5: The effect of optimization-based enforcement of local bounds on convergence rates.

5 PERFORMANCE STUDIES

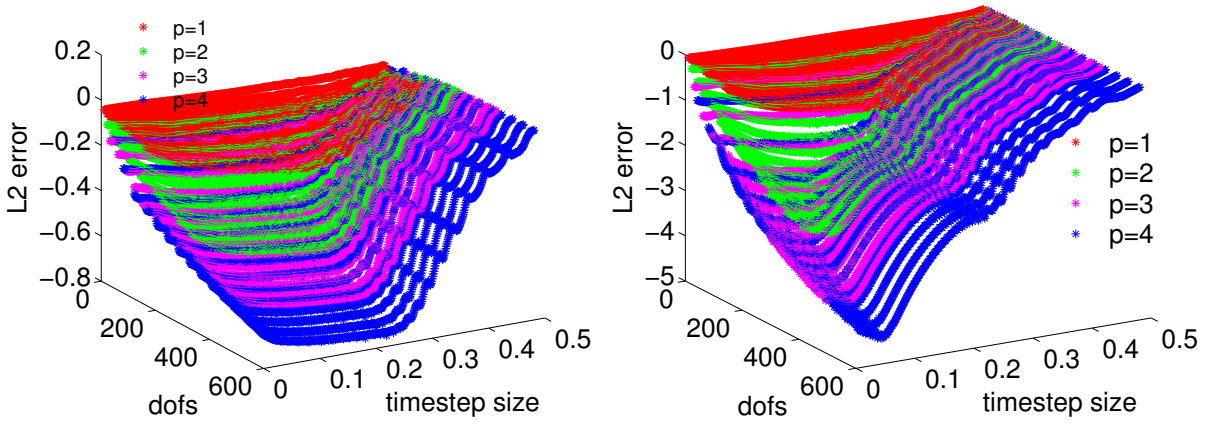
This section compares several realizations of the optimization-based SL-SEM with different orders of basis functions. The comparison is done using a Gaussian profile and the slotted Zalesak cylinder both advected by the deformational flow velocity field introduced in Section 4.2. In all cases we use fourth-order Runge Kutta for back-tracing the characteristics.

In this study the SL-SEM is run over a wide range of mesh and time step sizes using several different polynomial orders. The purpose of this study is to examine the accuracy of the method as a function of the spatial resolution, as measured by the number of spatial degrees of freedom and the temporal resolution, determined by a varying time step size. The accuracy is measured by the L_2-norm error. Figures 6(a) and 6(b) reveal an interesting behavior of the optimization-based SL-SEM. Unlike with Eulerian schemes, there is an optimal time step size. The reason for this is that every interpolation instance builds up some amount of error. Thus, there appears to be an optimal time step size which balances the accumulation of the errors with the available resolution. Also, as expected, increasing the number of spatial degrees of freedom tends to reduce errors.

All numerical computations in this section are done on a workstation with an Intel core i5-2500 processor and 8 Gigabytes of memory.

6 CONCLUSIONS

We presented an optimization-based semi-Lagrangian spectral element method for a scalar advection equation. The method combines a characteristic-based approach with optimization to enforce preservation of physical properties such as global conservation and local solution bounds. Convergence studies reveal minimal degradation of L_1-norm



(a) Discontinuous initial condition

(b) Gaussian initial condition

Figure 6: Parameter study of the approximation error in the optimization-based SL-SEM as function of the time step size and the number of spatial degrees-of-freedom for different polynomial degrees.

convergence rate for smooth solutions.

A parameter study was performed to characterize accuracy with respect to the number of degrees of freedom and the time step size. The study reveals the existence of an optimal time step for a given spatial resolution and indicates that higher polynomial degrees outperform lower degrees even on discontinuous profiles, which confirms the utility of the proposed approach in the context of high-order approximations.

Acknowledgements

This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research.

REFERENCES

[1] Pavel Bochev, Denis Ridzal, and Kara Peterson. Optimization-based remap and transport: A divide and conquer strategy for feature-preserving discretizations. *Journal of Computational Physics*, 257:1113–1139, 2014.

[2] Pavel Bochev, Denis Ridzal, and Mikhail Shashkov. Fast optimization-based conservative remap of scalar fields through aggregate mass transfer. *Journal of Computational Physics*, 246(0):37 – 57, 2013.

[3] Jim Douglas, Jr and Thomas F Russell. Numerical methods for convection-dominated diffusion problems based on combining the method of characteristics with finite element or finite difference procedures. *SIAM Journal on Numerical Analysis*, 19(5):871–885, 1982.

[4] Francis X Giraldo. The Lagrange–Galerkin spectral element method on unstructured quadrilateral grids. *Journal of Computational Physics*, 147(1):114–146, 1998.

[5] FX Giraldo, JB Perot, and PF Fischer. A spectral element semi-Lagrangian (sesl) method for the spherical shallow water equations. *Journal of Computational Physics*, 190(2):623–650, 2003.

[6] Oksana Guba, Mark Taylor, and Amik St-Cyr. Optimization-based limiters for the spectral element method. *Journal of Computational Physics*, 267:176–195, 2014.

[7] Randall J. LeVeque. High-resolution conservative algorithms for advection in incompressible flow. *SIAM Journal on Numerical Analysis*, 33(2):627–665, 1996.

[8] Randall J LeVeque. *Finite volume methods for hyperbolic problems*, volume 31. Cambridge university press, 2002.

[9] MA Taylor, J Edwards, S Thomas, and R Nair. A mass and energy conserving spectral element atmospheric dynamical core on the cubed-sphere grid. In *Journal of Physics: Conference Series*, volume 78, page 012074. IOP Publishing, 2007.

[10] Lloyd N Trefethen. *Approximation theory and approximation practice*. Siam, 2013.

[11] Dongbin Xiu and George Em Karniadakis. A semi-Lagrangian high-order method for Navier–Stokes equations. *Journal of computational physics*, 172(2):658–684, 2001.

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis(Eds)

ANALYSIS OF STRESS PARTITIONING IN BIPHASIC MIXTURES BASED ON A VARIATIONAL PURELY-MACROSCOPIC THEORY OF COMPRESSIBLE POROUS MEDIA: RECOVERY OF TERZAGHI'S LAW

ROBERTO SERPIERI*, FRANCESCO TRAVASCIO†, SHIHAB ASFOUR† AND LUCIANO ROSATI+

*Dipartimento di Ingegneria, Università degli Studi del Sannio
Piazza Roma, 21 - I, 82100, Benevento, Italy. e-mail: rserpier@unisannio.it

†Biomechanics Research Laboratory, Department of Industrial Engineering, University of Miami
1251 Memorial Drive, MCA268, Coral Gables, FL, 33146.
e-mail: f.travascio@miami.edu, sasfour@miami.edu

+Dipartimento di Strutture per l'Ingegneria e l'Architettura (DIST), Università di Napoli Federico II
Via Claudio, 21, 80125 Napoli, Italy. e-mail: rosati@unina.it

Key words: Variational poroelasticity, Effective stress, Terzaghi's law, Consolidation, Porous media

Abstract.

The mechanics of stress partitioning in two-phase porous media is predicted on the basis of a variational purely-macroscopic theory of porous media (VMTPM) with compressible constituents. Attention is focused on applications in which undrained flow (UF) conditions are relevant, e.g., consolidation of clay soils and fast deformations in cartilagineous tissues. In a study of the linearized version of VMTPM we have recently shown that, as UF conditions are approached (low permeability or fast loading), Terzaghi's effective stress law holds as a general property of rational continuum mechanics and is recovered as the characteristic stress partitioning law that a biphasic medium naturally complies with. The proof of this property is obtained under minimal constitutive hypotheses and no assumptions on internal microstructural features of a particular class of material. VMTPM predicts that such property is unrelated to compressibility moduli of phases and admits no deviations from Terzaghi's expression of effective stress, in contrast with most of the currently available poroelastic theoretical frameworks. This result is presently illustrated and discussed. Simulations of compressive consolidation tests are also presented; they are obtained via a combined analytical-numerical integration technique, based on the employment of Laplace transforms inverted numerically via de Hoog et al.'s algorithm. The computed solutions consistently describe a transition from drained to undrained flow which confirms that Terzaghi's law is recovered as the limit UF condition is approached and indicate a complex mechanical behavior.

1 INTRODUCTION

The analysis of the dynamic response of multiphase porous media in the limit of undrained flow conditions is relevant in all those applicative contexts where the nature of characteristic loadings, compared to the characteristic consolidation time of the medium, determines a flow regime close to the ideal conditions of complete prevention of fluid drainage. Canonical examples of applications in which this condition is met are the analysis of saturated clay soils subjected to seismic loading [1], or building-induced short time static loading, and the analysis of cartilaginous tissues subjected to physiological impulsive loading [2].

Among several approaches so far proposed for continuum modelling of multiphase porous media, variational approaches [3–5] provide a tool for addressing poroelastic multiphase problems by introducing the least possible number of mechanical assumptions and postulated balance laws.

We have recently proposed a least-action based macroscopic continuum description of two-phase poroelasticity to derive a general biphasic formulation at finite deformations based upon the inclusion of the intrinsic volume variations of the solid [6, 7] among the kinematic descriptors. This theory is shortly referred to as Variational Macroscopic Theory of Porous Media (VMTPM). In VMTPM no Lagrange multipliers are employed to obtain the macroscopic local balance equations. The main consequence is that work-association between stress and strain measures is naturally preserved, and the local macroscopic Euler-Lagrange momentum balance equations include a momentum balance associated with intrinsic volumetric strains.

For the subclass of undrained flow (UF) problems in which macroscopic fluid redistribution within the mixture is impeded and inertial forces can be neglected, the linearized version of VMTPM predicts that stress is partitioned in the two phases in strict compliance with Terzaghi's law. This relation is found to hold irrespective of thermodynamic constraints, constitutive or microstructural features of the medium, and independently from intrinsic compressibility properties of phases. Furthermore, such a property, provided UF conditions are met, admits no deviations from Terzaghi's expression of effective stress, in contrast with most of the currently available poroelastic theoretical frameworks.

The objective of this study is to complement the general law of stress partitioning in UF, derived in [8], with solutions of the general 1D consolidation problem. The objective is the analysis of the transition from a drained to an undrained behavior so as to investigate the behavior of systems which admit superposition of solutions, and for which solution techniques based on Laplace transforms can be exploited. Accordingly, the governing equations herein considered are those pertaining to the linearized VMTPM in which elastic stiffness moduli are introduced in a standard purely-variational form as the second derivatives of strain potentials. The constitutive parameters employed in the numerical simulations are selected in compliance with the bounds reported in [9, 10].

2 VMTPM IN QUASI STATIONARY FLOW

The purely mechanical linearized version of VMTPM is considered herein. Accordingly, all forces are represented by a potential so that Hamilton's principle, written for the solid porous phase, takes the form:

$$\delta \int_{t_1}^{t_2} \left(T^{(s)} - U^{(s)} \right) dt = 0. \tag{1}$$

In the previous relation $U^{(s)}$ is the total potential energy of the porous solid body, $T^{(s)}$ is the kinetic energy and t_1 and t_2 are two generic times.

In linearized kinematics, the infinitesimal deformation of the solid phase is defined, at the macroscale, by the solid macroscopic infinitesimal displacement field $\bar{\mathbf{u}}^{(s)}$, and by the macroscopic field of the intrinsic volumetric strain $\hat{e}^{(s)}$. The latter is related to the microscopic volumetric strain field of the pore-scale deformation, $e^{(s)}$, by the standard averaging relation:

$$\hat{e}^{(s)}\left(\boldsymbol{x}\right) = \frac{1}{V^{(s)}} \int_{\Omega^{(s)}_{RVE}(\boldsymbol{x})} e^{(s)}\, dV = \frac{dV^{(s)}}{V^{(s)}}. \tag{2}$$

where $V^{(s)}$ and $dV^{(s)}$ are the volumes of the subset of the representative volume element centered in $\boldsymbol{x}$, and its first-order variation, respectively. The addition of $\hat{e}^{(s)}$ as a kinematic descriptor beside displacements is the unique essential kinematic/contitutive enhancement introduced with respect to standard Cauchy continuum theory. The reader is referred to [6–8] for a more extensive description of the VMTPM.

Hereby, the kinematics and the linearized set of governing equations are briefly recalled. Specifically, employing the notation conventions and terminology used in [8], we recall from [7] the specialization of the linearized VMTPM, which stems by considering negligible inertia forces. This condition will be shortly termed henceforth Quasi Stationary Flow (QSF).

The extrinsic volumetric strain, $\bar{e}^{(s)}$, is ordinarily related to the macroscopic solid displacement field as the trace of the solid strain tensor:

$$\bar{e}^{(s)} = \mathrm{tr}\bar{\boldsymbol{\varepsilon}}^{(s)} = \nabla \cdot \bar{\mathbf{u}}^{(s)}, \tag{3}$$

where $\bar{\boldsymbol{\varepsilon}}^{(s)} = \mathrm{sym}\, \bar{\mathbf{u}}^{(s)}_{\nabla}$, being $\bar{\mathbf{u}}^{(s)}_{\nabla} = \bar{\mathbf{u}}^{(s)} \otimes \nabla$ the displacements gradient.

The governing equations of linearized VMTPM at QSF are synoptically recalled in Table 1 where the notation employed in [8, 10] is used. Equations (4) and (5) are, respectively, the so-called extrinsic momentum balance of the solid phase and the intrinsic momentum balances, both stemming from (1). These two equations express stationarity of the Action functional with respect to variations of the solid displacements and variations of $\hat{e}^{(s)}$, respectively. Equation (7) represents the saturation constraint in linearized kinematics (see [8], and Bedford and Drumheller [3]), written in dimensionless form for a biphasic compressible medium.

Accordingly, p is the fluid pressure while volumetric fractions $\phi_o^{(\alpha)}$ (with $\alpha = s,\ f$), due to the saturation hypothesis fulfill $\phi_o^{(f)} + \phi_o^{(s)} = 1$. Moreover, $\check{\boldsymbol{\sigma}}^{(s)}$ and $\hat{p}^{(s)}$, termed 'solid extrinsic stress tensor' and 'solid intrinsic scalar pressure', are the primary stress measures of the solid

Extrinsic linear momentum balance of solid phase

$$\nabla \cdot \check{\boldsymbol{\sigma}}^{(s)} + \boldsymbol{\pi}^{fs} = 0 \tag{4}$$

Intrinsic momentum balance of the solid phase

$$\hat{p}^{(s)} + \hat{p}^{(fs)} = 0 \tag{5}$$

Momentum balance of the fluid phase

$$-\phi_o^{(f)} \nabla p + \boldsymbol{\pi}^{sf} = 0 \tag{6}$$

Combined fluid mass balance and saturation constraint (Dimensionless saturation constraint)

$$\phi_o^{(f)} \bar{e}^{(f)} + \phi_o^{(s)} \bar{e}^{(s)} = \phi_o^{(f)} \hat{e}^{(f)} + \phi_o^{(s)} \hat{e}^{(s)} \tag{7}$$

Table 1: General set of equations that govern the mechanical behavior of the biphasic compressible poroelastic system in linearized VMTPM.

phase workassociated with $\bar{\boldsymbol{\varepsilon}}^{(s)}$ and $\hat{e}^{(s)}$, respectively while $\hat{p}^{(fs)} = -\frac{\partial \bar{\psi}^{(f)}}{\partial \hat{e}^{(s)}}$ where $\bar{\psi}^{(f)}$ denotes the strain energy density of the fluid phase. Mutual drag body volume forces acting over the solid and the fluid phase are respectively denoted by $\boldsymbol{\pi}^{fs}$ and $\boldsymbol{\pi}^{sf}$, with $\boldsymbol{\pi}^{fs} = -\boldsymbol{\pi}^{sf}$.

The linear theory in terms of elastic coefficients is derived introducing the elastic moduli in a standard form, as second order derivatives of strain potentials with respect to the primary strain measures [9, 10]. Compared to the general treatment, the QSF theory admits a simpler description in which only two additional moduli appear along with Lamé elastic moduli. These two moduli are a dimensionless coefficient $\bar{k}_r$ which is characteristic of VMTPM and couples extrinsic strains with intrinsic ones, plus an auxiliary intrinsic stiffness modulus, $\hat{k}_s$. They are defined as follows:

$$\bar{k}_r = \phi_o^{(s)} \frac{\partial^2 \bar{\psi}^{(s)}}{\partial \hat{e}^{(s)} \partial \bar{e}^{(s)}} \left(\frac{\partial^2 \bar{\psi}^{(s)}}{\partial \hat{e}^{(s)} \partial \hat{e}^{(s)}} \right)^{-1}, \qquad \hat{k}_s = \frac{1}{\phi_o^{(s)}} \frac{\partial^2 \bar{\psi}^{(s)}}{\partial \hat{e}^{(s)} \partial \hat{e}^{(s)}}. \tag{8}$$

Employing the coefficients defined in (8), a linear combination of the linear isotropic constitutive laws of the solid phase with the intrinsic momentum balance can be conveniently introduced in QSF, [10]:

$$\check{\boldsymbol{\sigma}}^{(s)} = 2\bar{\mu} \bar{\boldsymbol{\varepsilon}}^{(s)} + \bar{\lambda} \bar{e}^{(s)} \mathbf{I} - \bar{k}_r p \mathbf{I}, \qquad \frac{\phi_o^{(s)}}{\hat{k}_s} p = -\bar{k}_r \bar{e}^{(s)} - \phi_o^{(s)} \hat{e}^{(s)} \tag{9}$$

Moreover, in QSF a combination of governing balances (4)-(7) yields a u-p form of the governing equations in which the primary variables are the solid macroscopic displacements and the fluid pressure, [9, 10]. This form turns out to be more practical from an engineering computational point of view, although limited to QSF problems. Upon introducing a simple linear

Darcy law for the drag body volume forces $\boldsymbol{\pi}^{fs}$:

$$\boldsymbol{\pi}^{fs} = -\boldsymbol{\pi}^{sf} = K\left(\frac{\partial \bar{\mathbf{u}}^{(f)}}{\partial t} - \frac{\partial \bar{\mathbf{u}}^{(s)}}{\partial t}\right) \tag{10}$$

(with K measured in $\left[\frac{Ns}{m^4}\right]$) the u-p form reads:

$$\nabla \cdot \boldsymbol{\sigma}_D^{(s)} - \left(\phi_o^{(f)} + \bar{k}_r\right)\nabla p = 0 \tag{11}$$

$$-\frac{\left(\phi_o^{(f)}\right)^2}{K}\left(\nabla \cdot \nabla\right)p + \left(1 + \bar{k}_r\right)\frac{\partial}{\partial t}\nabla \cdot \bar{\mathbf{u}}^{(s)} + \frac{1}{\hat{k}_{sf}}\frac{\partial p}{\partial t} = 0 \tag{12}$$

where $\boldsymbol{\sigma}_D^{(s)} = \bar{\lambda}\bar{e}^{(s)}\mathbf{I} + 2\bar{\mu}\bar{\boldsymbol{\varepsilon}}^{(s)}$ is the *drained solid stress* and $\hat{k}_{sf} = \left(\frac{\phi_o^{(s)}}{\hat{k}_s} + \frac{\phi_o^{(f)}}{\hat{k}_f}\right)^{-1}$, being $\hat{k}_f$ the bulk modulus of the fluid phase.

3 STRESS PARTITITONING IN UNDRAINED FLOW

The UF condition refers to the flow of a biphasic medium contained in a closed region of space $\Omega_{bU} \subseteq \Omega_b$ such that macroscopic fluid mass flow across the boundary $\tilde{\mathcal{S}}$ of any closed region contained in Ω_{bU} is zero. Conditions of UF are met when 1) the characteristic consolidation time is much higher than the timescale of observation of the problem; 2) when drainage is prevented at the boundaries and/or in presence of special symmetry conditions (such as those typically encountered in jacketed tests). The characterization of UF conditions from a kinematic point of view corresponds to the existence of a unique macroscopic continuum field of undrained displacements $\bar{\mathbf{u}}^{(un)}$, common to both solid and fluid phases, defining the macroscopic displacements of the undrained medium:

$$\bar{\mathbf{u}}^{(un)}\left(\boldsymbol{x}\right) = \bar{\mathbf{u}}^{(f)}\left(\boldsymbol{x}\right) = \bar{\mathbf{u}}^{(s)}\left(\boldsymbol{x}\right), \quad \boldsymbol{x} \in \Omega_{bU}. \tag{13}$$

The dual static characterization of (13) has been determined, in linearized kinematics, in [8]. Upon deriving the characteristic forms, U_{bU}^{int}, U_{bU}^{ext}, achieved at UF by the internal and eternal energy potentials, the Principle of Virtual Deformations is applied in the framework of linearized VMTPM. The primary difference with respect to the derivation of the principle of virtual work for single phase continuum mechanics is that, in accordance with the presence in VMTPM of two macroscopic state fields $\bar{\mathbf{u}}^{(un)}$ and $\hat{e}^{(s)}$, two variational integral conditions (describing equilibrium in static UF configurations) are accordingly inferred. These equations are associated with virtual macroscopic displacements $\delta\bar{\mathbf{u}}^{(un)}$ and virtual intrinsic strain variations $\delta\hat{e}^{(s)}$:

$$\partial_{\left(\bar{\mathbf{u}}^{(un)},\,\hat{e}^{(s)}\right)}U_{bU}^{tot}\left[\delta\bar{\mathbf{u}}^{(un)},\,\delta\hat{e}^{(s)}\right] = 0, \qquad \forall\ \delta\bar{\mathbf{u}}^{(un)},\,\delta\hat{e}^{(s)}. \tag{14}$$

Derivation of the explicit strationarity integral equation stemming from (14) yields explicitly the following fundamental mechanical identification which holds within UF conditions:

$$\frac{\partial\bar{\psi}^{(f)}}{\partial\bar{e}^{(un)}} = -p. \tag{15}$$

As shown in [8], conversion from weak to strong form of the equilibrium equations yields a set of local equations which exactly match the classical tensorial statement of Terzaghi's principle. Specifically, these equations state that on the boundary of a region undergoing undrained flow, tractions, $\mathbf{t}^{(ext)}$, applied by the external environment are related to internal stress measures, $\check{\boldsymbol{\sigma}}^{(s)}$ and p, by

$$\mathbf{t}^{(ext)} = \left(\check{\boldsymbol{\sigma}}^{(s)} - p\mathbf{I}\right)\mathbf{n}, \tag{16}$$

where — it is worth being recalled — $\check{\boldsymbol{\sigma}}^{(s)}$ is the stress tensor naturally work-associated with isochoric (volume preserving) strains of the solid phase, consistently with its role of 'effective' stress. This relation holds irrespective of thermodynamic constraints, of the constitutive or microstructural features of the medium and independently from intrinsic compressibility of phases.

4 TRANSITION FROM DRAINED TO UNDRAINED FLOW

The transition from a more general condition of drained flow to UF conditions has been investigated, as predicted by VMTPM, under conditions of quasi stationary flow for a simple 1D consolidation problem. A semi analytical numerical solution method based on Laplace transforms was applied to solve the QSF problem of a generic biphasic specimen subjected to uni-axial confined compression in a stress-relaxation test. The examined problem is the following. A biphasic sample is laterally and inferiorly confined in an impermeable chamber, and compressed by a porous plug allowing for fluid exudation. The setup considered is schematized in Figure 1.

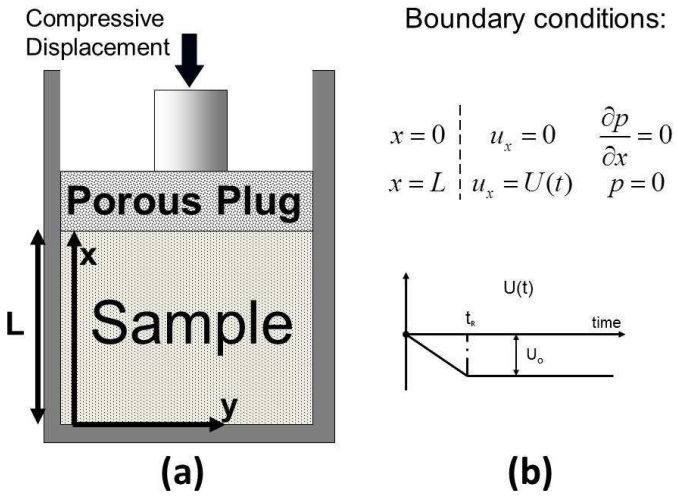


Figure 1: Simulated experimental setup for uni-axial confined compressive test: (a) biphasic mixture is confined in an impermeable chamber and axially compressed by a porous plug allowing for fluid exudation; (b) compressive displacement history applied to the plug for static tests (U) and associated boundary conditions.

A dimensionless treatment of this problem is exploited and tilde accents are used to indicate dimensionless quantities. Denoting by L and t_o, respectively, the characteristic sample length and time scale of observation, we introduce the following dimensionless quantities:

$$\tilde{p} = \frac{p}{(\bar{\lambda} + 2\bar{\mu})} \qquad \tilde{\sigma}^{(s)}_{Dxx} = \frac{\sigma^{(s)}_{Dxx}}{(\bar{\lambda} + 2\bar{\mu})}, \quad \tilde{u}^{(s)}_x = \frac{\bar{u}^{(s)}_x}{L}, \quad \tilde{x} = \frac{x}{L}, \quad \tilde{t} = \frac{t}{t_o} \tag{17}$$

so that the 1D counterpart of u-p equations (11), (12) can be recast in the form:

$$\frac{\partial^2 \tilde{u}^{(s)}_x}{\partial \tilde{x}^2}(\tilde{x}, \tilde{t}) - (\phi^{(f)}_o + \bar{k}_r)\frac{\partial \tilde{p}}{\partial \tilde{x}}(\tilde{x}, \tilde{t}) = 0 \tag{18}$$

$$(1 + \bar{k}_r)\frac{\partial^2 \tilde{u}^{(s)}_x}{\partial \tilde{t} \partial \tilde{x}}(\tilde{x}, \tilde{t}) + C_r \frac{\partial \tilde{p}}{\partial \tilde{t}}(\tilde{x}, \tilde{t}) - \frac{1}{De}\frac{\partial^2 \tilde{p}}{\partial \tilde{x}^2}(\tilde{x}, \tilde{t}) = 0 \tag{19}$$

Coefficients De and C_r are dimensionless characteristic parameters of the system, defined by:

$$De = \frac{\tau}{t_o}, \qquad \tau = \frac{L^2(\phi^{(f)}_o)^2}{K(\bar{\lambda} + 2\bar{\mu})}, \qquad C_r = \frac{\bar{\lambda} + 2\bar{\mu}}{\hat{k}_{sf}} \tag{20}$$

where τ is proportional to the consolidation time of the system.

The physical meaning of coefficient De is closely related to the Deborah number [13], since τ is a characteristic time proportional to the classic poromechanics definition of relaxation time (i.e. the time required for the biphasic medium to adapt to applied stresses or deformation). The range of this coefficient is $0 \leq De \leq \infty$. When $De >> 1$ the system tends to behave as an elastic completely undrained medium. The opposite condition of completely drained behavior corresponds to $De << 1$ while, for $De \simeq 1$, the fluid component of the mixture provides a significant contribution to time dependent stress/deformation phenomena.

The dimensionless quantity C_r represents the ratio of the aggregate modulus of the solid phase in the mixture (i.e., $\bar{\lambda}$ and $\bar{\mu}$) over the intrinsic stiffness ($\hat{k}_{sf}$) of the medium. Accordingly, when $C_r \to 0$, the mixture can be assumed to be intrinsically incompressible.

Space-time boundary conditions of the problem of Figure 1 are also directly examined in dimensionless form. At time $t = 0$ the mixture is in an undeformed state and in equilibrium with the environment so that the initial conditions are:

$$\tilde{u}^{(s)}_x = 0, \, \tilde{p} = 0 \qquad \tilde{t} = 0, \tilde{x} \in [0, \, 1] \tag{21}$$

At the bottom of the specimen (i.e., $\tilde{x} = 0$), solid displacement is null, and no fluid flows through the inferior wall of the chamber. This condition amounts to a null solid-fluid relative displacement and, recalling (6) and (10), implies a null pressure gradient. At the top ($\tilde{x} = 1$), fluid exudation occurs through the porous plug. In addition, for the case of displacement controlled compression, the displacement applied to the plug follows a ramp-and-hold time sequence, where the dimensionless ramp time is $\tilde{t}_R$, and the final compressive displacement is

$-U_o$, (with $U_o > 0$) see figure 1b. Based on these assumptions, the boundary conditions of the problem are collected below:

$$\tilde{u}_x^{(s)}(0,\tilde{t}) = 0, \quad \frac{\partial \tilde{p}}{\partial \tilde{x}}(0,\tilde{t}) = 0, \quad \tilde{u}_x^{(s)}(1,\tilde{t}) = \tilde{U}(\tilde{t}) = -\frac{U_o}{L}\left[\tilde{t} - H(\tilde{t}-\tilde{t}_R)\right], \quad \tilde{p}(1,\tilde{t}) = 0 \tag{22}$$

where H is the Heaviside function $\tilde{U} = U/L$ and $\tilde{t}_R = t_R/t_o$. The absolute value of the displacement applied to the porous plug, finally attained in the ramp-and-hold time sequence, is $U_o = 0.01\,L$ (i.e., 1 percent of the total length of the sample) while the ramp time is $\tilde{t}_R = 0.1$ (see figure 1b).

Solutions to the boundary value problem composed of equations (18), (19), (21), (22) are first determined in the complex Laplace space. Expressing (18), (19) in terms of Laplace transforms, denoted by star superscripts, yields:

$$\frac{\partial^2 (\tilde{u}_x^{(s)})^*}{\partial \tilde{x}^2}(\tilde{x},s) - (\phi_o^{(f)} + \bar{k}_r)\frac{\partial \tilde{p}^*}{\partial \tilde{x}}(\tilde{x},s) = 0 \tag{23}$$

$$s(1+\bar{k}_r)\frac{\partial (\tilde{u}_x^{(s)})^*}{\partial \tilde{x}}(\tilde{x},s) + sC_r\tilde{p}^*(\tilde{x},s) - \frac{1}{De}\frac{\partial^2 \tilde{p}^*}{\partial \tilde{x}^2}(\tilde{x},s) = 0 \tag{24}$$

Denoting by $\tilde{U}^*$ the Laplace transform of $\tilde{U}(\tilde{t})$, the transforms of the dimensionless solid displacement, of the solid displacements gradient and of the fluid pressure turn out to be:

$$(\tilde{u}_x^{(s)})^*(\tilde{x},s) = \tilde{U}^*\frac{\left[\beta\frac{1}{\alpha\sqrt{s}}\frac{\sinh(\alpha\sqrt{s}\tilde{x})}{\cosh(\alpha\sqrt{s})} + (1-\beta)\,\tilde{x}\right]}{\left[\beta\frac{1}{\alpha\sqrt{s}}\tanh(\alpha\sqrt{s}) + 1 - \beta\right]} \tag{25}$$

$$\frac{\partial \bar{u}_x^*}{\partial \tilde{x}}(\tilde{x},s) = \tilde{U}^*\frac{\left[\beta\frac{\cosh(\alpha\sqrt{s}\tilde{x})}{\cosh(\alpha\sqrt{s})} + 1 - \beta\right]}{\left[\beta\frac{1}{\alpha\sqrt{s}}\tanh(\alpha\sqrt{s}) + 1 - \beta\right]} \tag{26}$$

$$\bar{p}^*(\tilde{x},s) = \tilde{U}^*\frac{\left[\frac{\cosh(\alpha\sqrt{s}\tilde{x})}{\cosh(\alpha\sqrt{s})} - 1\right]}{\left[(\phi_o^{(f)} + \bar{k}_r)\frac{1}{\alpha\sqrt{s}}\tanh(\alpha\sqrt{s}) + \frac{C_r}{(1+\bar{k}_r)}\right]} \tag{27}$$

where parameters α, β are defined as follows:

$$\alpha = \sqrt{De\,(C_r+\gamma)}, \qquad \beta = \frac{\gamma}{C_r+\gamma}, \qquad \gamma = (\phi_o^{(f)} + \bar{k}_r)(1+\bar{k}_r). \tag{28}$$

Further mathematical details for obtaining the solution of equations (23) and (24) are reported in [11].

Except for special cases, Laplace anti-transforms of (25)-(27) cannot be performed analytically retaining a closed form. Thus, they were carried out by numerical computation via de Hoog et al's algorithm [15]. A special case is represented by the UF limit at $De \to \infty$. In such a limit the antitrasform of (27) to the time domain is easily obtained analytically. This solution

consistently corresponds to strain and pressure fields that are uniform in space. In particular the constant value attained by the pressure field, written in dimensional form, is:

$$p = -\left(1 + \bar{k}_r\right) \hat{k}_{sf} \frac{U_o(t)}{L} \tag{29}$$

and corresponds to the general undrained solution, compliant with Terzaghi's law (16), whose full expression is reported in [10].

For solutions outside of the UF limit, $De \neq \infty$, the observation of the system is carried out over a value of the dimensionless time equal to 1 and is obtained numerically.

Concerning the selection of the values of the dimensionless coefficients employed in the simulations ($\phi_o^{(f)}$, $\bar{\lambda} + 2\bar{\mu}$, $\bar{k}_r$, $\hat{k}_s$ for the solid, and $\hat{k}_f$ for the fluid phase), these were set so as to describe the behavior of a generic biphasic medium, without referring to a particular medium of a specific application. In absence of an experimental characterisation of these parameters, special attention has been payed in assigning to $\phi_o^{(f)}$, $\bar{\lambda} + 2\bar{\mu}$, $\bar{k}_r$, $\hat{k}_s$ consistent values preserving physical admissibility and meaningfulness. In this respect, it is important to observe that, while $\hat{k}_f$ is unrelated to the parameters pertaining to the solid phase ($\phi_o^{(f)}$, $\bar{\lambda} + 2\bar{\mu}$, $\bar{k}_r$, $\hat{k}_s$) the latter four parameters cannot be independently assigned since they all belong to the solid phase. To recognize this it can be easily verified that, if these values are improperly selected so that the condition $C_r + \gamma > 0$ is violated, the solution achieves a singularity with unbounded values for β and imaginary values for α.

It can be shown that, by considering standard Mori-Tanaka estimates for the macroscopic shear modulus and Hashin's Composites Spheres Assemblage (CSA) estimates for $\bar{k}_r$, $\hat{k}_s$ and $\bar{\lambda}$, the range of variation of $\bar{k}_r$ is $-1 \leq \bar{k}_r \leq 0$ [9, 10] and that the quantity $(C_r + \gamma)$ is strictly positive [11]. Conversely, as $\phi_o^{(f)}$ increases, function β turns from negative to positive values. In the example of Figure 3 below it is shown that this change of sign determines significant qualitative and quantitative differences in the strain profiles.

For $De = 1$ and different values of $\phi_o^{(f)}$, $\bar{k}_r$ and C_r (details on the employed data are reported in [11]) the shape of the curves of stress history at the plug is the same. Stress increases throughout the ramp phase of the displacement and reaches a peak value at $\tilde{t}_R$. Subsequently, while plug displacement is held fixed, stress relaxes to an equilibrium value corresponding to the condition in which no relative motion between solid and fluid phases occurs (see Figure 2).

To appreciate the role played by $\bar{k}_r$ in determining the character of the solution, sensitivity analyses of the space-time solutions have been carried out numerically [11] by employing dimensionless coefficients provided by CSA and Mori-Tanaka estimates [9–11]. Herein two solutions are presented from those considered in [11]; they have been obtained with the two sets of dimensionless parameters of Table 2, which in turn correspond to two different values of the porosity $\phi_o^{(f)} = 0.25$ (A) and $\phi_o^{(f)} = 0.5$ (B). Figure 3 shows, at several dimensionless times, the computed strain profiles. While the final stationary states of cases A and B are coincident, significant qualitative and quantitative differences are detected in the transient response. Details of these analyses and an extensive discussion of the related results are reported in [11].

Simulation	C_r	De	$\phi_o^{(f)}$	γ	$C_r + \gamma$	β	α
A	0.25	1	0.25	-0.08	0.17	-0.45	0.42
B	0.25	1	0.5	0.24	0.49	0.49	0.7

Table 2: Sets of dimensionless parameters employed in the analysis of transient strain profiles of Figure 3.

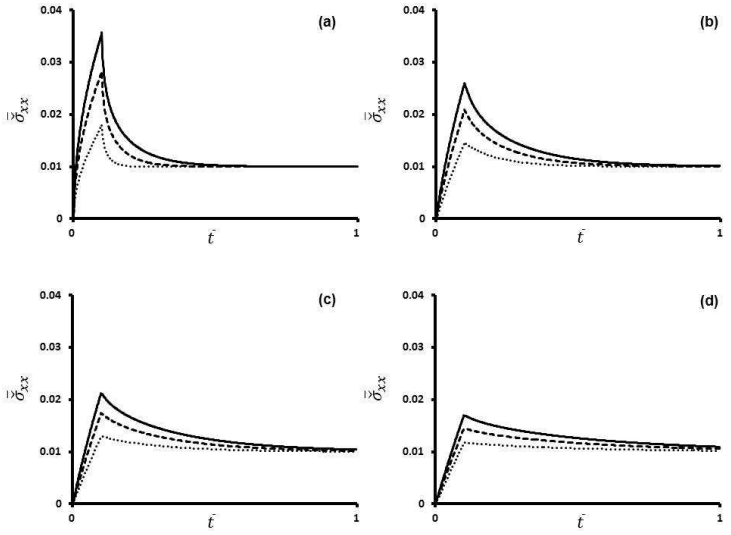


Figure 2: Apparent solid stress at plug for $De = 1$ and different $\phi_o^{(f)}$, $\bar{k}_r$ and C_r [11].

5 CONCLUSIONS

The analytical and numerical analyses developed in the present paper provide the following evidences.

- The computed solutions confirm that, in the transition from drained to undrained flow ($De \to \infty$), Terzaghi's law is recovered as the unique stress partitioning law as the limit UF condition is approached, in agreement with the general property shown in [8].
- Numerical results show that the ratio $\bar{k}_r$, which defines the elastic coupling between extrinsic strains and intrinsic ones, plays a significant role in determining the quantitative and qualitative character of the solution, in a way similar to the role played by Poisson ratio for single continuum mechanics. Since $\bar{k}_r$ is not affected by the intrinsic (grain) compressibility, and hence is a parameter unrelated to Biot' coefficient, this indicates that $\bar{k}_r$ should be carefully characterized in order to properly analyze the macroscopic mechanical response of a given biphasic medium.

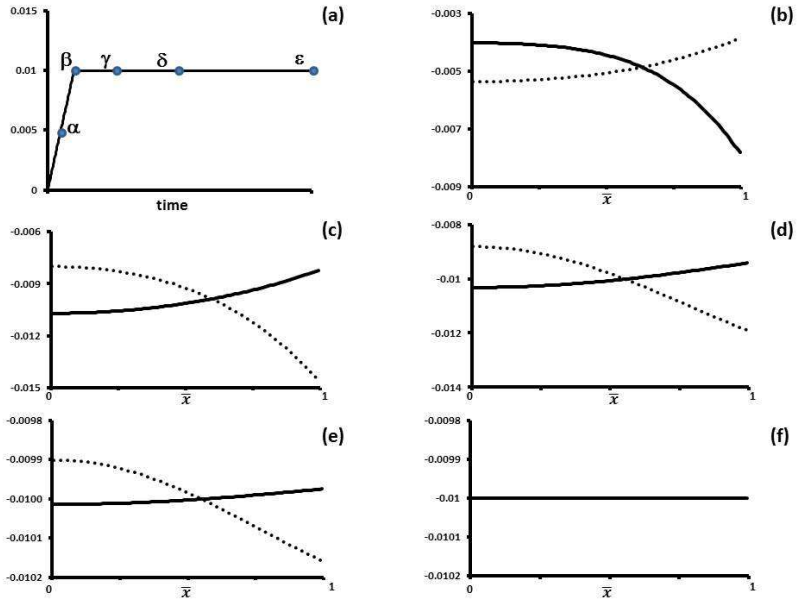


Figure 3: Transient strain profiles of the mixture at several dimensionless time frames: (a) $\tilde{U}(\tilde{t})$ plug displacement history with marks of time frames investigated; (b) strain profiles at α; (c) strain profiles at β; (d) strain profiles at γ; (e) strain profiles at δ; (f) strain profiles at ϵ. In all Figures (b)-(f): response for set A of parameters with $\phi_o^{(f)} = 0.25$ (solid line); response for set B of parameters with $\phi_o^{(f)} = 0.5$ (dotted line). Sets A and B are reported in Table 2.

REFERENCES

[1] Zienkiewicz, O. C., Chan, A. H. C., Pastor, M., Schrefler, B. A., and Shiomi, T. *Computational geomechanics*. Wiley, (1999).

[2] Hou, J.S., Holmes, M.H. Lai, W.M. and Mow, V.C. Boundary conditions at the cartilage-synovial fluid interface for joint lubrication and theoretical verifications *Journal of Biomechanical Engineering* (1989), **111**(1):78–87.

[3] Bedford, A. and Drumheller, D.S. A variational theory of porous media. (1979) *International Journal of Solids and Structures*, **15**(12):967–980.

[4] dell'Isola, F., Madeo, A. and Seppecher, P. Boundary conditions at fluid-permeable interfaces in porous media: A variational approach, *International Journal of Solids and Structures* (2009) **46**(17):3150–3164.

[5] Landau, .LD. and Lifshitz, E.M., *Mechanics: Volume 1 (Course of Theoretical Physics)*, Butterworth-Heinemann (1976).

[6] Serpieri, R. and Rosati, L., Formulation of a finite deformation model for the dynamic response of open cell biphasic media. *Jour. Mech. Phys. Sol.* (2011) **59**:841–862.

[7] Serpieri, R. A rational procedure for the experimental evaluation of the elastic coefficients in a linearized formulation of biphasic media with compressible constituents. *Transp. porous media* (2011) **90**:479–508.

[8] Serpieri, R., Travascio, F., Asfour, S. and Rosati, L. Variationally consistent derivation of the stress partitioning law in saturated porous media. *Intern. Jour. Solids Struct.* (2015) **56-57**:235–247.

[9] Serpieri, R., Travascio, F. and Asfour, S. Fundamental solutions for a coupled formulation of porous biphasic media with compressible solid and fluid phases., *V Conference on Computational Methods for Coupled Problems in Science and Engineering (COUPLED 2013)*, 1142-1153, Ibiza, Spain, 17-19 June, 2013.

[10] Serpieri, R., Travascio, F. General quantitative analysis of stress partitioning and boundary conditions in undrained biphasic porous media via a purely-macroscopic purely-variational approach. *Submitted*.

[11] Travascio, F., Serpieri, R., Asfour, S. and Rosati L. Semi analytical numerical solution of the general consolidation problem of compressible porous media by a purely-macroscopic variational continuum approach. *In preparation*.

[12] Terzaghi, K., The shearing resistance of saturated soils and the angle between the planes of shear, *International Conference on Soil Mechanics and Foundation Engineering, Cambridge (MA), USA*. (1936).

[13] Reiner, M. The Deborah number. (1964) *Physics Today*, **17**(1):62.

[14] Mow, V.C. Kuei, S.C., Lai, W.M. and Armstrong, C.G. Biphasic creep and stress relaxation of articular cartilage in compression: theory and experiments. (1980) *Journal of Biomechanical Engineering*, **102**(1):73–84.

[15] De Hoog, F.R., Knight, J.H., Stokes, A.N. An improved method for numerical inversion of Laplace transforms. *SIAM J. Sci. Stat. Comput.* (1982) **3**:357–366.

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis(Eds)

FINITE STRAINS FULLY COUPLED ANALYSIS OF A HORIZONTAL WELLBORE DRILLED THROUGH A POROUS ROCK FORMATION

N. Spiezia[†*], V. Salomoni[†] C. Majorana[†]

[†]Department of Civil, Environmental and Architectural Engineering (DICEA)
University of Padua
Via Marzolo 9, 35131 Padova, Italy
e-mail: nicolo.spiezia@dicea.unipd.it

Key words: Wellbore instability, solid-displacement/fluid-diffusion problem, finite strain elastoplasticity.

Abstract. Wellbore instability, in particular in deep perforations, continues to be one of the major problem in the oil and gas industry, that can dramatically increase production costs. Eventual instabilities may be prevented supporting temporarily the wellbore with mud circulation. If instability may occur, the value of the mud pressure needs to be sufficiently high to prevent compressional failure, but it should also be lower than a critical value that would cause tensile failure and unintentional hydraulic fracturing. Predicting faithfully the stress distribution around a borehole, and moreover the yielding and failure zones, is a challenging but fundamental task, essential to estimate the correct mud pressure and hence to prevent instabilities and sand production. This study focuses on quantifying the pressure distribution, stress field and plastic zones around a horizontal borehole drilled at great depth through a highly porous rock formation. The perforation of a wellbore in a saturated porous material is a coupled problem, which involves deformations of the solid phase and simultaneous diffusion of the fluid phase. A fully coupled finite element method is adopted, considering both material non linearity (elastoplasticity) and geometric nonlinearity (finite deformations) in the solid matrix, resulting in a so called $\boldsymbol{u}-p$ formulation. The variation of porosity and permeability, as consequence of the finite deformations of the solid matrix, is taken into account. The model adopts an elastoplastic constitutive law characterized by two yield surfaces, that is able to capture the dilatant and compactant plastic mechanism. The simulations investigate the quasi-static transient phenomenon associated with the perforation, until the steady state condition is reached. The model describes the evolution of the stress and pressure distribution, and moreover the propagation of the plastic zones around the borehole. The work demonstrates the capability of the finite deformations coupled approach to simulate the whole process, giving an instrument to determine the stability and sand production of the wellbore.

1 INTRODUCTION

Wellbore instability is still one of the major problem in the oil and gas industry, with disruptive consequence in the perforation and exploitation costs [1]. The instability of a wellbore is the effect of the failure of the rock surrounding the borehole, which may collapse under the new stress distribution due to the perforation.

In order to prevent instabilities during the drilling process, the wellbore is temporarily supported by the mud pressure [2]. The value of the mud pressure has to be properly determined: if the value is too low, compressional failure may occur around the wall of the hole. On the other hand, if the value of the mud pressure is too high, tensile failure may occur, causing unintentional hydraulic fracturing. Even if the risk of complete failure is avoided, the limitation of the failure zones is desirable, to reduce as much as possible the production of sand. For all these reasons, predicting faithfully the stress distribution around a borehole - and moreover the yielding and failure zones - is a fundamental task, essential to decide the most appropriate range of values for the mud pressure.

Traditionally, the stability of a well is determined using models based on linear elasticity, where failure is assumed to occur when the stress along the wall of the hole reaches the failure strength of the rock. An improvement in the assessment of the stability is given by numerical simulations, in particular with Finite Element Models. FEM analysis, endowed with elastoplastic constitutive law, can provide a more realistic evaluation of stability, with the advantage of being able to delineate the extent of the damaged region [3, 4, 5]. However, the complete drilling process can be properly simulated only if the interaction between the fluid phase and the solid phase is taken into account. Normally, these two fields are treated as separate issues and the tendency for each field is to simplify and make approximate assumptions for the other field. This is expected because of the complexity of treating geomechanics and multiphase fluid flow as coupled processes.

The aim of this study is to quantify the pressure and stress distribution and the propagation of plastic zones around a horizontal borehole drilled at great depth through a highly porous rock formation. The solid-displacements/fluid-diffusion problem is simulated with a fully coupled finite element method, resulting in a so called $\boldsymbol{u}-p$ formulation [7]. The model assumes finite strains formulation, which guarantees the correctness of the method, even in case of large deformation of the solid matrix. The variation of porosity and permeability, as consequence of the finite deformations of the solid matrix, is taken into account. The constitutive behavior of the porous rock formation is described by an elastoplastic cap model. In fact, highly porous rock are susceptible to different failure mechanisms, mainly due to shear-induced dilation and shear-enhanced compaction. Hence, the model is characterized by two yield surfaces, in order to capture the dilatant plastic mechanism and the compactant plastic mechanism.

The paper is organized as follow: section two briefly recall the finite strains coupled formulation for a fully saturated porous media. Section three deals with the constitutive laws adopted for the solid phase and the fluid phases. Section four presents the numerical

results of a wellbore considered ad example, drilled in an high porous rock formation, to illustrate the importance of coupling rock deformation and fluid flow. Finally, section five draws the conclusion of the presented work and gives an insight of the future developments of the research.

2 THEORETICAL FRAMEWORK

In this section we briefly recall the balance laws that govern the solid-displacement/fluid-diffusion problem in the fully saturated porous rock formation [7]. The fundamental idea consists in writing the mass balance and linear momentum equations for a two phase continuum body. We use mixture theory to formulate the kinematics of deformation of the solid matrix containing the fluid within its pores. We denote the motion of the solid matrix by $\varphi_s(\boldsymbol{X}_s, t)$ where $\boldsymbol{X}_s = \boldsymbol{X}$ is the position vector of the solid material point X in the reference configuration, the motion of fluid by $\varphi_f(\boldsymbol{X}_f, t)$, where $\boldsymbol{X}_f$ is the initial position vectors of fluid. Referring to the motion of the solid, we can write the material time derivatives following the fluid motion in terms of the material time derivative following the solid motion as

$$\frac{\mathrm{d}^f(\cdot)}{\mathrm{d}t} = \frac{\mathrm{d}(\cdot)}{\mathrm{d}t} + \nabla^{\boldsymbol{x}}(\cdot) \cdot \tilde{\boldsymbol{v}}, \tag{1}$$

where $\tilde{\boldsymbol{v}} := \boldsymbol{v}_f - \boldsymbol{v}$ is the velocity of the fluid with respect to the solid.
Let $\boldsymbol{x} = \varphi(\boldsymbol{X}, t)$ denote the position of the solid material point X and $\boldsymbol{F} = \partial \boldsymbol{x}/\partial \boldsymbol{X}$ the associated deformation gradient of the solid matrix, with the Jacobian J such that $J = \det(\boldsymbol{F}) = \mathrm{d}v/\mathrm{d}V$. We denote the volume fraction ϕ^α of the constituent α as the ratio between its volume $\mathrm{d}V^\alpha$ divided by the total volume of the mixture $\mathrm{d}V$, that is, $\phi^\alpha = \mathrm{d}V^\alpha/\mathrm{d}V$. Therefore,

$$\phi^s + \phi^f = 1 \tag{2}$$

The partial mass density of constituent α is given by $\rho^\alpha = \phi^\alpha \rho_\alpha$, where ρ_α is the intrinsic mass density of constituent α. This gives

$$\rho^s + \rho^f = \rho \tag{3}$$

where ρ is the total mass density of the mixture.
With these preliminaries, the conservation of mass equation for the solid and fluid phases are respectively

$$\dot{\rho}^s + \rho^s \nabla^{\boldsymbol{x}} \cdot \boldsymbol{v} = 0; \tag{4}$$

$$\dot{\rho}^f + \rho^f \nabla^{\boldsymbol{x}} \cdot \boldsymbol{v} = -\boldsymbol{\nabla}^{\boldsymbol{x}} \cdot \boldsymbol{q}, \tag{5}$$

where $\boldsymbol{q} \equiv \rho^f \tilde{\boldsymbol{v}}$ is the Eulerian relative flow vector of the fluid phase with respect to the solid matrix. Adding Eqs. (4) and (5), we get the basic conservation of mass equation for the system, i.e.

$$\dot{\rho}_0 = -J \nabla^{\boldsymbol{x}} \cdot \boldsymbol{q}, \tag{6}$$

where $\rho_0 \equiv J\rho$ is the pull-back mass density of the mixture in the reference configuration. In order to formulate the balance of linear momentum, a thermodynamically consistent effective stress equation may be written in terms of the total Cauchy stress tensor $\boldsymbol{\sigma}$, effective Cauchy stress tensor $\boldsymbol{\sigma}'$ and pore fluid pressure p as

$$\boldsymbol{\sigma} = \boldsymbol{\sigma}' - Bp\mathbf{1}, \tag{7}$$

where $B = 1 - K/K_s$ is the Biot coefficient and K is the bulk modulus of the solid matrix. To impose the balance of linear momentum with respect to the undeformed configuration, we need to introduce the first Piola-Kirchhoff stress tensor $\boldsymbol{P}$, which holds

$$\boldsymbol{P} = \boldsymbol{\tau} \cdot \boldsymbol{F}^{-T} = \boldsymbol{\tau}' \cdot \boldsymbol{F}^{-T} - B\theta_f \boldsymbol{F}^{-T} \tag{8}$$

where $\boldsymbol{\tau} = J\boldsymbol{\sigma}$ is the symmetric total Kirchhoff stress tensor, $\boldsymbol{\tau}' = J\boldsymbol{\sigma}'$ is the effective Kirchhoff stress tensor and $\theta_f = Jp$ is the Kirchhoff pore fluid pressure. Ignoring inertia forces, the balance of linear momentum in Lagrangian form can be written as

$$\nabla^{\boldsymbol{X}}(\boldsymbol{P}) + \rho_0 \boldsymbol{g} = \mathbf{0}. \tag{9}$$

The balance laws (6) and (9), in addition with the constitutive laws developed in the next section, provide a complete set of governing equations, which allow for the solution of quasi-static deformation-diffusion boundary-value problems. After developing the variational form, the problem results in a parabolic system where the displacements of the solid phase and the pore pressures are the basic unknowns in an updated Lagrangian finite element scheme.

3 CONSTITUTIVE MODELS

For the solid phase, we employ an elastoplastic model based on the multiplicative decomposition of deformation gradient and product formula algorithm described by Simo [9]. The elastic region is assumed to be governed by an isotropic strain-energy function as follow

$$\Psi = \frac{1}{2}\lambda[\epsilon_1^e + \epsilon_2^e + \epsilon_3^e]^2 + \mu[(\epsilon_1^e)^2 + (\epsilon_2^e)^2 + (\epsilon_3^e)^2], \tag{10}$$

where λ and μ are the classical Lamé constants and $\varepsilon_a^e = \ln \lambda_a^e$, with λ_a^e $(a = 1, 2, 3)$ the principal elastic stretches. The Kirchhoff effective stress tensor $\boldsymbol{\tau}'$ is coaxial with the left elastic Cauchy-Green deformation tensor $\boldsymbol{b}^e = \boldsymbol{F}^e \cdot \boldsymbol{F}^{eT}$ and defined such that

$$\boldsymbol{\tau}' = 2\frac{\partial \Psi}{\partial \boldsymbol{b}^e} \cdot \boldsymbol{b}^e. \tag{11}$$

The plastic region is defined in terms of stress invariants of the effective Kirchhoff stress tensor

$$P = \frac{1}{3}\mathrm{tr}(\boldsymbol{\tau}') \qquad Q = \sqrt{\frac{3}{2}}\,\|\boldsymbol{s}\|\,, \tag{12}$$

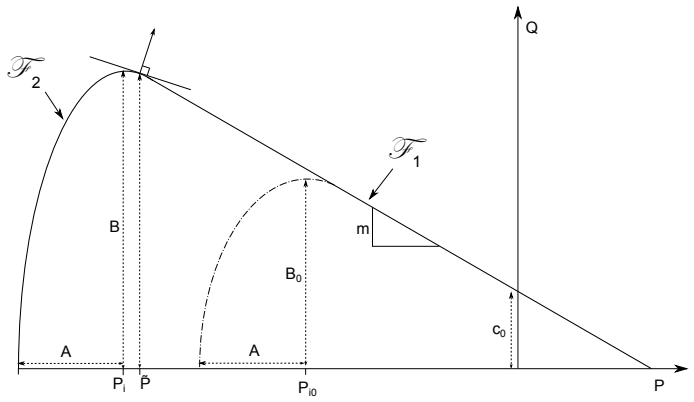


Figure 1: Schematic representation of the elastoplastic constitutive model for the high porous rock.

where $\boldsymbol{s} = \boldsymbol{\tau}' - P\mathbf{1}$. The yield function is characterized by two surfaces $\mathscr{F}_1$ and $\mathscr{F}_2$, that intersect smoothly: a linear surface to capture the dilatant shear failure and an elliptical surface to capture the compactant mechanical behavior of high porous rock. The yield functions, in the Kirchhoff stress invariants space, take the form

$$\mathscr{F}_1(P,Q) = Q - mP - c_0 = 0 \qquad \mathscr{F}_2(P,Q,P_i) = B^2(P-P_i)^2 + A^2Q^2 - A^2B^2 = 0 \tag{13}$$

where m is the slope of the linear yield surface, associated with the frictional angle, c_0 is the cohesion, A and B are respectively the minor and major semi-axis of the ellipse and P_i is the centroid. The intersection between the two surfaces is defined as the point in which $\mathscr{F}_1$ is tangent to $\mathscr{F}_2$, ensuring that the two surfaces produce a unique surface without any angular point. Hardening is admissible only along the compactant side, since the elliptical surface contracts and expands depending on the accumulated plastic volumetric strain ϵ_v^p, maintaining a constant width (i.e. $A = A_0$) [10] and remaining always tangent to the linear surface, as follow

$$P_i = P_{i_0}\left(\frac{\epsilon^*}{\epsilon^* - \epsilon_v^p}\right)^r \tag{14}$$

where P_{i_0} is the initial value, ϵ^* is the volumetric deformation at ultimate compaction and r is the exponent that control the rate of volumetric hardening [11]. A non associative flow rule is assumed only for the linear side, with a plastic potential of the form

$$\bar{\mathscr{G}} = Q - \bar{m}P - \bar{c} = 0. \tag{15}$$

This non-associative law, assuming that $\bar{m} < m$, avoid the often excessive dilatancy predicted by the associative rule. Fig. 1 summarizes all the relevant information of the elastoplastic model.

The constitutive relation for the fluid phase is given by the generalized Darcy's law, which takes the form

$$\phi^f \tilde{\boldsymbol{v}}_{\boldsymbol{f}} = -\boldsymbol{k} \cdot \boldsymbol{\nabla} \left(\frac{\theta_f}{\rho_f g} + z \right), \tag{16}$$

where g is the gravity acceleration constant and z is the elevation potential. We consider the following evolution of the hydraulic conductivity isotropic tensor, to express the change of permeability with deformation through the Jacobian J with the Kozeny-Carman equation

$$\boldsymbol{k} = k_s(J) = \frac{g\rho_f}{\mu} \frac{D^2}{180} \frac{(J - \phi_0^s)^3}{J(\phi_0^s)^2} \mathbf{1}, \tag{17}$$

with μ is the dynamic viscosity of the fluid, D is the diameter of the grains, and $\phi_0^s = 1 - \phi_0^f$.

4 NUMERICAL SIMULATIONS

The balance equations and the constitutive laws proposed have been implemented in a non linear finite element code, following standard procedures. To show the capability of the method, we investigate a horizontal wellbore drilled in Campos Basin field [12], a reservoir located 290 Km offshore Brazil cost. The parameters of numerical simulations are as follow: $E = 3500$ MPa, $\nu = 0.15$, $\phi_0^f = 0.3$, $m = -1.08$, $c_0 = 10$ MPa, $\bar{m} = -0.15$, $\mu = 0.1$ Kg/ms, $D = 0.01$ mm, $B = 0.6$. As far as the in-situ geostatic stresses in the reservoir production region, the total horizontal stresses are equal in both principal directions, and read $S_H = S_h = 41.4$ MPa. The total vertical stress is assumed to be $S_V = 64.5$ MPa. The pore fluid pressure in the reservoir is equal to $P_p = 32$ Mpa. The open-hole wellbore radius is $R = 107.95$ mm. The wellbore axis is parallel to the direction of the principal horizontal far-field stress S_h. Numerical analysis assumes a plane strain condition. Hence, the principal stress acting on the plane of the borehole section are the vertical and the horizontal stress.

The domain is discretized with quadrilateral elements, endowed with 9 nodes for displacements unknowns and 4 nodes for pressure unknowns, integrated over 9 Gauss point. The finite element geometry, with the boundary conditions, is represented in Fig. 2.

The analysis consists in two phases: in the first phase, the in-situ stresses and the reservoir pore pressure are applied to the complete domain - i.e. as the rock formation is still intact - in order to simulate the in-situ condition before the drilling process. In the second phase, in order to simulate the drilling process, the elements corresponding to the borehole are progressively removed from the domain, decreasing the stiffness and increasing the permeability. At the same time, the value of the mud pressure is applied to the nodes along the wall. This process is simulated instantaneously, assuming just for computational reason a period of time of 0.02 s.

The Fig. 3 represents the results in term of pore pressure (top row) and vertical displacements (bottom row) in case of balanced drilling ($\Delta P = 0$ MPa), i.e when the mud pressure equals the reservoir pore pressure. When the rock is drilled, the pressure distri-

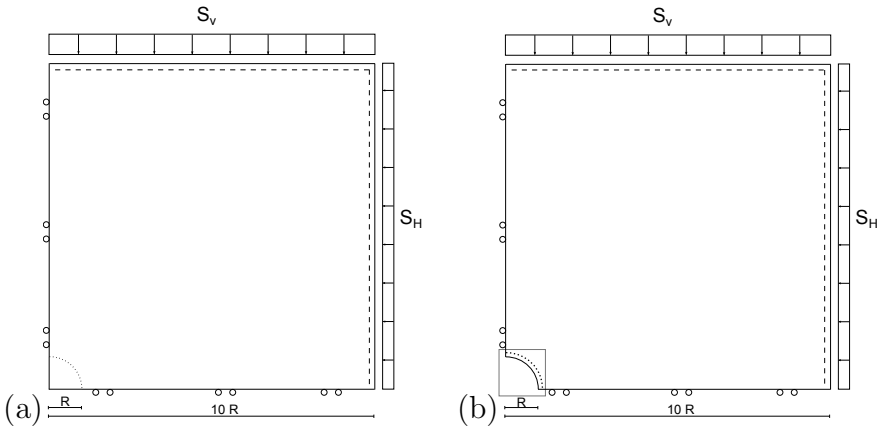


Figure 2: Plane strain domain of a quarter of the borehole. (a) In-situ configuration (no hole): far field stress and reservoir pressure applied (dashed line). (b) Drilling configuration (with hole): far field stress, reservoir pressure (dashed line) and mud pressure (dotted line) applied, with the hole. The box is zoomed in Fig. 5

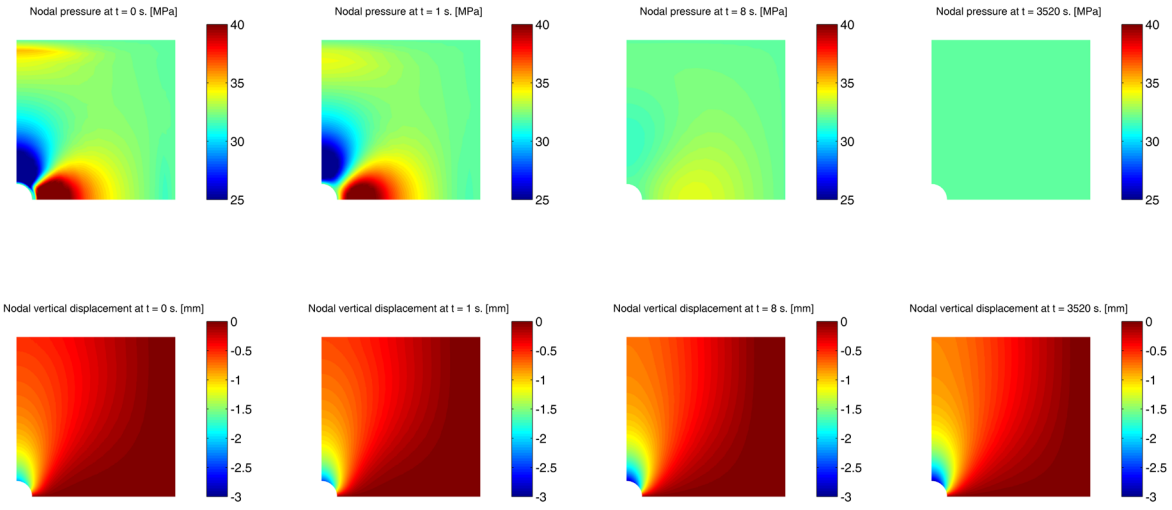


Figure 3: Wellbore drilling simulation with $\Delta P = 0$ MPa. Top images: nodal pore pressure distribution over time. Bottom images: vertical displacements over time.

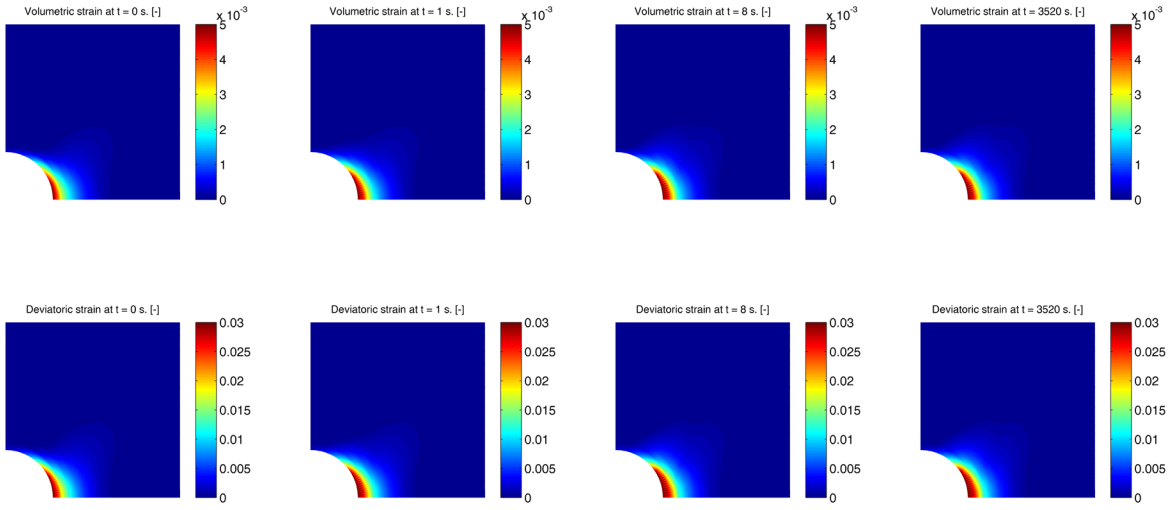


Figure 4: Wellbore drilling simulation with $\Delta P = 0$ MPa. Top images: Plastic volumetric strains over time. Bottom images: Plastic deviatoric strains over time.

bution immediately change significantly. An increment of pressure along the direction of the minimum stress is observed, associated with a decrease of pressure along the direction of maximum stress. This is due to the fact that initially the new load configuration in the rock formation is transferred only to the fluid. Then, the fluid diffuses through the rock until the considered domain reaches an uniform pressure distribution. The diffusion process is associated with the increase of vertical displacements (Fig. 3, bottom row), in close analogy with a consolidation process. The displacements are concentrated along the direction of the maximum stress, increasing in the area along the wall of the hole.
The Fig. 4 represents the results in terms of volumetric (top row) and deviatoric (bottom row) plastic strain. The fluid diffusion determines the expansion of the plastic zone along the wall of the wellbore, as consequence of the increase of the effective stress. The plastic zone is located along the direction of the minimum in-situ stress and involve a significant portion of the rock along the hole. This plastic zone is associated exclusively with the dilatant plastic mechanism (demonstrated by the positiveness of the volumetric plastic strain ϵ_v^p), which means that, for the considered stress configuration, the rock is not subjected to a compactant plastic mechanism.

Following results show the effect of the mud pressure on the drilling process. Two additional configurations are considered, namely with the mud pressure equal to $P_m = 30$ MPa (under-balanced drilling with $\Delta P = 2$ MPa) and the mud pressure equal to $P_m = 36$ MPa (over-balanced drilling with $\Delta P = 4$ MPa). The vertical displacement of point A (see Fig. 5) over time is represented in Fig. 6 in the three different configurations. The severe increase of displacements (about 2 mm) in the first instants is associated with

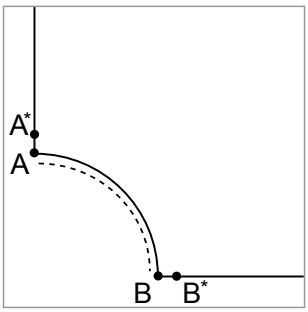


Figure 5: Detail of the domain around the hole. The point with star is located 18 mm far from the hole.

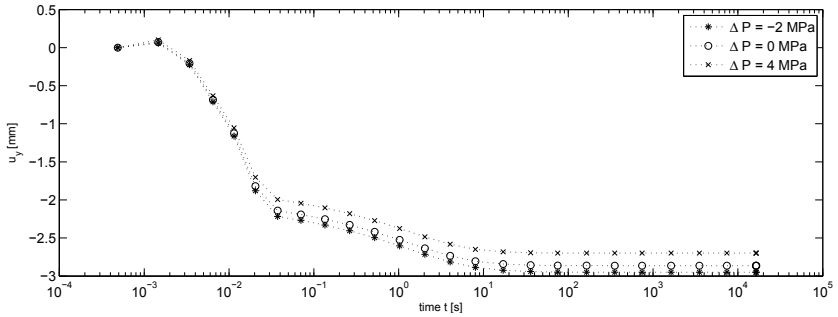


Figure 6: Vertical displacement of point A with different value of mud pressure.

the drilling phase (simulated instantaneously, assuming just for computational reason a period of time of 0.02 s). Then, there is a progressive increase of the vertical displacement due to the diffusion process of the fluid (about 1 mm). The incrementation of the mud pressure, as expected, reduces the vertical displacements, thanks to the contrast furnished to the settlements.

The Fig. 7 represents the evolution of the pore pressure in the point A* (i.e. along the direction of the maximum stress), for the three different values of mud pressure. As can be observed, there is an initial drop of the pressure, associated with the drilling process, until the pore pressure reaches the lowest value. Subsequently, the pore pressure increase progressively, due to the diffusion of the fluid, until the steady state condition, which depends on the mud pressure applied (respectively 30, 32 and 36 MPa).

The Fig. 8 represents the evolution of the pore pressure in the point B* (i.e. along the direction of the minimum stress), for the three different configurations. The evolution of the pore pressure is characterized by an initial increase, corresponding to the drilling phase. Then there is a progressive decrease of the pressure, until the steady state condition. It is interesting to observe that the path is not monotonic, resembling a behavior comparable with the so called Mandel-Cryer effect observed in the two-dimensional consolidation.

The Fig. 9 represents the evolution of the deviatoric plastic strain in the point B,

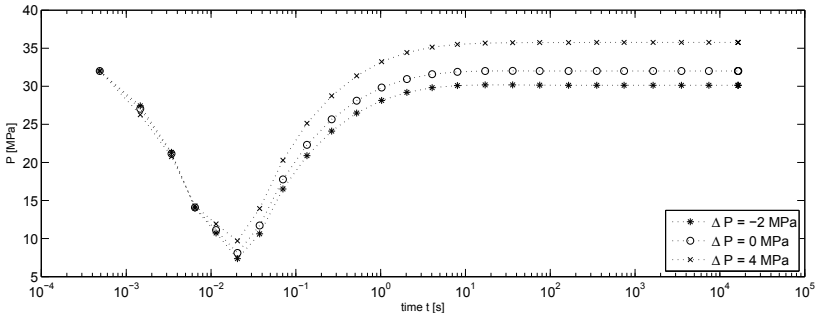


Figure 7: Pore pressure of point A* with different value of mud pressure.

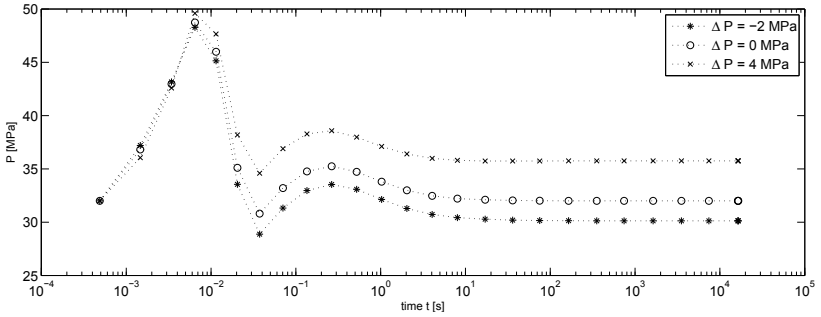


Figure 8: Pore pressure of point B* with different value of mud pressure.

extrapolated from the closest Gauss Point. The point has been selected since it is situated in the region where plastic deformations are localized. As can be observed, there is firstly an increase of plastic strain associated with the drilling process, and secondly a further increase due to the diffusion of the fluid, until the domain reaches the steady state condition. Depending on the applied mud pressure, the amount of the accumulated plastic strain changes accordingly: an increase of the mud pressure reduces the value of the plastic deviatoric strain, limiting the plastic zone. Since the plasticity is associated with the linear side of the yield surface, the evolution of the deviatoric and volumetric plastic strain are proportional, and therefore the volumetric strain follows the same path.

5 CONCLUSION

This work investigates the drilling process of a horizontal wellbore in a fully saturated rock formation, taking into account the coupling effects between the fluid and solid phases. Results of the analysis show the capability of the numerical simulations to describe the complete process. Fundamental is the detection of the plastic zones, which can be considered to assess the stability of the wellbore and the sand production.

Due to the impact that this topic has in the oil and gas industry, this subject may be extensively further developed. In particular, the analysis of different in-situ stress

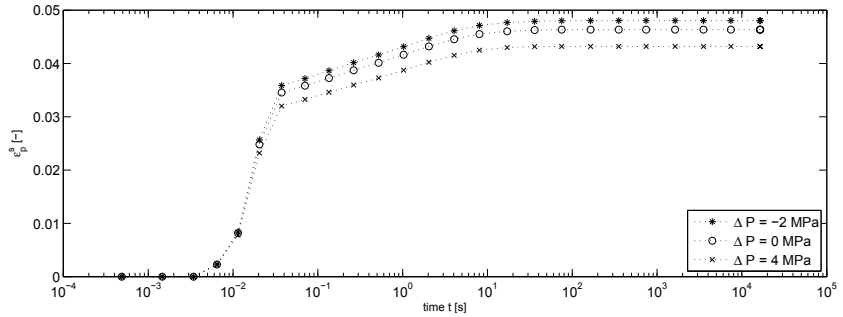


Figure 9: Deviatoric plastic strain of point B with different value of mud pressure.

and pressure configurations should be considered. The evolution of the porosity and permeability in the rock formation associated with the deformations should be assessed. Furthermore, this proposed formulation, based on finite strains assumptions, can be particularly suitable to analysis the eventual band localization around the wellbore, which can have dramatic consequences in the reservoir exploitation.

AKNOWLEDGMENTS

Financial support was provided to the first author by Gini Foundation, while part of this research was being conducted at Stanford University. Prof. R.I. Borja (Stanford University), provided the much appreciated motivation and inspirational support.

REFERENCES

[1] Moss, D., Peska, P., Finkbeiner, T., and Zoback, M., Comprehensive wellbore stability analysis using quantitative risk assessment, *Jour. Petrol. Sci. and Eng., Spec. Issue on Wellbore stability* (2003) **38**:97–109.

[2] Zoback, M. D., *Reservoir Geomechanics*, Cambridge University Press (2010).

[3] Chen, S. L., Abousleiman, Y. N. and Muraleetharan, Closed-form elastoplastic solution for wellbore problem in strain hardening/softening rock formations, *Int. J. Geomech.* (2012) **12**:494–507.

[4] Preisig M. and Prvost J.H., Fully coupled simulation of fluid injection into geomaterials with focus on nonlinear near-well behavior, *Int. J. Numer. Anal. Meth. Geomech.* (2012) **36**:1023–1040.

[5] Rutqvist, J., et al. Coupled multiphase fluid flow and wellbore stability analysis associated with gas production from oceanic hydrate-bearing sediments. *Journal of Petroleum Science and Engineering* (2012) **92**:65–81.

[6] Macinia P., Mesinia E., Salomoni V.A., Schrefler B.A., Casing influence while measuring in situ reservoir compaction, *Journal of Petroleum Science and Engineering* (2006) **50**:40–54.

[7] Song X., and Borja R.I., Mathematical framework for unsaturated flow in the finite deformation range. *International Journal for Numerical Methods in Engineering* (2014) **97**:658–682.

[8] Cuss, R. J., Rutter, E. H., and Holloway, R. F., Experimental observations of the mechanics of borehole failure in porous sandstone., *International Journal of Rock Mechanics and Mining Sciences* (2003) **40**:747–761.

[9] Simo, J.C., Algorithms for static and dynamic multiplicative plasticity that preserve the classical return mapping schemes of the infinitesimal theory, *Computer Methods in Applied Mechanics and Engineering* (1992) **99**:61–112.

[10] Carroll, M. M., A critical state plasticity theory for porous reservoir rock, *Recent Advances in Mechanics of Structured Continua* (1971) **97**:935–950.

[11] Stefanov, Y. P. and Chertov, M. A. and Aidagulov, G. R. and Myasnikov, A. V., Dynamics of inelastic deformation of porous rock and formulation of localized compaction zones studied by numerical modeling, *Journal of the Mechanics and Physics of Solids* (2011) **59**:2323–2340.

[12] Coelho, L. C., Soares, A. C, Ebecken, N. F. F., Alves, J. L. D., Lau, L., The impact of constitutive modeling of porous rocks on 2-D wellbore stability analysis *Journal of Petroleum Science and Engineering* (2005) **46**:81–100.

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis(Eds)

HYGROTHERMAL BEHAVIOUR OF HEMP CONCRETE; EXPERIMENTAL EVIDENCES AND MODELLING

Antonin FABBRI*, Pierre-Antoine CHABRIAC*, Dac Thuong NGO*, Frederic SALLET*, Emmanuel GOURDON*, Henry WONG* and Jean-Claude MOREL*

*LGCB-LTDS, UMR 5513 CNRS, ENTPE
Universit de Lyon, F69100 Vaulx-en-Velin, France
e-mail: antonin.fabbri@entpe.fr, pierre-antoine.chabriac@entpe.fr, dacthuong.ngo@entpe.fr, frederic.sallet@entpe.fr, emmanuel.gourdon@entpe.fr, KwaiKwan.wong@entpe.fr, jean-claude.morel@entpe.fr

Key words: Instructions, Coupled Problems, Multiphysics Problems, Applications, Computing Methods

Abstract. This paper presents experimental hygrothermal data of an hemp concrete wall of dimensions $0.9 \times 0.9 \times 0.1$ [m^3]. The wall is instrumented with sensors to monitor temperature, relative humidity at the middle of the wall and incoming heat flows at the external surfaces. It is placed in a double climatic chamber that allows the regulation of temperature and relative humidity on each side of the wall, independently to each other. The experimental results leads to a clear identification of the coupling between the variation of the relative humidity inside the wall and its temperature. The validity of the commonly adopted assumptions for hygrothermal simulation are finally analyzed in the light of these experimental results. The material parameters used for the simulations are measured separately on decimetric samples of the same hemp concrete, which comes from the same mix and with the same apparent density.

1 INTRODUCTION

Buildings now represent 23.5% of the greenhouse gas emissions in France [1].This energy cost is due to buildings consumptions (heat, ventilation...), but also to the embodied energy required for their construction, their rheabilitation and their dismantling. In this context, the use of bio-based insulating materials like hemp concrete are gaining interest. Indeed, they allow to drastically reduce fossil energy consumption and greenhouse gas emissions associated with the manufacture compared to conventionally used materials for building insulation such as rock wool or fiberglass [2].
They are renewable, extracted from biomass and have promising thermal and acoustical

characteristics [3], [4], [5]. In addition, thanks to their ability to allow diffusion of water vapor within their porous network [6], these materials present an alternative option to isolate and rehabilitate buildings made with non-industrial materials (rammed earth, cob, adobe,...) and whose stability requires to maintain a water exchange with the outside [7]. The thermal conductivity of hemp concrete depends on its formulation, its implementation and its water content. It commonly ranges from 70 mW/(mK) to 300 mW/(mK) [8], which is quite high for an insulating material. In comparison, the thermal conductivities of expended polystyrene and glass wool are about 40 mW/(mK). However, the thermal performance of hygroscopic insulating systems like hemp concrete is commonly above that what can be expected by considering their sole characteristics on thermal conductance and their thickness. This higher performance can be attributed to the latent heat associated with the liquid-to-vapor phase change of the water contained in the porous network of the material.
Another impact of this in-pore water phase change is the moisture buffering abilities of these materials. For example, at building scale, it was shown that the use of hygroscopic materials leads to a significant reduction of moisture variation amplitudes, which thus induces energy saves on ventilation [9].
In this context, this paper focus on experimental identification and quantification of the couplings between the flows of heat and water mass within a hemp concrete wall of dimensions $0.9 \times 0.9 \times 0.1$ [m^3]. The tested material, the experimental device and the implementation of the wall are presented in a first section. Then, the results obtained for two kind of hygrothermal loadings are presented and discussed in the light of numerical simulations based on the system of equation developed by [10]. The test conditions are chosen in order to clearly observe the impact of hygrometry on temperature variations.

2 MATERIALS AND METHODS

2.1 Material

The binder used in this study is the Tradical PF70. It is mixed with water and hemp by the amounts indicated in Table 1. The mixture is then compacted in a framework of internal dimensions $0.9 \times 0.9 \times 0.11$ [m^3] by 8 successive layers.

Hemp/Binder kg/kg	water/Binder kg/kg	ρ_d kg/m^3	ϕ -	μ -	μ^e -	λ at $w = 0\%$ mW/(m.K)	λ at $w = 5\%$ mW/(m.K)
0,33	0,81	504	0.46	0.42	0.25	147	162

Table 1: Main characteristics of the tested material

In addition to the wall, the same batch is used to realize cylindrical samples of 16cm diameter and 32cm length and plates of dimensions $0.3 \times 0.3 \times 0.06$ [m^3]. These samples are used for porous, hydric and thermal characterizations. The wet density just after

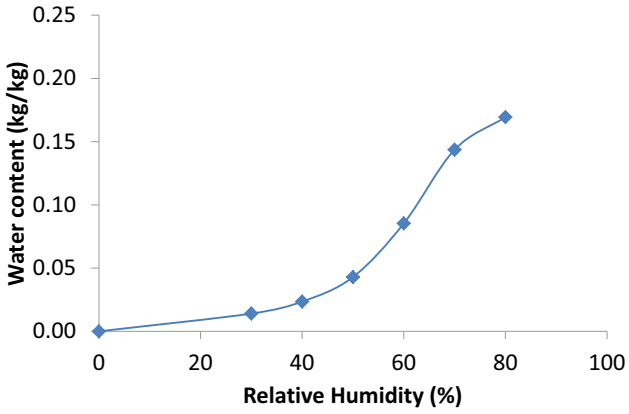


Figure 1: Sorption isotherm of the test hemp concrete at 20°C

the mixing of the wall, of the cylinders and of the plates are almost the same (800 $^{+}/_{-}$ 50 [kg/m^3]). The apparent dry density of the hemp concrete, ρ_d, is measured from the weighting of dried cylinders at 50°C, and the porosity, ϕ is estimated with a water pycnometer, through the relation : $\phi = (\rho_s - \rho_d)/\rho_s$, where ρ_s is the density of the skeleton estimated with the pycnometer.
The hydric characterization consists on a isothermal sorption isotherm following the NF EN ISO 12571 standard at 20°C , water vapor resistance (μ) following the NF EN ISO 12572 standard between relative humidities of 0% and 43.2%, and apparent water vapor resistance (μ^e) which is realized using the same protocol but at higher humidity (between 43.2% and 97%). These two last measurements allow to estimate the vapor and liquid transport characteristics, which are noted respectively δ_p and D_φ [10]:

$$\delta_p = \frac{\delta_a}{\mu} \; ; \; D_\varphi = p_{VS}\delta_a \left(\frac{1}{\mu^e} - \frac{1}{\mu} \right) \tag{1}$$

where δ_a is the vapor permeability of stagnant air (equal to 2.0 $\times 10^{-10}$ s at 20°C), and p_{VS} the saturated vapor pressure. This latter is an increasing function of temperature whose expression is well known and can be found, for example, in [11]. Let us underline that the value of D_φ obtained in this way remains valid only in the range of relative humidity used to estimate μ^e. A precise determination of this parameter would require the measurement of evolution of the relative permeability of the material with its water content.

Finally, the thermal conductivity, λ, is measured with a hot wire apparatus (Neotim - FP2C) and a hot plate apparatus (HFM - Netsch). The hot wire apparatus only allows

local measurements. Then, its accuracy is questionable when heterogeneous materials like hemp concrete are tested. To overcome this problem the thermal conductivities determined with this device are the average of a least five measurements at several locations of the sample. Using this protocol, the obtained result for the dried sample is similar to the one obtained with a standardized hot plate apparatus. In consequence, we can assume that the hot wire apparatus is suitable to estimate the thermal conductivity of the studied hemp concrete.
Thanks to its rapidity, the hot wire technique allows the determination of thermal conductivity of wet samples with a limited risk of water content variation during the measurement. The thermal conductivity of the tested material is then measured at both $w = 0\%$ and $w = 5\%$, where w is the gravimetric water content (defined as the mass of water within the pores per unit of solid skeleton mass). Hydric and thermal properties of the hemp concrete are reported in table 1 and in Figure 1 for the sorption curve. Let us underline that the obtained results are in the range of the values commonly observed in literature for this kind of material.

2.2 Calibration and design of the instrumentation scheme

The wall is instrumented with internal temperature and relative humidity sensors and with surface heat flow meters. It is placed in a double climatic chamber that allows the regulation of temperature and relative humidity on each side of the wall, independently to each other. The instrumentation of the wall is illustrated in Figure 2.
The accuracy of the temperature and heat measurements is first tested on a expended polystyrene wall whose thermal porperties are known and are equal to $\lambda = 40$ mW/(m.K) and $\rho C = 24$ kJ/(m^3.K). As illustrated in Figure 2, two temperature sensor are placed in the middle of the wall. The first one, called top-middle temperature sensor, is inserted from the top of the sample, while the other one, called side-middle temperature sensor, is inserted from the left side of the sample. The aim of these two measurements is to scan the impact of the temperature sensor implementation on the measurement accuracy.
For this preparatory test, the thermal solicitation consists in three steps. During the first one, the left climatic chamber temperature is held at 40°C and right climatic chamber temperature is held at 20°C. During the second stage, the temperature of both climatic chambers is fixed at 20°C. Finally, the last stage is the symmetric of the first one: the left climatic chamber temperature is kept at 20°C while the right climatic chamber temperature is increased to 40°C. Each stage lasts approximatively 4 hours.

The temperature and heat flow measurements are compared to a pure thermal one-dimensional calculation. The thermal equation, namely $\rho C \partial T/\partial t = \underline{\nabla} \cdot (\lambda \underline{\nabla T})$ where T is the absolut temperature in Kelvin and $\underline{\nabla}$ is the symbol Nabla, is solved with COMSOL Multiphysics using the PDE module. The simulation is made for a 1D geometry of length $L = 0.1$m (x direction). It represents the lateral section of the tested wall. The Fourier's boundary conditions are set according to the temperature measurement within climatic

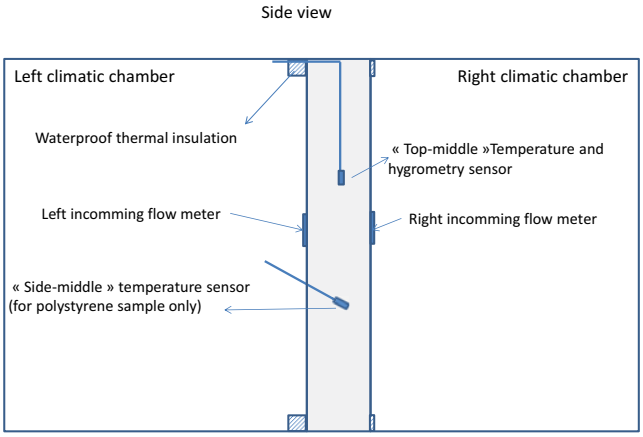


Figure 2: Schematic representation of the instrumentation of the wall

chambers and the heat transfer coefficients are supposed constant and respectively equal to 6 W/(m^2.K) for the left chamber and 7 W/(m^2.K) for the right one. These values are estimated from the average air velocity measurements within the climatic chambers at the vicinity of the wall which are equal to 0.1 m/s for the left one and 0.4 m/s for the right one.

Results, reported in Figure 3, show a good correlation between the measured and calculated incoming heat flows at both surface. The same conclusion can also be hold for the top-middle temperature sensor measurements. This comparison gives some confidence on the reliability of the temperature and heat flows which are measured. On the opposite, a significant difference is observed between the calculation and the measurements of the side-middle temperature sensor. Indeed, this latter remains very close to the temperature of the left climatic chamber, which suggest that, in that case, a significant thermal bridge is created between the sensor and the climatic chamber. In consequence, even if this experimental bias should be lowered when the thickness of the wall increases, the insertion of temperature sensor along the thermal gradient direction should be, whenever possible, avoided, and the measurement from temperature sensor instrumented in that way should be analyzed with care.

2.3 Definition of the loading paths

This paper focus on the coupling between heat and mass transfers within hemp concrete. To do so, two simple hygrothermal loadings are performed. The first one consists in a sudden increase of relative humidity within the two climatic chambers from 35% to

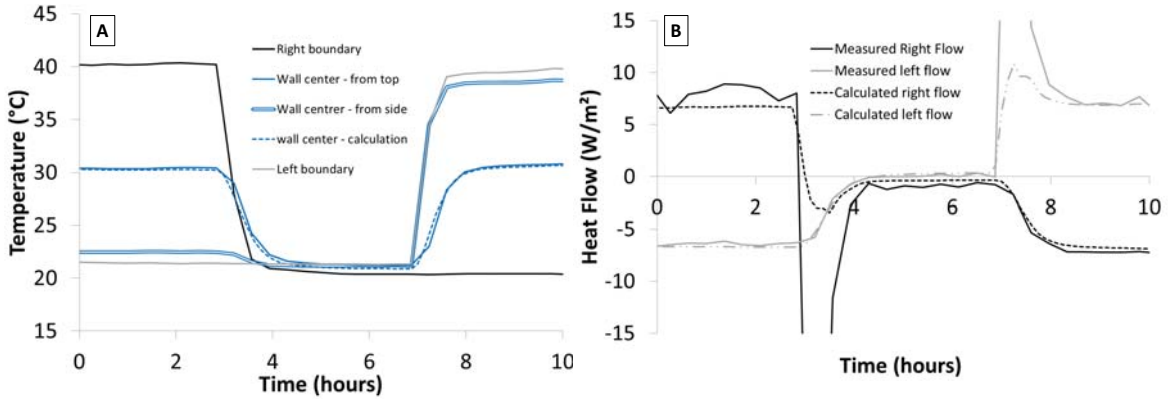


Figure 3: Comparison between the measured and the calculated temperature (A) and heat flows (B) for the polystyrene wall

80% at a constant temperature of 30°C. In that way, any variation of temperature in the middle of the wall and any heat flow across the wall surfaces would be induced by the condensation of water within the wall.
In the second experiment the left climatic chamber temperature is held at 30°C while the right one is held at 40°C. The two climatic chambers are initially at a relative humidity of 50%. After a stabilization period (that last about one week), a sudden increase of relative humidity from 50% to 80% is performed within both climatic chambers. Then, after another stabilisation period of 8 days, a sudden decrease of relative humidity from 80% to 50% is imposed within both climatic chambers. The purpose of this second experiment is to scan the effect of the phase change phenomena when the wall is submitted to temperature and vapor pressure differences between its two sides, and to compare the amplitude of the heat flow caused by this temperature difference and the one cause by the phase change processes. Let us underline that each loading are preceded by an equilibration period that last until the hygrometry and temperature measured within the wall are the same than the ones within the two climatic chambers.

3 RESULTS

The obtained variation of temperature and relative humidity within the two climatic chambers and at the middle of the wall during the first test are reported in Figure 4.

At first, even if the temperature within both climatic chambers remains fairly stable, the increase of relative humidity leads to a significant increase of temperature of 3°C within the wall. The shape of the wall temperature variation seems to follows the temporal derivation of the wall internal relative humidity (double-lined curve in Figure 4B).
The Figure 5A shows that this temperature increase go along with a decrease of the incoming heat flow at both sides down to significant negative values. Assuming that the

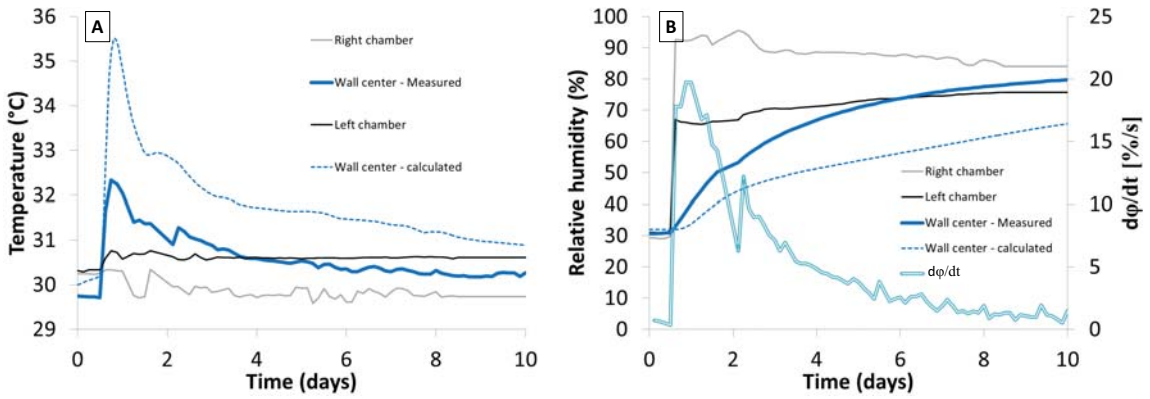


Figure 4: Temperature (A) and relative humidity (B) variations within the two climatic chamber and at the middle of the wall during the first test and their comparison with numerical results

heat flow through the isolated surfaces of the wall are negligible, the power provided by the wall due to the condensation process reads:

$$W = -(q_{left} + q_{right}) \tag{2}$$

where W is the supplied power per unit of wall lateral surface, while q_{left} and q_{right} are respectively the measured incoming heat flow through the left and right sides of the wall. It leads to values up to 10W/m^2. Given the fact that this supplied power remains higher than 6W/m^2 for at least 24 hours, the induced energy provided by the wall can become significant if an important surface of an hygroscopic wall is submitted to condensation.

Let us now observe the results from the second experiment which are reported in Figure 6 for the temperature and relative humidity variations and in Figure 5B for the heat flows. Similarly to the previous experiment, the wall temperature variation follows the same tendency than the variation of temporal derivation of relative humidity within the wall. An increase of the wall temperature (up to 1.5°C) is observed when the relative humidity increases and a decrease of wall temperature (down to 1.5°C) is observed when the wall relative humidity decreases. When the wall relative humidity returns to equilibrium with the relative humidity within the climatic chambers, no more phase change phenomena occurs and the temperature at the middle of the wall becomes equal to the average of the climatic chambers temperatures. In that case, the Figure 5B shows that the incoming heat flows from the left and right sides are opposite and equal to respectively -10W/m^2 and 10W/m^2. During the phase change processes, these two flows are no more directly opposite. The condensation process leads to a strong reduction of the incoming flow through the "hot" surface, while the evaporation process leads to a strong reduction of the outgoing flow though the "cold" surface. For at least 1 day, these reductions remain

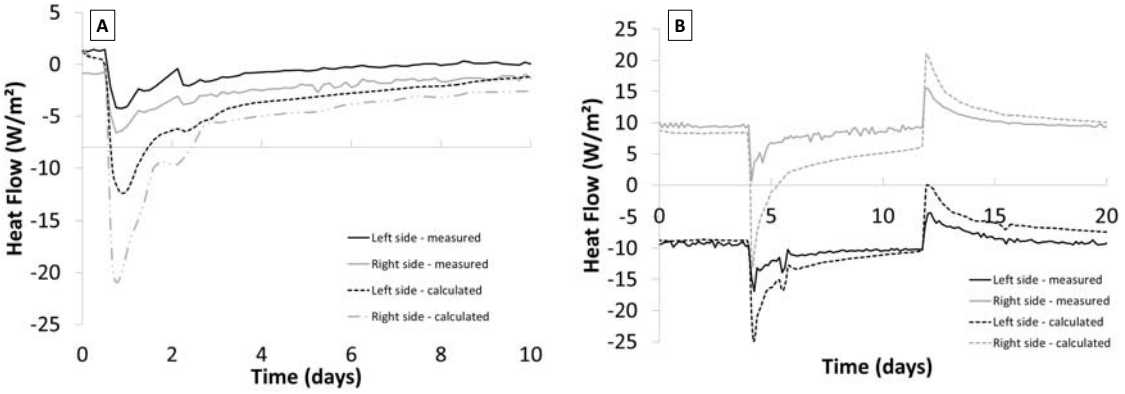


Figure 5: Heat flow at the left and right sides of the wall during the first (A) and the second (B) tests and their comparison with numerical results

higher than 75% of the heat flow induced by a temperature difference of 10°C between the chambers.

4 DISCUSSION

A simple way to explain the variation of temperature with time derivation of relative humidity is to consider the thermal equation in its unidimensional form for porous media with in-pore water phase change (proof reported in [12] and [11]):

$$\langle \rho C \rangle \frac{\partial T}{\partial t} = \underline{\nabla} \cdot (\langle \lambda \rangle \underline{\nabla T}) - L_V \overset{o}{m}_V \tag{3}$$

where L_V stands for the latent heat of evaporation/condensation, $\overset{o}{m}_V$ is the rate of water mass which evaporates per unit of material volume, $\langle \rho C \rangle$ is the average volumetric heat capacity of the material, and $\langle \lambda \rangle$ is its average thermal conductivity. These two latter vary with water content and temperature. However, only their variations with water content is considered here following the relations:

$$\langle \rho C \rangle = (1-\phi)\rho_s C_s + \rho_d w C_L + \left(\phi - \frac{w\rho_d}{\rho_L}\right) C_A \rho_A \ ; \ \langle \lambda \rangle = \lambda_0 + \delta_\lambda w \tag{4}$$

where $C_s = 1.6\text{kJ/(kg.K)}$ [13], $C_L = 4.2\text{kJ/(kg.K)}$ and $C_A = 1.0\text{kJ/(kg.K)}$ are the specific heat capacities of the solid skeleton, the liquid water and the wet air, while $\rho_A = 1\text{kg/m}^3$ and $\rho_L = 1000\text{kg/m}^3$ are the density of air and of liquid water. Finally, $\lambda_0 = 147\text{mW/(m.K)}$ is the dried thermal conductivity and $\delta_\lambda = 300\text{mW/(m.K)}$ is the variation of $\langle \lambda \rangle$ with water content. For illustration purpose, these two relations leads to $\langle \rho C \rangle = 981\text{kJ/(m}^3\text{.K)}$ and $\langle \lambda \rangle = 162\text{mW/(m.K)}$ at $w = 0.05$.

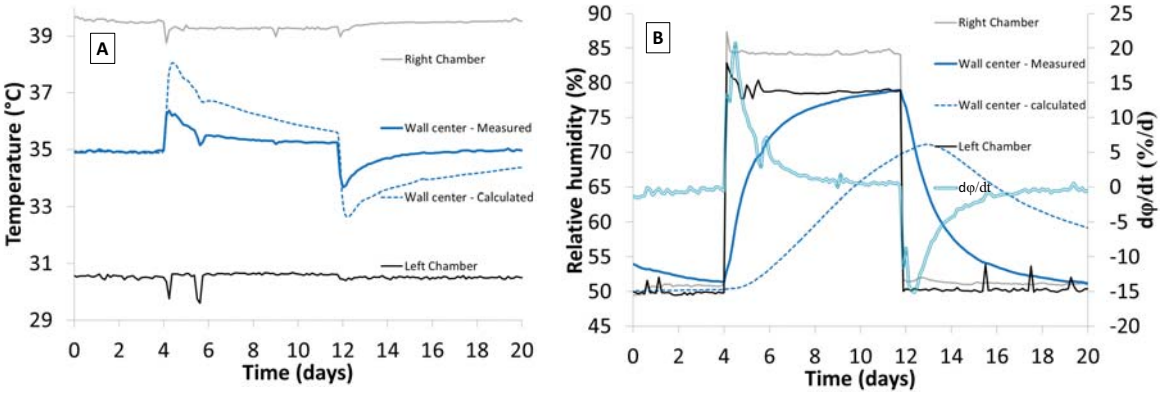


Figure 6: Temperature (A) and relative humidity (B) variations within the two climatic chamber and at the middle of the wall during the second test and their comparison with numerical results

According to the sorption curve of the material reported in Figure 1, the water content of the material increases with relative humidity. The obtained results are then consistent with eq. (3). Indeed, the relative humidity increases within the climatic chambers will lead to an increase of the water content of the wall and thus to condensation process within the porous network. In that case, $\overset{o}{m}_V$ is negative, and the last term of (3), which correspond to the heat supplied by the phase change process will be positive. On the opposite, a decrease of the chambers relative humidity leads to a positive $\overset{o}{m}_V$, and thus to heat consumption. However, the quantification of this effect is not so simple and a precise evaluation of $\overset{o}{m}_V$ requires a fully coupled hygrothermal model. To do so, it is commonly assumed that the equilibrium between the liquid water and its vapor is reach at anytime and anywhere within the porous network of the wall, and that the overall mass variation of water vapor can be neglected. Under these two assumptions, the mass conservation equation of water vapor leads to equality between the evaporation mass rate and the outgoing mass flow of vapor within the material :

$$\overset{o}{m}_V = -\underline{\nabla} \cdot \left(\delta_p \underline{\nabla \varphi p_{VS}}\right) \tag{5}$$

where φ is the relative humidity and p_{VS} is the saturated vapor pressure. Eq. (5) allows rewriting the thermal equation in the form :

$$\langle \rho C \rangle \frac{\partial T}{\partial t} = \underline{\nabla} \cdot \left(\langle \lambda \rangle \underline{\nabla T}\right) + L_V \underline{\nabla} \cdot \left(\delta_p \underline{\nabla \varphi p_{VS}}\right) \tag{6}$$

The outgoing mass flow of vapor within the material can eventually be estimated from the overall mass conservation of water. Assuming that the liquid water mass flow induced by temperature gradient at constant relative humidity is negligible and recalling that the water content is defined as $w = (m_V + m_L)/m_S$ where m_V, m_L and m_S are respectively

the mass of vapor, liquid water and solid skeleton, this latter reads:

$$\frac{\partial m_V + m_L}{\partial t} = \rho_d \frac{\partial w}{\partial t} = \underline{\nabla} \cdot \left(\delta_p \underline{\nabla \varphi p_{VS}} + D_\varphi \underline{\nabla \varphi} \right) \tag{7}$$

The system of equation (6)-(7), which was initially developed by [10], is the one which is commonly used to compute hygrothermal problems. It is solved with COMSOL Multiphysics using the PDE module. The simulations are made for a 1D geometry of length L=0.11m (x direction). It represents a cross section of the tested wall. Similarly to the purely thermal simulation realized on polystyrene, the Fourier's boundary conditions are set according to the measurement within the two climatic chambers and with heat exchange coefficients of 6 W/(m^2.K) for the left chamber and of 7 W/(m^2.K) for the right one. As suggested by [10] the moisture exchange coefficient is estimated to be equal to 25 10^{-9} kg/(m^2.s.Pa). No exchange of liquid water is assumed between the wall and the exterior. The material parameters used for the simulation are the ones reported in Table 1. Comparisons of calculated temperature, relative humidity and heat flows with their measurements are reported in Figures 4, 5 and 6.
Whatever the considered test, the system of equation (6)-(7) seems to overestimate the heat flows and the temperature variations while it tends to underestimate the variation of relative humidity at the middle of the sample. These two observation are contradictory. Indeed, to reach approximatively the same variation of the relative humidity, we need to consider a vapor resistance coefficient μ around 0.5 (which is, by the way, unrealistic since, by definition, μ should not be lower than 1). This decrease of μ will increase the vapor diffusion coefficient δ_p up to 4 10^{-10} kg/m/s/Pa and thus increase the heat source term, namely $L_V \underline{\nabla} \cdot \left(\delta_p \underline{\nabla \varphi p_{VS}} \right)$ in the thermal equation. In consequence, the difference between the calculated and measured heat flows and temperature would become even higher. To counterbalance this effect, we need to consider an specific thermal capacity C_d of about 150 000 J/(kg.K), which is clearly unrealistic and which leads to a significantly smoother temperature variation than the measured one.
In conclusion, the classical hygrothermal equation gives the good tendencies but does not seems to quantify correctly the hygrothermal behavior of the tested wall. Consequently, like it was recently made for raw earth materials by [11], the relevance of the simplifying assumptions which are commonly made in hygrothermal computations should be tested in the case of hemp concrete.

5 CONCLUSIONS

An hygrothermal experiment of an hemp concrete wall of dimensions $0.9 \times 0.9 \times 0.1$ [m^3] is presented. The wall is instrumented with sensors to monitor temperature, relative humidity at the middle of the wall and incoming heat flows at the external surfaces. It is placed in a double climatic chamber that allows the regulation of temperature and relative humidity on each side of the wall, independently to each other. Before the experiment, the accuracy of the measurement are tested on a calibrated material. This preliminary

test underlines the bias on temperature measurements when the sensors are introduced in the wall following the thermal gradient direction.
The test on the hemp concrete wall allows to observed its hygrothermal ability with a heat power supplied by the wall higher than 6W/m^2 during 24 hours when the relative humidity of the ambiance increase from 30% to 80%.
When a 10°C difference of temperature is imposed between the left and the right climatic chambers, the condensation process leads to a strong reduction of the incoming flow through the "hot" surface, while the evaporation process leads to a strong reduction of the outgoing flow though the "cold" surface. For at least 1 day, these reductions remain higher than 75% of the heat flow induced by a temperature difference of 10°C between the chambers.
Finally, the numerical computation of this experimental results underlines that the classical hygrothermal equation leads to the good tendencies, but does not seems to quantify correctly the behavior of the tested wall. The relevance of the simplifying assumptions which are commonly made in hygrothermal computations should thus be checked in the case of hemp concrete.

Acknowledgments : The present work is supported by the French Agency for Environment and energy Management (ADEME) through project IBIS (PIA - Projets d'Investissement d'Avenir).

REFERENCES

[1] Ministère de l'Ecologie, du Développement Durable et de l'Energie *Emissions de gaz à effet de serre (Monde, Europe, France)* (2013) URL http://www.developpement-durable.gouv.fr/Part-et-evolution-des-secteurs.htm.

[2] Harris, D. A quantitative approach to the assessment of the environmental impact of building materials. *Building and Environment* (1999) **34**:751–758.

[3] Pierre, T, Colinart, T, Glouannec, P. Measurment of thermal properties of biosourced building materials. *International Journal of Thermophysics* (2013).

[4] Asdrubali, F., Schiavoni, S., Horoshenkov, K.V. A review of sustainable materials for acoustic applications. *Build. Acoust.* (2013) **19**:283–312.

[5] Glé, P., Gourdon, E., Arnaud, L. Acoustical properties of materials made of vegetable particles with several scales of porosity. *Appl. Acoust.* (2011) **72**:249–259.

[6] Walker, R, Pavia, S. Moisture transfer and thermal properties of hemp-lime concretes. *Construction and Buildings Materials* (2014) **64**:270–276.

[7] Champiré, F, Fabbri, A, Morel, J-C and Wong, H. Experimental evidences of the atypical behavior of compacted earth materials. *Materials and Structures* (submitted).

[8] Collet, F, Pretot, S. Thermal conductivity of hemp concretes: variation with formulation, density and water content. *Construction and Buildings Materials* (2014) **65**:612–619.

[9] Woloszyn, M, Kalamees, T, Abadie, MO, Steeman, M, Sasic Kalagasidis A. The effect of combining a relative-humidity-sensituve ventilation system with the moisture buffering capacity of materials on indoor climate and energy efficiency of buildings. *Building and Environment* (2009) **44**:515–524.

[10] Künzel, M. *Simultaneous heat and moisture transport in building components one - and two-dimensional calculation using simple parameters* (1995) PhD thesis, Fraunhofer Institute of Building Physics.

[11] Soudani, L, Fabbri,A, Morel, J-C, Woloszyn, M, Chabriac, P-A, Wong, H, and Grillet, A-C. Experimental evidences of the atypical behavior of compacted earth materials. *Energy and Buildings* (In revision).

[12] Coussy, O. *Poromechanics* (2004) John Wiley & sons.

[13] Evrard, A. *Transient hygrothermal behaviour of Lime-Hemp Materials.* (2008) PhD thesis, Ecole polytechnique de Louvain, Belgique.

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis (Eds)

NUMERICAL COUPLING BETWEEN DAMAGE AND GAS PERMEABILITY FOR CONCRETE APPLIED ON A 3D SPLITTING TEST

M. Ezzedine El Dandachy, M. Briffaut, F. Dufour and S. Dal Pont

Univ. Grenoble-Alpes,
CNRS, 3SR UMR 5521, BP 53, 38041 Grenoble, France
E-mail: mohamad.ezzedineeldandachy@3sr-grenoble.fr, www.3sr-grenoble.fr

Key words: Damage, Gas Permeability, Coupling, Splitting Test, Mortar.

Abstract. Due to the industrial needs, one of the key issues nowadays is to develop numerical tools that are able to predict the leakage rate through a cracked structure. This paper presents a validation of a numerical modelling of leakage rate through a mortar specimen in a splitting test versus experimental results. The mechanical state of the material is described by means of an enhanced non-local damage model which takes into account the stress state and provides a realistic damage field at failure. Two continuous approaches based on permeability-damage law are considered to study the coupling between the mechanical state of the material and the permeability. The first one lies on coupling the permeability of a crack with the mechanical state by means of the damage variable, and the second one is based on the regularized equivalent strain. Results show that the first approach isn't acceptable in case of discontinuous macrocracking while the second one predict well the permeability in the same case.

1 INTRODUCTION

During their service life, due to external loading (mechanical and/or environmental), concrete structures may undergo damage in a diffuse manner (microcracking) at the material scale and/or localised (macrocracking) at the structural level. The estimation of the evolution of transfer properties in such a cracked material is a key issue for structural durability analysis. Choinska et al. **[1]** observed in their experimental study three different regimes of permeability evolution. A first regime exhibiting relatively a slight permeability increase, which is due to the presence of microcracks spread out in the material. Coupling between permeability and microcracking (diffuse damage), has been proposed by Picandet et al. **[2]** when the material is in compression. A second regime is also observed where the permeability of the material increase rapidly due to strain localization. This regime is an intermediate phase between diffuse damage and discontinuous macrocracking. A third regime is observed where macrocracks are formed and permeability is governed at the macrostructural level by Poiseuille's law and mainly depends on the crack opening, this regime is characterized by a slower rate of permeability increase with respect to the second regime.

Two approaches are considered in order to numerically simulate the coupling between the mechanical state of the material and the permeability:

- A first approach was proposed by Pijaudier-Cabot et al. **[3]** that is based on a permeability-damage law, which is a combination of Picandet and Poiseuille's permeability. Where the two permeabilites are directly related to damage.

- A second approach proposed in this paper that is also based on a permeability-damage law, Picandet's permeability is retained but Poiseuille's permeability is no longer function of damage but related to the regularized equivalent strain.

Those two approaches allow to predict a leakage rate without need to calculate the CODs. The permeability is obtained at each integration point and is used to compute the leakage rate point wise in the volume. The mechanical state of the material is described by means of an enhanced non-local damage model which takes into account the stress state and provides a very realistic damage field representing micro cracking and macro cracking at failure **[4]**.

A physical experiment has been performed on a mortar specimen subjected to splitting test; the gas permeability of the specimen is measured during the test at different load levels. The validation of the approach against experimental results is performed on the leakage rate perpendicular to the disk for different load stages.

The physical experiment of the splitting test will be detailed in section 2 and the numerical models (damage and continuous approach for leakage rate) will be presented in the section 3. In section 4 the application of the numerical models on the splitting test will be detailed.

2 PHYSICAL EXPERIMENT

The experiment **[5]** consists of performing a brazilian splitting test applied on mortar specimens. The Brazilian splitting test is an indirect tension test used to measure tensile strength of concrete, rocks and other geomaterials. It consists of loading a cylindrical specimen along a diametral plane by means of steel or wood bearing plates, as shown in **Fig**.**1**. Gas flow rate measurements are taken after partially unloading the sample to avoid brutal rupture (See **Fig**. **2**). The sample has a cylindrical shape of 40 mm of height and 110 mm of diameter. It is worth noting that the sample wasn't a perfect cylinder, a difference of 0.1 mm in diameter was observed. The Young modulus is equal to 18 MPa and the Poisson's coefficient equal to 0.2.

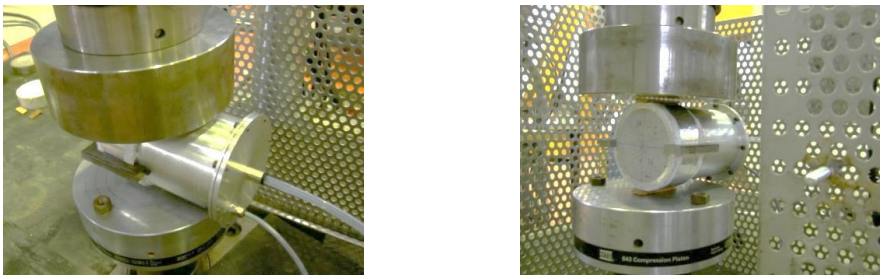


Figure1: Permeability analysis while performing a Brazilian test on a mortar sample. **[5]**

The splitting test is controlled and directed by the crack opening located on the horizontal transversal line of the bigger face of the specimen, a displacement sensor is placed in the

central part of this face to do the controlling. The other face (the smaller) is discretized by means of a speckle pattern in order to perform Digital Image Correlation and get the 2D displacement field on the surface.

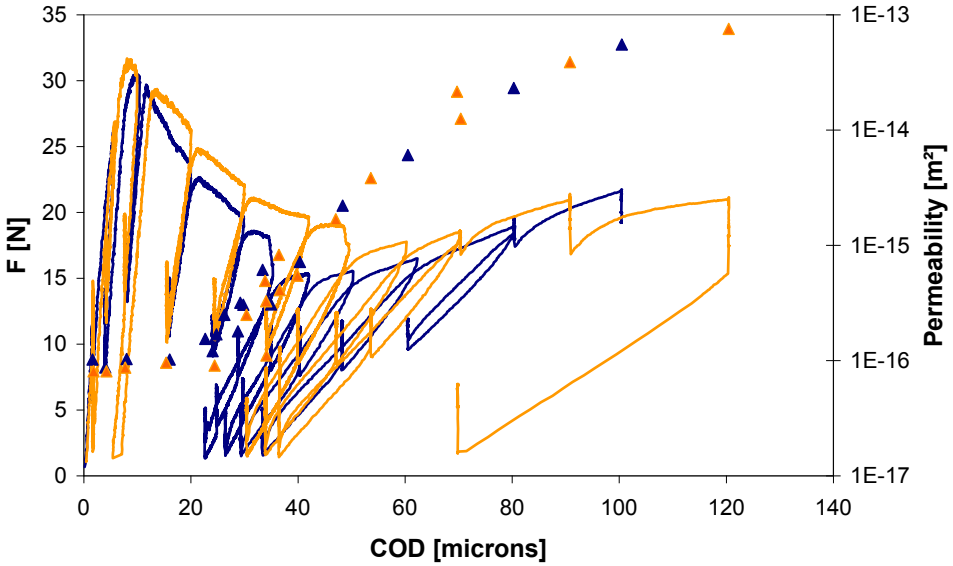


Figure2: Mechanical behavior and permeability evolution versus the crack opening displacement for two mortar specimens. **[5]**

The Mechanical response and the permeability evolution of two mortar specimens are given in **Fig**. 2. The mechanical response is composed of an elastic part, and after reaching the peak the softening behavior ends by a complete split of the specimen. At COD equal to approximately 34 microns the specimen is split and the behavior afterwards is described by the half portion of the specimen subjected to compression loading. An analysis has proven that there was a 3D effect on the force-COD behavior of the specimen. The crack is initiated and propagated firstly on the bigger face and then is propagated in the longitudinal direction to reach the smaller face and propagate on it until the total split of the specimen. Therefore 3D simulations are needed.

The permeability evolution can be described by the three regimes proposed by Choinska: For crack openings varying from 0 to 30 microns a slight increase in the permeability is recorded which corresponds to the first regime. Afterwards the second regime appears when a large increase in the permeability begins for crack openings higher than 30 microns. Finally, one can see a decrease of the rate permeability increase. In the third regime, a coefficient ξ that includes the roughness of crack, the turtuosity and the constrictivity of the material is calculated at COD equal to 100 microns. This coefficient is an adjustment factor for the Poiseuille's law (See **Fig. 3**).

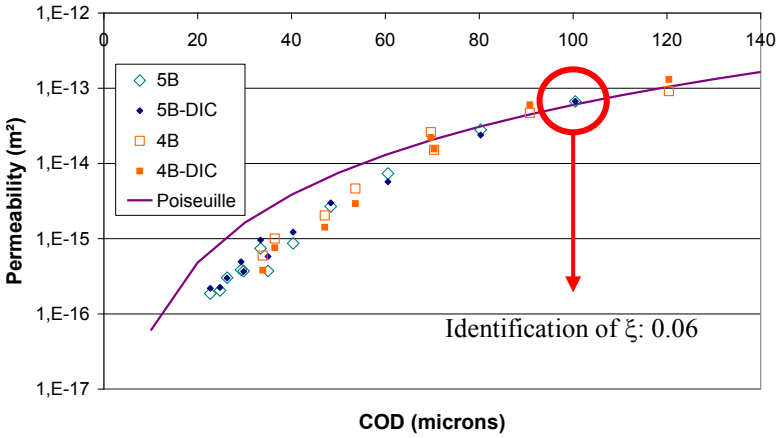


Figure3: Identification of an average roughness coefficient at COD equal to 100 microns. **[5]**

3 CONTINUOUS APPROACH BASED ON DAMAGE

The advanced numerical modelling of cracking in reinforced concrete structures may be dealt with either by means of advanced discontinuous models (G-FEM, etc.) or using continuous constitutive models such as damage mechanics. In this approach damage models are considered since the permeability-mechanical state coupling is based on damage. The enhanced non-local damage model that is based on the stress state will be reviewed in the following. In the second subsection the continuous approach is presented.

3.1 Regularized damage models

The loss of stiffness associated to mechanical degradation of the material is represented by D:

$$\sigma = (1 - D)C : \varepsilon \tag{1}$$

Where σ and ε are respectively the Cauchy stress tensor and the small strain tensor, and C is the tensor of elastic moduli.The parameter D range between 0 for virgin materials and 1 for completely damaged materials. It is assumed that D depends on a state variable Y, which in turn depends on the strains, Y = Y (ε).The basic idea of nonlocal damage models is averaging the state variable Y in the neighborhood for each point. In this manner, the nonlocal state variable $\tilde{Y}$is obtained:

$$\tilde{Y} = \int_v \alpha(d) Y dV \Big/ \int_v \alpha(d) dV \tag{2}$$

The weight function α depends on the distance d to the point under consideration. Generally a Gaussian function is used:

$$\alpha(d) = \exp\left[-\left(\frac{2d}{l_{c0}}\right)^2\right] \tag{3}$$

Where l_{c0} is a material parameter of the nonlocal damage model called characteristic length, and Y is the state variable, that drives the damage (D=D($\tilde{Y}$)) according to Mazars law:

$$Y = \sqrt{\sum_i [\max(0, \varepsilon_i)]^2} \tag{4}$$

Damage evolution follow a law which distinguishes tensile damage D_t and compressive damage D_c:

$$D = \alpha_t D_t + \alpha_c D_c \tag{5}$$

Where α_t and α_c are the weights computed from the strain tensor. The tensile damage D_t and compressive damage D_c are calculated as follows:

$$D_{t,c} = 1 - \frac{Y_{D0}(1 - A_{t,c})}{Y} - \frac{A_{t,c}}{e^{[B_{t,c}(Y - Y_{D0})]}} \tag{6}$$

The newly version of this damage model, the non local stress based (**NLSB**) is characterized by the regularization that takes into account the stress state of the material. A modification on the Gaussian function is applied and the function becomes:

$$\alpha(d) = \exp[-\left(\frac{2d}{l_{c0} * \rho}\right)^2 \quad] \tag{7}$$

Where ρ a factor that is calculated for each point and depends on the principal stresses of the medium.

3.2 Permeability-Continuous approach

The continuous approach is based on a permeability-damage law that is presumed to describe the three different regimes. The following equation is one of the matching laws proposed by Pijaudier-Cabot et al. which was proven to be the most convenient to describe the permeability evolution.

$$\log(k) = (1 - D)\log(k_D) + D\log(k_f) \tag{8}$$

This law is a logarithmic combination of Picandet's permeability for diffuse damage (low damage) and Poiseuille's permeability when a macrocrack is formed. k_D and k_f are two local permeabilites defined for each integration point in the medium and D is the value of damage at the integration point. Regarding diffuse damage, the first term of the modified Picandet's permeability is retained. Hence k_D is equal to:

$$k_D = k_0\left[1 + (\alpha D)^\beta\right] \tag{9}$$

Where α and β are two parameters to be determined from experimental results. They were fitted by Picandet to 11.3 and 1.64 respectively on experimental results of axial compressive damage on gas permeability of ordinary and high-performance concrete. However for a splitting test the damage is mainly generated in tension and is localized in the crack surface,

Consequently a negligible variation of the mean structural permeability due to diffuse damage is seen in the physical experiment.
Concerning the permeability of a macrocrack, Poiseuille's permeability of two parallel planes ($k_P = \xi\frac{[u]^2}{12}$) is considered with taken into account the coefficient (ξ). Pijaudier-Cabot et al. supposed equivalence at failure between continuous damaged domain with a damaged zone of width λl_c and discrete macrocracked domain with a crack opening $[u]$. This assumption was followed in order to substitute crack opening by damage field. Then the Poiseuille's permeability is equal to:

$$k_P = \xi\frac{(\lambda l_c)^2}{12}\left(F^{-1}(D) - Y_{D0}\right)^2 \tag{10}$$

Where l_c is the characteristic length, λ is a unit less factor that influences the width of the damaged band given equal to 2 and Y_{D0} is the damage threshold. The inverse damage evolution law $F^{-1}(D)$, based on equation (6), is used in order to represent the permeability k_f:

$$F^{-1}(D) = Y_{D0} - \frac{\ln(1-D)}{B_t} \tag{11}$$

The second approach consists firstly of using a function F($\tilde{Y}$) instead of the inverse function of damage. Then the crack opening is related to this function as follows:

$$[u] = Dens*(\mathrm{F}(\widetilde{\mathrm{Y}}) - \mathrm{Y}_{\mathrm{D0}}) \tag{12}$$

Where Y_{D0} is the threshold of damage (See **Table 1**) and *Dens* is an average of the density of the mesh of the fracture zone. F($\widetilde{\mathrm{Y}}$) for a cracked element is the component of the regularized equivalent strain $\widetilde{\mathrm{Y}}$ for this element that is in the orthogonal direction to the element surface. That means that before applying **Eq. 12** the cracked elements and its orientations has to be identified using numerical tools such as the global tracking algorithm **[6]**. In case of splitting test, for sake of simplicity one can suppose that the plane of symmetry is the surface of the crack, it means all the elements of this surface are cracked and the orthogonal direction to the elements in this case is the horizontal direction parallel to the x axis. Thus, in this case F($\widetilde{\mathrm{Y}}$) is the horizontal component of the regularized equivalent strain $\widetilde{\mathrm{Y}}$.
The permeability of a cracked element at the integration point k_f is defined as follows:

$$k_f = k_P\frac{[u]L}{S} \tag{13}$$

Where $[u]$ the crack opening of the element, and $\frac{L}{S}$ is the fraction between the length of the element and its surface. For sake of simplicity $\frac{L}{S}$ is calculated as $\frac{1}{l_e}$ where l_e equal to $3\sqrt{V_e}$, (V_e is the volume of the element).Thus k_f is equal to:

$$k_f = \xi \frac{[u]^3}{12 l_e} \tag{14}$$

To determine the total leakage rate, the permeability problem by imposing a pressure gradient has to be solved. Once the leakage rate is determined, by means of Darcy's law, the mean structural permeability can be calculated and will be compared to the permeability measured experimentally.

4 APPLICATION ON A 3D SPLITTING TEST

This numerical study is a simulation of the physical experiment presented in **section 2**. The steel bearing plates are modelled as rigid plates, with high Young's modulus (E = 300 GPa) and Poisson's ratio ν of mortar (ν=0.2) in order to avoid a confinement effect of mortar. Numerical simulations are performed in the FE code *Cast3m* with 4-nodes tetrahedral elements in 3D. Due to the symmetry of the problem, the computation domain consists of only quarter of the specimen. The mesh is shown in **Fig. 4** (b). It should be noted that the mesh is generated with the same conicity of the real specimen in order to reproduce the 3D effect in the simulations. **S_s** and **S_b** corresponds respectively to the smaller and bigger edge surfaces of the specimen. The post-peak behavior in splitting test includes a snap-back in the force displacement curve and therefore an arc-length control (by maximum strain or crack opening) is required to solve the numerical problem **[7,8]**.

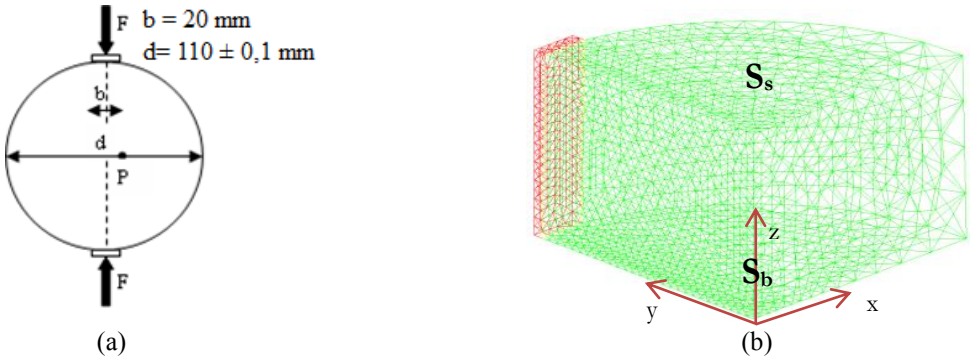


Figure 4: Brazilian splitting test (a) problem statement (b) 3D conic FE mesh.

4.1 Mechanical description

In this subsection the mechanical description by the stress based non local **NLSB** is concerned. Young's Modulus and Poisson's ratio of the mortar are taken from physical experiment. A calibration of the parameters is done for the model after finding a good

numerical description of the physical response and the set of Mazars' parameters is given in table 1.

Parameter	NLSB
Y_{D0}	$3.5.10^{-4}$
A_c	1.4
A_t	0.88
A_c	800
A_t	4050
β	1.06
l_{c0}	7.5 mm

Table 1:Set of Mazars' parameters for the damage models

The point P on which the crack opening displacement is calculated is located on the central horizontal line of the bigger face distant of 10 mm from the center (See **Fig. 4** (a)). The numerical simulation of the mechanical response (Force versus COD at P) of the physical experiment presented in **section 2** is given in **Fig. 5.** This result shows that the stress based non local model NLSB is in a good agreement with the physical response in describing the behavior even after the total split. However, the model is elasto-damageable model that do not take into account the plastic deformation therefore the exaggerated snap-back where COD decreases from about 35 to about 30.5 microns is probably due to elastic discharge of the face where the COD of P is calculated. This drawback in the model will eventually affect the coupling since the permeability is directly linked to the mechanical state.

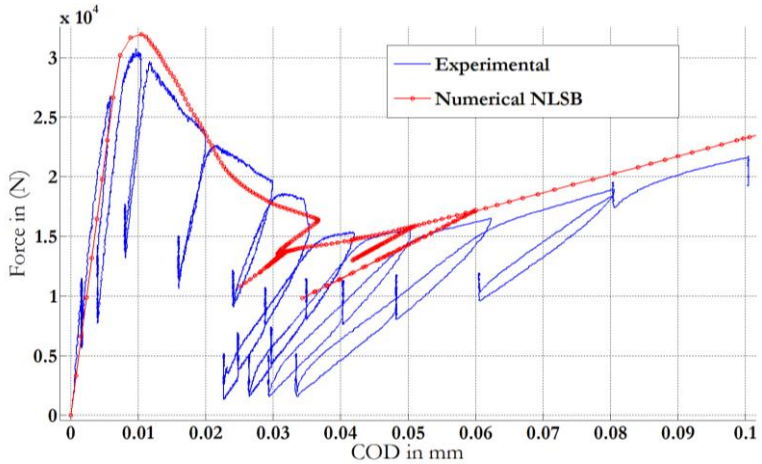


Figure 5: Mechanical response (Force versus COD) described by NLSB compared to experimental response.

4.2 Damage propagation

Damage fields at 4 loading levels given by the NLSB model are displayed in **Fig. 6**. At the peak A, damage is initiated in the center of the surface $\mathbf{S_b}$ and diffused in the volume. In the post-peak behavior (between A and B) localized damage propagates from the center on the surface $\mathbf{S_b}$ and in the transversal direction toward the surface $\mathbf{S_s}$. At point B the damage arrives

to the surface S_s and between the points B and C damage propagates on the surface S_s. At point C the behavior becomes the one of the half portion and almost all the crack surface is totally damaged. D corresponds to a total split with COD on P equal to 105 microns.

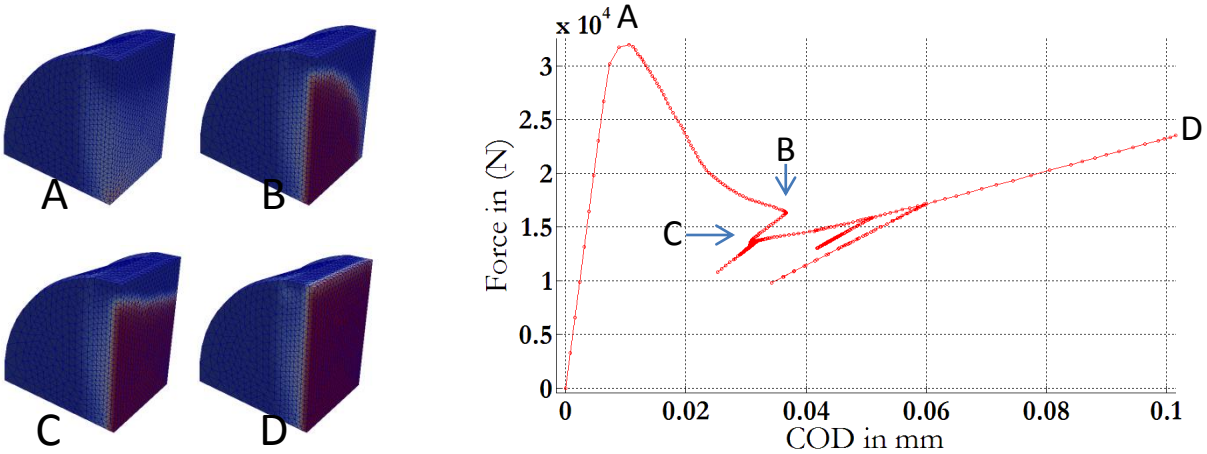


Figure 6: Damage fields given by NLSB at 4 loading levels A B C and D.

5 COUPLING PERMEABILITY-DAMAGE

Once the numerical simulation of the mechanical response is achieved, one can apply the matching law in the post treatment phase in order to obtain the permeability at each integration point and then compute the leakage rate point wise in the volume by applying a pressure gradient and solving the permeability problem (See **Fig. 7**). Hence, the mean structural permeability is calculated using Darcy's law and compared to experimental measurements.

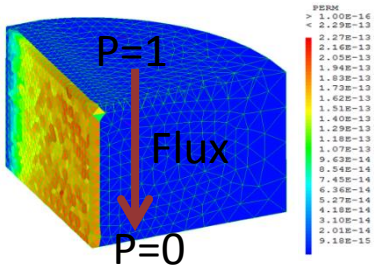


Figure 7: Solving the permeability problem by applying a pressure gradient.

The evolution of the mean permeability calculated by applying the matching law proposed by Pijaudier-Cabot et al. versus the COD at point P is presented in **Fig. 8**. It is shown that the estimated mean permeability for low level of damage is acceptable. However in the intermediate zone the permeability is overestimated due to an overestimation of the crack opening (COD at P ranging from 380 to 600 microns). Moreover for higher COD, the evolution of the mean permeability becomes asymptotic. This is due to the stabilization of the damage around 0.99. Consequently the relation between the crack opening and the inverse function of D $F^{-1}(D)$ (Eq. 11) could be discussed. Another limit is that for this model the damage state do not evolves during unloading. One can see on **Fig. 8** that during unloading a constant permeability is predicted.

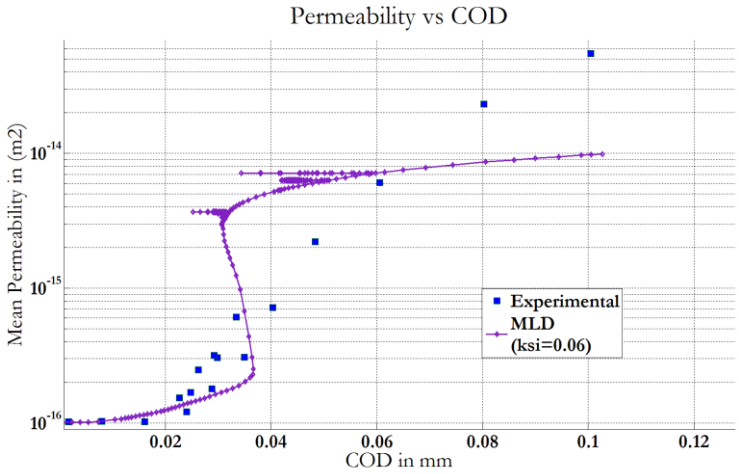


Figure 8: Mean permeability calculated by applying Matching law based on the first approach versus COD of point P.

The evolution of the mean permeability calculated by the proposed method as well as a comparison between the result of the two approaches are presented in **Fig. 9**. The coefficient ξ is supposed to be constant for any COD and is taken as calculated from the physical experiment at COD=0.1 mm equal to 0.06. This result shows that the mean permeability is very well predicted at the COD=0.1 mm where the coefficient ξ were calculated, however it is overestimated for lower CODs. The issue of constant permeability when unloading is resolved since the equivalent strain decreases and unlike the damage the equivalent strain do not stabilize but continue to increase with the crack opening.

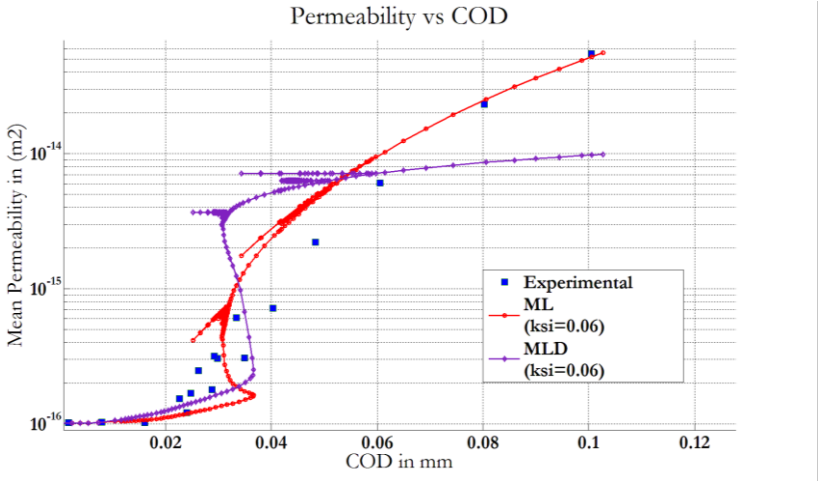


Figure 9: Mean permeability calculated by applying Matching law based on Damage MLD and by the proposed one based on equivalent strains ML versus COD of point P.

6 CONCLUSIONS

The mechanical simulation of the physical experiment of a splitting test performed on a mortar specimen were done using the NLSB model. Two approaches were considered in order to do the permeability-mechanical state coupling. A first approach that was proposed by Pijaudier-Cabot et al that is based on damage. The second one proposed in this article based on the regularized equivalent strain. The study has shown that the crack opening of an element cannot be directly related to its damage because the damage stabilize around 0.99 and do not evolves during unloading. It has been shown also that the crack opening can be related to the regularized equivalent strain since the prediction of the structural permeability using this approach is good compared to the experimental measurements.

REFERENCES

[1] Choinska M, Khelidj A, Chatzigeorgiou G, Pijaudier-Cabot G. Effects and interactions of temperature and stress-level related damage on permeability of concrete. CemConcr Res 2007;37(1):79–88.

[2] Picandet V, Khelidj A, Bastian G. Effect of axial compressive damage on gas permeability of ordinary and high-performance concrete. CemConcr Res 2001;31(11):1525–32.

[3] Pijaudier-Cabot G.; Dufour F.; Choinska M., Permeability due to the Increase of Damage in Concrete: From Diffuse to Localized Damage Distributions, journal of engineering mechanics - ASCE Volume: 135 Issue: 9 Pages: 1022-1028 (2009).

[4] Giry C., Dufour F., Mazars J. (2011), Stress-based nonlocal damage model International Journal of Solids and Structures, Vol. 48, Issues 25–26, 15, Pages 3431-3443.

[5] Dufour F. (2007); Contributions à la modélisation numérique de la fissuration des structures en béton avec prise en compte du fluage et estimation de la permeabilité.

[6] F. Dufour, G. Legrain, G. Pijaudier-Cabot and A. Huerta, "Estimate of crack opening from a 2D continuum-based FE computation", Int. J. for Num. and Anal. Meth.inGeomech, http ://dx.doi.org/10.1002/nag.1097 (2012).

[7] Rodríguez-Ferran, A., Huerta, A., (2000) Error estimation and adaptivity for nonlocal damage models. International Journal of Solids and Structures 37 (48-50),7501-7528.

[8] Pegon, P., Anthoine, A., (1997) Numerical strategies for solving continuum damage problems with softening: Application to the homogenization of masonry. Computers & Structures 64, 623 - 642.

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis(Eds)

SIMULATION OF CONSOLIDATION IN LARGE STRAINS: A COMPARISON BETWEEN FINITE ELEMENT METHOD AND MATERIAL POINT METHOD

N. Spiezia†, F. Ceccato†*, V. Salomoni†, P. Simonini†

†Department of Civil, Environmental and Architectural Engineering (DICEA)
University of Padua
Via Marzolo 9, 35131 Padova, Italy
e-mail: francesca.ceccato@dicea.unipd.it

Key words: Large strains, consolidation, Finite Element, Material Point Method

Abstract. The numerical simulation of a consolidation process undergoing large strains is a challenging task that requires the formulation and the solution of the coupled solid-deformation/fluid-diffusion problem within a changing geometry. Deformations can be rigorously taken into account with the classical Finite Element Method (FEM) based on continuum mechanics of porous media at finite strains. Alternatively, several innovative methods (SPH, MPM, PFEM, MLPG, etc.), which have been specifically developed to simulate large deformations, can be used. Among them, the Material Point Method (MPM) has been recently grown in popularity. With the MPM, large deformations are simulated with material points (MP) moving through a fixed mesh. The MP trace all the properties of the continuum (mass, velocity, stress, strain as well as external loads), while the mesh is used to solve the governing equations, but does not store any permanent information thus it can be redefined at the end of each time step, preventing problems of element distortions. The aim of this work is to investigate the analogies and the differences in terms of theoretical formulation and numerical results between FEM and MPM in consolidation processes undergoing large deformations. In fact, the simulations of the one-dimensional consolidation process in case of small deformations demonstrate that the two methods give identical results, which are also in agreement with the analytical solution demonstrated by Terzaghi. On the contrary, differences are observed in the case of large deformations. The results obtained with the two formulations are compared and discussed, enlightening the most probable source of differences.

1 INTRODUCTION

The numerical simulation of consolidation processes is a fundamental task, which can have important implications in the design of geotechnical structures. The simulation

of these phenomena, which involve the coupled interaction between the solid and the fluid phases, becomes particularly challenging when the solid phase is subjected to large strains, introducing a source of geometric non linearity in the formulation. There are many classical geotechnical applications where the finite deformations effects on a consolidation process could critically influence the results of a numerical analysis. For example, the consolidation in highly compressible clays, or the consolidation under a tower, due to the $P - \delta$ effect [1].

The consolidation process can be accurately simulated using the classical coupled finite element formulation, based on continuum theory of mixtures, considering the soil-water mixture as a two-phase continuum. Finite strains can be rigorously taken into account as an extension of the infinitesimal framework [1, 2], adopting an adequate formulation for the constitutive model. However, Updated Lagrangian Finite Element Methods (UL-FEM) encounter numerical problems for severe element distortions, with the consequent lack of convergence of the algorithm.

The problem of severe mesh distortion in UL-FEM may be circumvented by remeshing the domain or refining the regions where the elements are most deformed [3, 4]. Alternatively, some stabilization techniques have been proposed to prevent excessive element distortion [5, 6]. These strategies have the disadvantage to increase considerably the computational cost; moreover, numerical stability is still not ensured in case of extreme distortions.

To overcome the distortion problems encountered by FEM undergoing large deformations, in the last decades several innovative methods have been proposed in the literature, such as SPH [7], MPM [8], PFEM [9], MLPG [10]. In particular, the Material Point Method (MPM) has been recently grown in popularity.

The MPM belongs to the family of particle-based methods. It derives from the Particle-in-cell method (PIC) used for fluid mechanics [11]. Schreyer, Sulsky and co-workers, extended it for problems of solid mechanics [8]. Since then, the method has been widely applied to many fields of engineering and science and extended with advanced features. The application of MPM to multiphase problems is recent [12–17] and the research is in progress.

Although the MPM has been successfully applied to study coupled large deformation problems [18–20], not much work has been done to compare the numerical results obtained with the FEM based on finite strain theory and the MPM, in particular to simulate large deformation phenomena.

This work investigates the analogies and the differences in terms of theoretical formulation and numerical results between the FEM and MPM, in consolidation process undergoing large deformation. These methods give identical results in case of small deformation, coincident with the classical Terzaghi's solution. On the contrary, differences are observed in case of large deformations. The goal of this study is to identify the possible sources of differences, suggesting a discussion on the topic likely leading to future improvements of the methods.

The article is organized as follow. Section two and three briefly introduce respectively the FEM and MPM formulation. Section four presents the numerical results obtained by the two methods. One-dimensional consolidation with elastic and elasto-plastic material is solved, in case of small and large deformations. Section five draws the conclusion, discussing the differences among the two methods.

2 FINITE ELEMENT FORMULATION

This section recalls the complete set of governing equations, which allow for the solution of quasi-static deformation-diffusion boundary-value problems, such as consolidation, based on the classical mixture theory [21, 22].

Let $\boldsymbol{x} = \varphi(\boldsymbol{X}, t)$ denote the position of the solid material point X and $\boldsymbol{F} = \partial \boldsymbol{x}/\partial \boldsymbol{X}$ the associated deformation gradient of the solid matrix, with the Jacobian J such that $J = \det(\boldsymbol{F}) = \mathrm{d}v/\mathrm{d}V$. We denote the volume fraction ϕ^α of the constituent α as the ratio between its volume $\mathrm{d}V^\alpha$ divided by the total volume of the mixture $\mathrm{d}V$, that is, $\phi^s = \mathrm{d}V^s/\mathrm{d}V$ and $\phi^w = n = \mathrm{d}V^w/\mathrm{d}V$, with n denotes the porosity. Therefore,

$$\phi^s + \phi^w = 1 \tag{1}$$

The partial mass density of constituent α is given by $\rho^\alpha = \phi^\alpha \rho_\alpha$, where ρ_α is the intrinsic mass density of constituent α. This gives

$$(1-n)\rho_s + n\rho_w = \rho^s + \rho^w = \rho \tag{2}$$

where ρ is the total mass density of the mixture.

With these preliminaries, the conservation of mass equation for the solid and water phases are respectively

$$\dot{\rho}^s + \rho^s \nabla^{\boldsymbol{x}} \cdot \boldsymbol{v} = 0; \tag{3}$$

$$\dot{\rho}^w + \rho^w \nabla^{\boldsymbol{x}} \cdot \boldsymbol{v} = -\boldsymbol{\nabla}^{\boldsymbol{x}} \cdot \boldsymbol{q}, \tag{4}$$

where $\boldsymbol{\nabla}^{\boldsymbol{x}} \cdot (\cdot)$ is the divergence operator with respect to the current configuration, $\boldsymbol{v}$ is the velocity of the solid phase, $\boldsymbol{w}$ is the velocity of the water phase, $\boldsymbol{q} \equiv \rho^f \tilde{\boldsymbol{v}} = \rho^w(\boldsymbol{w} - \boldsymbol{v})$ is the Eulerian relative flow vector of the water phase with respect to the solid matrix. Adding Eqs. (3) and (4), we get the basic conservation of mass equation for the system, i.e.

$$\dot{\rho}_0 = -J \nabla^{\boldsymbol{x}} \cdot \boldsymbol{q}, \tag{5}$$

where $\rho_0 \equiv J\rho$ is the pull-back mass density of the mixture in the reference configuration.

Ignoring inertial forces, the balance of linear momentum in Lagrangian form can be written as

$$\nabla^{\boldsymbol{X}}(\boldsymbol{P}) + \rho_0 \boldsymbol{g} = \boldsymbol{0}. \tag{6}$$

where $\boldsymbol{g}$ is the gravity vector. The first Piola-Kirchhoff stress tensor $\boldsymbol{P}$ reads

$$\boldsymbol{P} = \boldsymbol{\tau} \cdot \boldsymbol{F}^{-T} = \boldsymbol{\tau}' \cdot \boldsymbol{F}^{-T} - \theta_w \boldsymbol{F}^{-T} \tag{7}$$

where $\boldsymbol{\tau} = J\boldsymbol{\sigma}$ is the symmetric total Kirchhoff stress tensor, $\boldsymbol{\tau}' = J\boldsymbol{\sigma}'$ is the effective Kirchhoff stress tensor and $\theta_w = Jp_w$ is the Kirchhoff pore pressure, in accordance with the principle of effective stress.

The links between the state of stress and the deformations and between the flow vector and the water pressure in the porous media are provided by constitutive relationships. The stresses are assumed to be a nonlinear function of the deformations via an elastoplastic constitutive response. On the other hand, the relative flow vector is related to the fluid pressure using Darcy's law, assuming constant permeability. The constitutive law for the solid phase is cast within the framework of nonlinear kinematics where the total deformation gradient is assumed to allow the multiplicative decomposition into elastic and plastic parts, $\boldsymbol{F} = \boldsymbol{F}^e \cdot \boldsymbol{F}^p$, where $\boldsymbol{F}^e$ and $\boldsymbol{F}^p$ are defined as the elastic and plastic deformation gradient, respectively. The constitutive law is given in terms of the Kirchhoff effective stress tensor $\boldsymbol{\tau}'$ and the left elastic Cauchy-Green deformation tensor $\boldsymbol{b}^e = \boldsymbol{F}^e \cdot \boldsymbol{F}^{eT}$ through the relation

$$\boldsymbol{\tau}' = 2\frac{\partial \Psi}{\partial \boldsymbol{b}^e} \cdot \boldsymbol{b}^e. \tag{8}$$

where Ψ is the strain-energy function.

The balance laws (5) and (6), together with constitutive relations, provide a complete set of governing equations, which allow for the solution of quasi-static deformation-diffusion boundary-value problems. The discretized equations are obtained following standard Galerkin procedure, approximating the nodal displacements $\boldsymbol{u}$ and pore pressure p_w by means of shape functions ($\boldsymbol{N}$, $\boldsymbol{N}^p$), and integrating implicitly over time. Hence, the two balance equations can be written as

$$\boldsymbol{H}^{ext} - \boldsymbol{H}^{int}(\boldsymbol{u}, p_w) = \boldsymbol{0} \tag{9}$$

$$\boldsymbol{F}^{ext} - \boldsymbol{F}^{int}(\boldsymbol{u}, p_w) = \boldsymbol{0} \tag{10}$$

where $\boldsymbol{H}^{int} = \int_{V_0} [\boldsymbol{N}^{pT} \Delta\rho_0 - \Delta t(J\boldsymbol{\Gamma}^T \boldsymbol{q})] \mathrm{d}V_0$ and $\boldsymbol{F}^{int} = \int_{V_0} [\boldsymbol{B}^{\boldsymbol{T}}(\boldsymbol{\tau}' - Jp_w\boldsymbol{I}) - \rho_0 \boldsymbol{N}^T \boldsymbol{g}] \mathrm{d}V_0$. For more details the reader should refer to [1, 2].

3 THE MATERIAL POINT METHOD

In the MPM, arbitrary large deformations of a body are simulated by a set of material points (MP) which move through a computational finite element mesh. The MP carries all the information of the continuum such as velocity, acceleration, stress, strain, material parameters as well as external loads. It can be regarded as an extension of the FEM, because the underlying finite element grid is used, as with the FEM, to solve the system

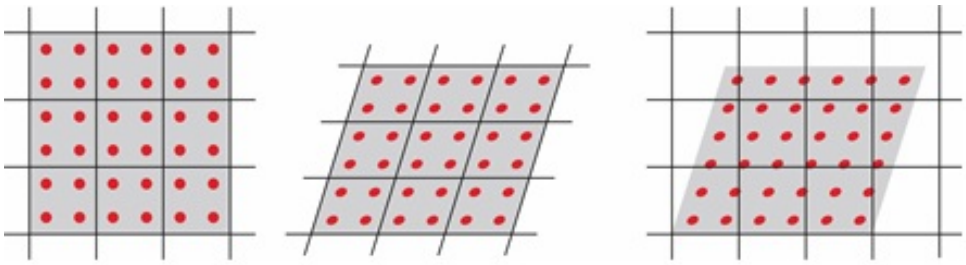


Figure 1: (left) configuration at the beginning of a time step in which the red dots are the MP; (center) incrementally deformed mesh; (right) reset mesh at the end of a time step

of equilibrium equations. However, information is mapped from nodes to MP at the end of each time step and the mesh is usually reset into its original state (Fig. 1). The mesh does not follow the deformations of the body as in the FEM, thus preventing problems of element distortion.

An available dynamic MPM code is used in this study. The primary unknown variables of the implemented two-phase formulation are the solid velocity $\boldsymbol{v}$ and fluid velocity $\boldsymbol{w}$. This formulation follows [23] and has been preferred for the adopted MPM because it is able to properly capture the dynamic response for general loading [24]. The $(\boldsymbol{v} - \boldsymbol{w})$ formulation is derived from the momentum equations of the water phase and the soil-water mixture as follows. The momentum equation of the water phase is

$$\rho_w \dot{\boldsymbol{w}} + ((n\gamma_w)/k)(\boldsymbol{v} - \boldsymbol{w}) = \nabla p_w + \rho_w \boldsymbol{g} \tag{11}$$

where γ_w is the unit weight of the water and k is the Darcy's permeability; the latter is assumed constant. The second term on the left hand side represents the interaction between solid and fluid.

The momentum equation for the mixture is

$$(1-n)\rho_s \dot{\boldsymbol{v}} + n\rho_w \dot{\boldsymbol{w}} = \nabla \cdot (\boldsymbol{\sigma'} + \boldsymbol{I} p_w) + \rho \boldsymbol{g} \tag{12}$$

where $\boldsymbol{\sigma'}$ is the effective stress and $\boldsymbol{I} = [1, 1, 1, 0, 0, 0]$.

The excess pore pressure increment can be calculated from the mass balance equation for the water phase

$$\dot{p}_w = K_w/n[(1-n)\nabla \cdot \boldsymbol{v} + n\nabla \cdot \boldsymbol{w}] \tag{13}$$

The effective stress increment is calculated with the soil constitutive model as

$$\dot{\boldsymbol{\sigma}}' = \boldsymbol{D}\dot{\boldsymbol{\epsilon}} \tag{14}$$

where $\boldsymbol{D}$ is the constitutive tensor and $\dot{\boldsymbol{\sigma}}'$ is the Jaumann stress rate tensor defined as

$$\dot{\boldsymbol{\sigma}}' = \boldsymbol{\sigma'} - \boldsymbol{W} \cdot \dot{\boldsymbol{\sigma}}' + \dot{\boldsymbol{\sigma}}' \cdot \boldsymbol{W}^{\boldsymbol{T}} \tag{15}$$

in which $\dot{\boldsymbol{\sigma}}$ is the Cauchy stress and $\boldsymbol{W}$ is the spin tensor.

Again, the discretized equations are obtained by deriving the weak form of the momentum conservation, using the Galerkin procedure, and approximating the velocity field by means of shape functions ($\boldsymbol{N}$) [16]. They can be written as

$$\boldsymbol{M_w}\dot{\boldsymbol{w}} = \boldsymbol{F_w^{ext}} - \boldsymbol{F_w^{int}} - \boldsymbol{F_w^{drag}} \tag{16}$$

$$\boldsymbol{M_s}\dot{\boldsymbol{v}} + \bar{\boldsymbol{M}}_{\boldsymbol{w}}\dot{\boldsymbol{w}} = \boldsymbol{F^{ext}} - \boldsymbol{F^{int}} \tag{17}$$

where the subscripts s and w indicate the soil and water phase respectively; no subscript indicates that the quantity belongs to the mixture.

$\boldsymbol{F_w^{drag}} = \int_V n\gamma_w k^{-1}\boldsymbol{N}^T\boldsymbol{N}dV(\boldsymbol{v}-\boldsymbol{w})$ denotes a drag force computed from the relative water velocity $(\boldsymbol{w}-\boldsymbol{v})$ which takes into account the solid-fluid interaction. The mass matrices for the fluid and the soil skeleton are defined as: $\boldsymbol{M}_w = \int_V \rho_w \boldsymbol{N}^T\boldsymbol{N}dV$ and $\boldsymbol{M_s} = \int_V (1-n)\rho_s\boldsymbol{N}^T\boldsymbol{N}dV$. Matrix $\bar{\boldsymbol{M}}_{\boldsymbol{w}}$ is formed using the density $n\rho_w$ in place of ρ_w. For numerical implementation, the lumped mass matrices are used. The internal forces are calculated as $\boldsymbol{F_w^{int}} = \int_V \boldsymbol{B^T I}p_w dV$ and $\boldsymbol{F^{int}} = \int_V \boldsymbol{B^T}\boldsymbol{\sigma}dV$ where $\boldsymbol{\sigma}$ is the total stress. For more details the reader should refer to [16].

In the present study the explicit Euler-Cromer scheme is used. This means that the acceleration is calculated explicitly and the velocity is updated from it implicitly.

The MPM solution procedure follows [25]. The acceleration of the fluid is calculated by solving Eq. (16). It is subsequently used to obtain the acceleration of the solid from Eq. (17). The velocities and the momentum of the MP are updated from the nodal accelerations of each phase. The nodal velocities are then calculated from the nodal momentum and used to compute the strain rate ($\dot{\boldsymbol{\epsilon}}$) at the material point location. The excess pore pressure increment can be calculated from Eq. (13). The effective stress increment is calculated with the soil constitutive model (Eq. 14). The displacement and position of each MP is updated according to the velocity of the solid phase.

4 NUMERICAL RESULTS

The consolidation problem is studied considering a 1m-long column. The head of the column is permeable and the bottom is impermeable, therefore the water can flow out of the column from the top surface and the drainage length h is 1m. Firstly, the case of small deformations is studied and the numerical solution is compared to the Terzaghi's analytical solution. Secondly, large deformations are taken into account. Thirdly, the plastic behavior of soil is considered using the Modified Cam Clay model [26].

In case of small strain, a linear elastic material with Young's modulus $E = 10000kPa$, and Poisson's ratio $\nu = 0$ is used. The permeability is $k = 10^{-3}m/s$. A load of $10kPa$ is applied at the head of the column. In the FEM model the column is discretized with 40 quadrilateral elements, with 9 nodes for the displacements unknowns and 4 nodes for the pressure unknowns, integrated over 9 Gauss Point. On the other hand, in the MPM model the column is discretized with 40 rows of 6 tetrahedral elements containing 4 MP each.

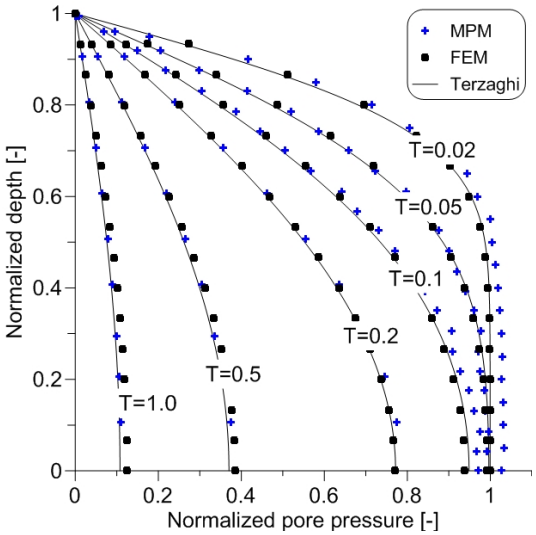


Figure 2: Normalized pore pressure along depth, comparison between numerical and analytical solution.

A non-dimensional time factor can be defined as:

$$T = \frac{c_v t}{h^2} \tag{18}$$

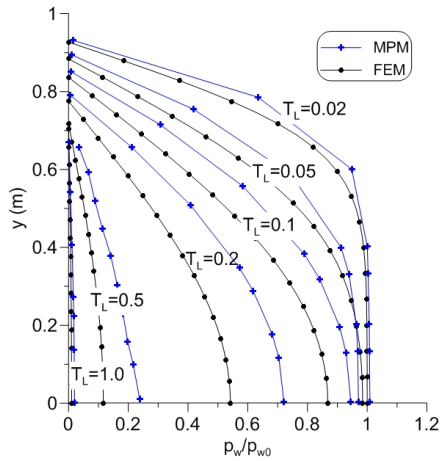
Figure 2 plots the normalized pore pressure p_w/p_{w0} ($p_{w0} = 10kPa$) against the normalized depth y/h as function of the non-dimensional time. The numerical solutions obtained with FEM and MPM are in perfect agreement with the Terzaghi's analytical solution.

The case of large deformations is considered assuming $E = 100kPa$ and applying a load of $\sigma_y = 50kPa$. Figure 3a shows the normalized pore pressure distribution at different normalized time T_L, which calculated with Eq. 18 where h is the initial length of the column. Regarding the pore pressure evolution with time, the MPM predicts higher pore pressures than the FEM, especially for time factors between 0.05 and 0.5. The maximum difference is about 20% at the bottom of the column. Some differences are also observed in the decrease of column height during time (Fig. 3b). In particular, the final column high obtained with the MPM is 0.674m and with the FEM is 0.704m.

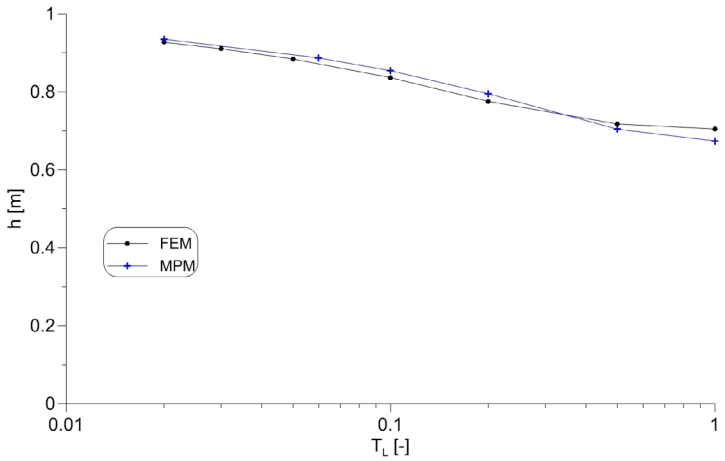
The plastic behavior of soil is considered using the modified Cam Clay model with the input parameter summarized in Table 1. The adopted FEM implements the modified Cam Clay model proposed in [27]; the reader should refer to this work for a correct interpretation of the input parameters of the model.

The material is assumed normally consolidated; the initial vertical and horizontal effective stresses are 50kPa. A vertical load of 150kPa is applied. In this case, the consolidation coefficient used to compute the time factor is calculated as:

$$c_v = \frac{k(1+e_0)\sigma'_{v0}}{\lambda\gamma_w} \tag{19}$$

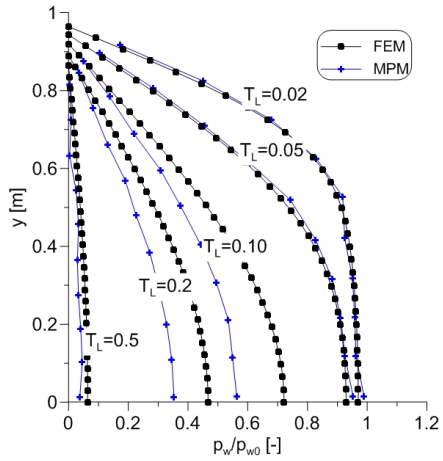


(a) Normalized pore pressure along depth.

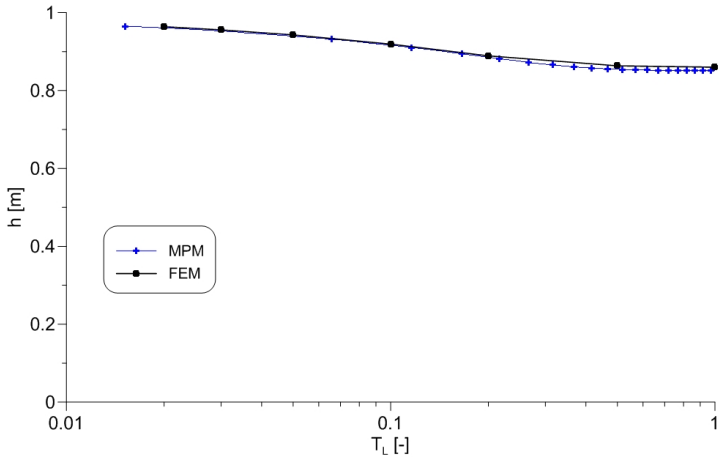


(b) Column height as function of the time factor T_L.

Figure 3: Large strain consolidation with linear elastic material; comparison between MPM and FEM solutions.



(a) Normalized pore pressure along depth.



(b) Column height as function of the time factor T_L.

Figure 4: Large strain consolidation with elastoplastic material; comparison between MPM and FEM solutions.

Material property	Symbol	Value
Virgin compression index [-]	λ	0.3
Recompression index [-]	κ	0.04
Effective poisson ratio [-]	ν'	0.25
Slope of CSL in p-q plane [-]	M	0.92
Initial void ratio [-]	e_0	1.0
Darcy's permeability [m/s]	k	$3 \cdot 10^{-3}$

Table 1: Material properties used in CPT analyses with MCC material model

where σ'_{v0} is the initial vertical effective stress.

Figure 4 compares the results obtained with FEM and MPM in terms of pore pressure distribution and column height. With regard to the pore pressure, better agreement is obtained compared to the elastic case. A difference of about 20% is observed for $T_L = 0.10$ and $T_L = 0.2$, while the pore pressure distribution for the other time factors are almost coincident. This peculiar result will be deeply investigated in the near future. The small differences in terms of pressure do not affect significantly the evolution of vertical displacements with time. Both method predict a final settlement of 0.142m.

5 DISCUSSION AND CONCLUSIONS

This paper compares the fully coupled FEM and the MPM in the simulation of one dimensional consolidation undergoing large strains.

Both method, even though adopting different primary unknowns (namely $\boldsymbol{u} - p_w$ and $\boldsymbol{v} - \boldsymbol{w}$), can describe efficiently the solid-displacement/fluid-diffusion process, typical of consolidation problem. Identical results are obtained with the two methods for small deformations. This observation is in agreement with the comparison made in [28] (for harmonic loading and small strain) of the two different ways of coupling solid and fluid interaction. Furthermore, the choice of different primary unknowns involves also the adoption of different time integration algorithms. In fact, the quasi-static FEM adopts an implicit time integration scheme; on the contrary, the dynamic MPM uses an explicit scheme. However, the adopted time increments ensure a sufficiently accurate solution for both methods.

In case of large deformations, the displacements predicted by the two method are in good agreement. However, discrepancies are observed in the evolution of pore pressure distribution.

Probably, the major source of differences between the MPM and the adopted FEM lies in the way large deformations are treated, in particular with respect to the constitutive equations. In MPM the hypothesis of small strain is assumed within the time increment, thus allowing the use of the Jaumann objective stress rate. This formulation restricts the validity of the rate constitutive equation to small elastic strains [29]. As regard the

FEM, the formulation of the constitutive law is based on hyperelasticity combined with the multiplicative decomposition of the deformation gradient. This formulation is more accurate for the development of large elastic strains, and also circumvents the so-called rate issue in finite deformation analysis [27]. The introduction of plasticity reduces the differences between the two methods, because if the elastic strains remain small the results based on the Jaumann rate should be close to results of formulations that are based on the multiplicative decomposition [29].

These considerations do not probably fully explain the differences observed in the pore pressure distributions. Further investigations, focusing on the diffusion problem combined with large deformations, should be carried out, possibly comparing the results with experimental data and also with available analytical solution such as [30].

This paper doesn't accomplish all the aspects in the comparison of the two numerical methods. Since considerable effort has been devoted in the last years to large strain coupled formulation in the MPM, with promising outcomes, a deeper investigation in the results should be carried out. The research should consider bidimensional and three dimensional problems, transient boundary conditions, cyclic loads, etc. In particular, simulations leading to a significant element distortions should be investigated, comparing FEM (enhanced with remeshing techniques) and MPM.

AKNOWLEDGMENTS

The authors thank the MPM research group of Deltares, Delf, The Netherlands for the kind support in the use of the MPM code.

REFERENCES

[1] Borja, R. I., and Alarcon, E., A mathematical framework for finite strain elasto-plastic consolidation. Part 1: balance laws, variational formulation, and linearization *Computer Methods in Applied Mechanics and Engineering* (1995) **122**:145–171.

[2] Borja, R. I., Tamagnini, C., Alarcon, E., Elastoplastic consolidation at finite strain. Part 2: finite element implementation and numerical examples *Computer Methods in Applied Mechanics and Engineering* (1998) **159**:103–122.

[3] Kato, K., Lee N. S., and Bathe, K. J. Adaptive finite element analysis of large strain elastic response *Computers and Structures* (1993) **47**:829–855.

[4] Lee, N.-S. and Bathe, K.-J. Error indicators and adaptive remeshing in large deformation finite element analysis *Finite Elements in Analysis and Design* (1994) **16**:99–139.

[5] Simo, J.C., Armero, F. and Taylor, R.L. Improved versions of assumed enhanced strain tri-linear elements for 3D finite deformation problems *Computer Methods in Applied Mechanics and Engineering* (1993) **110**:359–386.

[6] Nagtegaal, J.C. and Fox, D.D. Using assumed enhanced strain elements for large compressive deformation *International Journal of Solids and Structures* (1996) **33**:3151–3159.

[7] Lucy, L. B. A numerical approach to the testing of the fission hypothesis.*The astronomical journal* (1977) **82**, 1013–1024.

[8] Sulsky, D. Chen, Z. and Schreyer, H.L. A particle method for hystory-dependent materials*Computer Methods in Applied Mechanics and Engineering* (1994) **118**: 179–196

[9] Idelsohn, S., E. Oñate, and F. D. Pin. The particle finite element method: a powerful tool to solve incompressible fows with free-surfaces and breaking waves. *International Journal for Numerical Methods in Engineering* (2004) **61** (7), 964–989.

[10] Atluri, S. and T. Zhu. A new meshless local petrov-galerkin (MLPG) approach in computational mechanics. *Computational mechanics* (1998) **22** (2), 117–127.

[11] Harlow, F. H. *A Machine Calculation Method for Hydrodynamic Problems.* Los Alamos Scientific Laboratory report LAMS-1956.(1955)

[12] Zhang, H., K. Wang, and Z. Zhang. Material point method for numerical simulation of failure phenomena in multiphase porous media. *Computational Mechanics* (2007) p.p. 36–47. Springer.

[13] Zhang, D. Z., Q. Zou, W. B. VanderHeyden, and X. Ma. Material point method applied to multiphase flows. *Journal of Computational Physics* (2008) **227** (6): 3159–3173.

[14] H. W. Zhang, K. P. Wang, and Z. Chen. Material point method for dynamic analysis of saturated porous media under external contact/impact of solid bodies. *Computer Methods in Applied Mechanics and Engineering* (2009) **198**.17 : 1456-1472.

[15] Higo, Y., et al. A coupled MPM-FDM analysis method for multi-phase elasto-plastic soils. *Soils and foundations* (2010) **50**.4:515-532.

[16] Jassim, I. and Stolle, D. and Vermeer, P.Two-phase dynamic analysis by material point method. *International Journal for Numerical and Analytical Methods in Geomechanics* (2013) **37**:2502–2522

[17] Abe, K., K. Soga, and S. Bandara. Material Point Method for Coupled Hydromechanical Problems. *Journal of Geotechnical and Geoenvironmental Engineering* (2013) **140**.3.

[18] Alonso, E.E. and Zabala, F. Progressive failure of Aznalcóllar dam using the material point method, *Géotechnique* (2011) **61** 9: 795–808

[19] Bandara, S.; Soga, K. Coupling of soil deformation and pore fluid flow using material point method. *Computers and Geotechnics* (2015) **63** 199-214.

[20] Ceccato, F., Beuth, L., Vermeer, P. A., Simonini P. Two-phase Material Point Method applied to cone penetration for different drainage conditions. *IS on Geomechanics form micro to macro* (2014) CRC Press - Taylor and Francis Group

[21] Bowen, R.M. *Theory of mixtures*, in: A.C. Eringen (Ed.), Continuum Physics, Volume IIIMixtures and EM Field Theories Academic Press, (1976), New York, NY.

[22] Atkin, R.J. and Craine, R.E. Continuum theories of mixture: basic theory and historical development, *Q. J. Mech. Appl. Math.* (1976) **29**:209-244.

[23] Verruijt, A. *Soil dynamics*. Technische Universiteit, Faculteit Civiele Techniek, 1996.

[24] Van Esch, J., D. Stolle and I. Jassim. Finite element method for coupled dynamic flow-deformation simulation. In 2nd International Symposium on Computational Gaomechanics (COMGEO II), Cavat-Dubrovnik, Croatia (2011).

[25] Al-Kafaji, Issam K. J. *Formulation of a Dynamic Material Point Method (MPM) for Geomechanical Problems*, PhD thesis, University of Struttgart, Germay, 2013

[26] Schofield, A. and P. Wroth. *Critical state soil mechanics*. McGraw-Hill London, 1968.

[27] Borja, R. I., and Tamagnini, C. ,Cam-Clay plasticity Part III: Extension of the infinitesimal model to include finite strains *Computer Methods in Applied Mechanics and Engineering* (1998) **155**:73–95.

[28] Zienkiewicz, O. C., C. T. Chang, and P. Bettess. Drained, undrained, consolidating and dynamic behaviour assumptions in soils. *Geotechnique* (1980) **30**.4 : 385-395.

[29] De Borst, R., Crisfield, M. A., Remmers, J. JC., Verhoosel, C. V. *Nonlinear finite element analysis of solids and structures*. John Wiley & Sons, 2012

[30] Xie, K. H.; Leo, C. J. Analytical solutions of one-dimensional large strain consolidation of saturated and homogeneous clays. *Computers and geotechnics* (2004) **31**.4: 301-314.

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis(Eds)

THERMO-MECHANICAL PROPERTIES OF MAGNESIA CARBON FOAM COMPOSITES

A. Jung* **and S. Diebels***

*Lehrstuhl für Technische Mechanik
Universität des Saarlandes
Campus A4 2, 66123 Saarbrücken, Germany
e-mail: anne.jung@mx.uni-saarland.de, web page: http://www.ltm.uni-saarland.de

Key words: Thermo-mechanical Coupling, Multi-physics Problem, Multi-phase, Cellular Refractory, Carbon Foam

Abstract. Refractory materials have a wide range of applications in steel-making industry. Often, magnesia carbon bricks (MgO-C) are used. These consist of a periclase phase (MgO) with inclusions of carbon and gas filled pores. The thermo-mechanical properties of MgO-C composites could significantly be improved using cellular MgO-C composites based on carbon foams. Modelling of MgO-C composite foams is not only a multi-phase, but also a multi-physics problem, in which both the displacement field and the temperature field have to be taken into account. In the present contribution, a fully coupled phenomenological thermo-mechanical continuum model was developed. The theory of porous media (TPM) with a kinematic coupling of the displacement and temperature fields of all constituents was used. Linear thermoelasticity with a multiplicative decomposition of the deformation gradient into an elastic and a thermal part for isotropic materials was extended to the mixture of MgO and C phase. The total macroscopic stress was calculated using the theory of mixture, including the contributions from the pore pressure.

1 INTRODUCTION

Refractory materials have a wide range of applications in steel-making industry as lining of furnaces, oxygen converters or for ladles. Common refractories are made of a periclase phase (MgO) with inclusions of carbon and gas filled pores. In their applications, refractories are subjected to thermal and mechanical loads causing damage. There are two possible types of thermally induced stress. Permanent thermal stress arises only in heterogeneous materials due to the coupling of materials having different coefficients of thermal expansion (CTE). The corresponding thermal damage results from the isotropic expansion of both phases. Temporary thermal stresses emerge by temperature gradients

in homogeneous and also in heterogeneous materials causing mechanical damage and vanish for elastically loaded materials at thermal equilibrium. Mechanical damage occurs not only in multi-phase materials with differences in their CTE but also in homogeneous single phase materials. In the quasi-static heating of a furnace, no temperature gradients arise, hence the only possible thermal stresses are permanent thermal stresses due to the coupling of materials with a mismatch in their CTEs.
The next generation of carbon containing refractories should show improved characteristics in terms of microstructural design, material processing, and material sustainability. In recent studies, the thermo-mechanical properties of MgO-C composites could significantly be improved using cellular MgO-C composites based on carbon foams [1]. Foams consist of three different hierarchical levels. The macro scale deals with whole components, the meso scale with several pores and the micro level comprises single struts. In experiments and modelling, foams can be described on these different scales.
Modelling of MgO-C composite foams is not only a multi-phase and multi-physics problem, but also a multi-scale problem. In previous work, a mesoscopic elastic thermo-mechanically coupled model was used to investigate the structure-property relationship of MgO-C hybrid foams, in order to reduce thermally induced stresses and accompanying damage [2]. In the present contribution, the thermo-mechanical behavior of cellular ceramic composites will be modelled by a new multiphase approach of porous media using a fully coupled phenomenological thermo-mechanical continuum model. The key assumption is the use of the theory of porous media (TPM) with a kinematic coupling of the displacement and temperature fields of all constituents.

2 KINEMATICS FOR MgO-C FOAMS BY A MULTIPHYSICS APPROACH

Figure 1 shows the typical repesentative elementary volume (REV) of the micro structure of a common MgO-C brick and of a cellular MgO-C foam. Whereas common MgO-C bricks consist of a carbon matrix with inclusions of MgO and irregular pores, the new MgO-C foams consit of a carbon skeleton and more or less circular pores. The carbon skeleton is coated with a layer of MgO. There are several ways to model such problems. On the one hand, each constituent can be modeled according a single phase theory. Due to the coupling and contact of the constituents, this is very expensive. For each micro structure a different approach is needed. On the other hand, both micro structures can be described and modelled using the TPM by homogenization of the mixture and solving the problem as smeared out multiphase continuum. The theory of mixture according to Truesdell & Topin [3] uses a continuum with statistically distributed, immiscible constituents, but there is not a volumetric measure of the constituents included. Hence, for porous media, the concept of volume fractions can be introduced [4].

According to the concept of volume fractions, the volume V of the body $\mathcal{B}$ is the sum of the partial volumes V^β of the constituents φ^β in the body $\mathcal{B}$. In the case of the MgO-C foams, a volume element $\mathrm{d}v$ of the mixtures consists of the sum of the partial volume elements $\mathrm{d}v^\beta$ of all phases φ^β. Here, the three constituents carbon (φ^1), magnesia (φ^2)

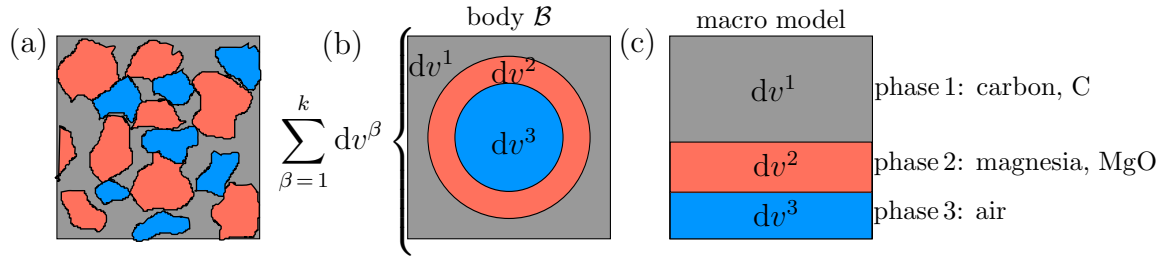


Figure 1: REV of a common MgO-C brick (a) and of a new cellular MgO-C foam (b) and substitute model according to the concept of volume fractions (c).

and pore gas (φ^3) are taken into account.

$$\mathrm{d}v = \mathrm{d}v^1 + \mathrm{d}v^2 + \mathrm{d}v^3 = \sum_{\beta=1}^{k} \mathrm{d}v^\beta \qquad \text{for} \quad k = 1, 2, 3 \tag{1}$$

$$V = \int_{\mathcal{B}} \mathrm{d}v = \sum_{\beta=1}^{k} V^\beta \qquad \text{with} \quad V^\beta = \int_{\mathcal{B}} \mathrm{d}v^\beta = \int_{\mathcal{B}} n^\beta \, \mathrm{d}v \qquad \text{and} \quad n^\beta = \frac{\mathrm{d}v^\beta}{\mathrm{d}v} \tag{2}$$

The volume fraction n^β is the local quotient of the volume element $\mathrm{d}v^\beta$ of φ^β refered to the volume element $\mathrm{d}v$ of the mixture. The volume fractions of all constituents must fulfill the saturation condition, whereas the sum of all volume fractions is equal one. Hence, the volume fraction of one phase can be calculated from the volume fractions of the other phases

$$\sum_{\beta=1}^{k} n^\beta = 1 \qquad \text{with} \quad k = 1, 2, 3 \qquad \text{hence,} \quad n_1 + n_2 + n_3 = 1. \tag{3}$$

The REV of the MgO-C foam can be substituted in a homogenised sense by the concept of volume fractions according to Figure 1 (c). For multiphase materials, two different densities can be defined. The real density $\rho^{\beta R} = \mathrm{d}m^\beta / \mathrm{d}v^\beta$ is the local quotient of the local mass element $\mathrm{d}m^\beta$ and the local volume element $\mathrm{d}v^\beta$. Whereas, the partial density $\rho^\beta = \mathrm{d}m^\beta / \mathrm{d}v$ is the quotient of the local mass element $\mathrm{d}m^\beta$ and the volume element of the mixture $\mathrm{d}v$. Both densities are coupled by the volume fraction

$$\rho^\beta = n^\beta \, \rho^{\beta R}. \tag{4}$$

It is assumed, that the two solid phases (phase (1), C and phase (2), MgO) of the cellular refractories are materially incompressible but due to the compressibility of the pore phase (phase (3), air), they possess a structural compressibility in the macro model. This leads to the geometrical constraint, that the real densities $\rho^{\beta R}$ of the solid phases

are constant. Due to the compressibility of the pore gas, the mixture is also compressible. Carbon and MgO have a changing partial density ρ^β, since the volume fractions n^β and hence the partial densities change with compression of the mixture. The total volume reduces by compression, the pore volume vanishs at the point of compaction, but the solid body mass m^β stays constant. As a result, the partial density ρ^β increases.

$$\rho^{\beta R} = \text{const.} \quad \beta = 1, 2 \quad \text{and} \qquad \rho^\beta = \frac{\mathrm{d}m^\beta}{\mathrm{d}v} \neq \text{const.} \tag{5}$$

The TPM combines the theory of mixture with the concept of volume fractions and assumes that all constituents are statistically distributed in the REV leading to a statistic substitute model for the mixture with superimposed and interacting constituents [5]. The spatial point $\mathbf{x}$ of the current configuration is at the same time occupied by material points of all constituents φ^β. Material points of individual constituents follow their own motion function $\boldsymbol{\chi_\beta}$ and hence, originate from different initial positions $\mathbf{X}^\beta$.
In contrast to the TPM, in our multiphase approach of porous media, it is assumed, that individual phases do not interpenetrate. In these modified superimposed continua for MgO-C foams, material points of the three phases (C, MgO, air) originate from the same initial position $\mathbf{X}$, follow the same motion function $\boldsymbol{\chi}$ and occupy the same point $\mathbf{x}$ at the time t (see Figure 2).

$$\mathbf{x} = \boldsymbol{\chi_\beta}(\mathbf{X}_\beta, t) = \boldsymbol{\chi}(\mathbf{X}, t) \qquad \forall\, \beta \tag{6}$$

$$\text{with} \quad \mathbf{X}_\beta = \boldsymbol{\chi_\beta}(\mathbf{X}_\beta, t_0) = \boldsymbol{\chi}(\mathbf{X}, t_0) = \mathbf{X} = \boldsymbol{\chi}^{-1}(\mathbf{x}, t) \tag{7}$$

For the velocity refered to the reference configuration, one has

$$\mathbf{v}_\beta(\mathbf{X}_\beta, t) = \frac{\partial \boldsymbol{\chi_\beta}(\mathbf{X}_\beta, t)}{\partial t} = \frac{\partial \boldsymbol{\chi}(\mathbf{X}, t)}{\partial t} \tag{8}$$

And with respect to the current configuration, one obtains

$$\mathbf{v}_\beta(\mathbf{x}, t) = \mathbf{x}'_\beta\left(\boldsymbol{\chi}_\beta^{-1}(\mathbf{x}, t), t\right) = \mathbf{x}'_\beta(\mathbf{x}, t) \tag{9}$$

As a result, the velocities $\mathbf{v}_\beta$ of all constituents are equal $\mathbf{v} = \mathbf{v}_\beta = \mathbf{v_1} = \mathbf{v_2} = \mathbf{v_3}$.
The deformation gradient $\mathbf{F}_\beta$ with respect to a motion $\boldsymbol{\chi_\beta}$ of a constituent φ^β is

$$\mathbf{F}_\beta = \frac{\partial \boldsymbol{\chi_\beta}(\mathbf{X}_\beta, t)}{\partial \mathbf{X}_\beta} = \operatorname{Grad}_\beta \boldsymbol{\chi_\beta} = \operatorname{Grad} \mathbf{x} = \mathbf{F}. \tag{10}$$

According to the assumption that the motion functions of all constituents φ^β are equal, the deformation gradient and the Jacobi determinant of all phases are equal, as well.

$$\mathbf{F} = \mathbf{F}_\beta \quad \text{and} \quad J_\beta = \det \mathbf{F}_\beta = \det \mathbf{F} \qquad \text{with} \qquad \beta = 1, 2, 3 \tag{11}$$

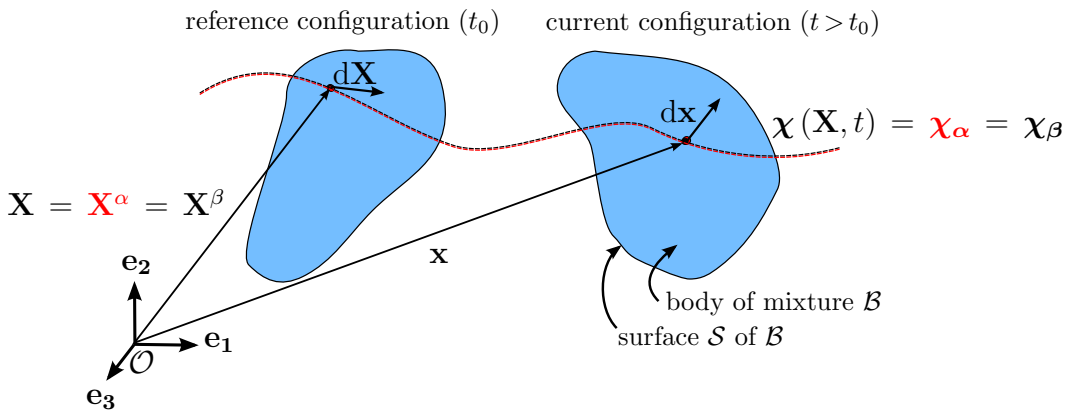


Figure 2: Kinematics for the modified superimposed continua of MgO-C foams

3 BALANCE EQUATIONS FOR MgO-C FOAMS

Balance equations for mixtures can be deduced from Truesdells three methaphysic principes [3]. All properties of the mixture must be mathematical consequences of the properties of all constituents. To describe the motion of one constituent, conceptually, one can separate this constituent from the residual constituents of the mixture, on condition that the influence of the other constituents is considered. By the introduction of the production terms in the balance equations, the interaction between the constituents φ_β can be described. And finally, the motion of the mixture is governed by the same equations as the motion of an one-phase material. The general local form of balance equations for one phase materials is [5]

$$\dot{\boldsymbol{\psi}}_{\boldsymbol{\beta}} + \boldsymbol{\psi} \operatorname{div} \dot{\mathbf{x}} = \operatorname{div} \boldsymbol{\phi} + \boldsymbol{\sigma} + \hat{\boldsymbol{\psi}} \tag{12}$$

with $()^{\bullet}$ representing the total time derivative with respect to the baroccentric velocity. Hence, the local form of the balance equations for the various constituents φ^β is

$$(\boldsymbol{\psi}_{\boldsymbol{\beta}})'_{\beta} + \boldsymbol{\psi}_{\boldsymbol{\beta}} \operatorname{div} \mathbf{x}'_{\beta} = \operatorname{div} \boldsymbol{\phi}_{\boldsymbol{\beta}} + \boldsymbol{\sigma}_{\boldsymbol{\beta}} + \hat{\boldsymbol{\psi}}_{\boldsymbol{\beta}}, \tag{13}$$

whereas $(\cdot)'_\beta$ is the material time derivative, which describes the changes of Ψ that an observer experienced who moved with the material point $\mathbf{X}_\beta$.
Before deducing the individual balance equations for the mixture and the various constituents, there is a short summary of all assumptions made up to now and further assumptions for the evaluation of the balance equations.

[A 1] Concept of volume fractions and superimposed continua

[A 2] Material incompressibility but structural compressibility for phase (1), C & (2),MgO

[A 3] Material compressibility for phase (3), the pore fluid, air

[A 4] Superimposed continua with equal motion functions $\boldsymbol{\chi}_{\boldsymbol{\beta}} = \boldsymbol{\chi}$

[A 5] For each point in time, the constituents occupy the same place

[A 6] Truesdells metaphysical principle

[A 7] No mass exchange between the various constituents

3.1 Mass balances

The local form of the mass balance of the mixture is

$$\dot{\rho} + \rho \operatorname{div} \dot{\mathbf{x}} = 0 \quad \text{yielding} \quad \rho = \rho_0 \det \mathbf{F}^{-1} \qquad \text{with} \quad \rho_0 : \quad \text{reference density} \tag{14}$$

Because of the fact, that the pore fluid is compressible, the mixture is compressible as well, although the phase (1) and phase (2) are incompressible.

The partial mass balances for each phase φ^β deduce by the extension of the axiom of mass conservation by production terms. In general, the mass of one constituent stays not constant, it changes by mass exchange between the various constituents of the mixture, e. g. by melting of ice or chemical reactions.

$$\left(\rho^\beta\right)'_\beta + \rho^\beta \operatorname{div} \mathbf{v}_\beta = \hat{\rho}^\beta \qquad \text{with} \qquad \mathbf{v}_\beta = \mathbf{x}'_\beta \tag{15}$$

With assumption [A 4] of equal motion functions and [A 7], no mass exchange between the various constituents, the mass balance for the various constituents φ^β becomes to

$$\left(\rho^\beta\right)'_\beta + \rho^\beta \operatorname{div} \mathbf{v}_\beta = 0 \qquad \text{with} \qquad \rho^\beta = n^\beta \rho^{\beta R} \tag{16}$$

$$\Rightarrow \quad n^\beta \left(\rho^{\beta R}\right)'_\beta + \rho^{\beta R} \left(n^\beta\right)'_\beta + n^\beta \rho^{\beta R} \operatorname{div} \mathbf{v}_\beta = 0 \tag{17}$$

For the three phases, different assumptions concerning the compressibility were taken, leading to different mass balances. Phase (1), C and (2) MgO are material incompressible, hence, the mass balance reduces to a volume balance

$$\left(\rho^{\beta R}\right)'_\beta = 0 \quad \Rightarrow \quad \left(n^\beta\right)'_\beta + n^\beta \operatorname{div} \mathbf{v}_\beta = 0 \qquad \text{for} \quad \beta = 1, 2 \tag{18}$$

Since all velocities of the phases are equal, the material time derivative is equal to the total time derivative $\left(n^\beta\right)'_\beta = \dot{n}^\beta$ and it follows directly by integration of Eq. 18

$$n^\beta = n_0^\beta \det \mathbf{F}_\beta^{-1} = n_0^\beta \det \mathbf{F}^{-1} \tag{19}$$

In contrast to the carbon and magnesia phase, phase (3), the pore gas is materially compressible. In the mass balance of Eq. (17), n^β will be replaced by $n^3 = 1 - n^1 - n^2$ according to the saturation condition (Eq. (3)). The material time derivative $(n^3)'_3$ can be replaced using the definition of the material time derivative and the volume balance for phase (1) and (2) (Eq. (18)) by

$$\left(n^3\right)'_3 = n^1 \operatorname{div} \mathbf{v_1} - \operatorname{grad}\left(n^1\right) \left(\mathbf{v_3} - \mathbf{v_1}\right) + n^2 \operatorname{div} \mathbf{v_2} - \operatorname{grad}\left(n^2\right) \left(\mathbf{v_3} - \mathbf{v_2}\right) \tag{20}$$

Based on assumption [A 4] with equal motion functions, the velocities are equal, as well $\mathbf{v} = \mathbf{v_1} = \mathbf{v_2} = \mathbf{v_3}$. Hence, there is no relative velocity between the various phases and the seepage velocity between the phases became zero, whereas all diffusion terms and gradients vanish. The material time derivative reduces to

$$\left(n^3\right)'_3 = \left(n^1 + n^2\right) \operatorname{div} \mathbf{v}. \tag{21}$$

Hence, the mass balance of the pore gas reduces to

$$\left(1 - n^1 - n^2\right) \left(\rho^{3R}\right)'_3 + \rho^{3R} \operatorname{div} \mathbf{v} = 0. \tag{22}$$

With $\operatorname{div} \mathbf{v} = (\det \mathbf{F})^{\bullet} / \det \mathbf{F}$ and $n^\beta = n_0^\beta \det \mathbf{F}^{-1}$ the mass balance for the compressible pore gas follows by direct integration

$$\rho^{3R} = \rho_0^{3R} \frac{1 - n_0^1 - n_0^2}{\det \mathbf{F} - n_0^1 - n_0^2} \tag{23}$$

3.2 Momentum balance

The change of the momentum of a body is caused by any force acting on the body. Disregarding external forces, for the momentum balance of the mixture, one has

$$\rho \ddot{\mathbf{x}} = \operatorname{div} \mathbf{T} + \mathbf{b} \tag{24}$$

The partial momentum balance of the constituent φ^β with the partial stress $\mathbf{T}^\beta$ is

$$\rho^\beta \mathbf{x}''_\beta = \operatorname{div} \mathbf{T}^\beta + \mathbf{b}^\beta + \hat{\mathbf{p}}^\beta \tag{25}$$

The momentum production term $\hat{\mathbf{p}}^\beta$ describes the interaction forces between the various constituents φ^β. The stress tensor of the mixture consists of the sum of the partial stresses and the diffusion terms, where $\mathbf{x}'_\beta$ is the velocity of the constituent φ_β and d_β is the relative motion by the diffusion velocity

$$\mathbf{T} = \sum_\beta \left(\mathbf{T}^\beta - \rho^\beta \mathbf{x}'_\beta \otimes \mathrm{d}_\beta\right) \quad \text{with [A 4]} \quad \Rightarrow \quad \mathbf{T} = \sum_\beta \mathbf{T}^\beta \tag{26}$$

With assumption [A 4], there is no relative motion between the various phases, whereas the diffusion part $\rho^\beta \mathbf{x}'_\beta \otimes \mathrm{d}_\beta$ is zero. The stress tensor of the mixture is the sum of the stresses of its constituents. The constituents are allowed to exchange momentum, but a production of momentum for the whole mixture is not acceptable, hence $\sum_\beta \hat{\mathbf{p}}^\beta = \mathbf{0}$.

4 LINEAR THERMOELASTICITY FOR MgO-C FOAMS

4.1 Strain measures by multiplicative decomposition of F

For the evaluation of strain measures of the phases (1), C and (2), MgO, the deformation gradient $\mathbf{F}$ will be multiplicatively decomposed, according to Figure 3. The intermediate configuration $\mathcal{B}_\theta$ at a nonuniform temperature θ is obtained from the deformed

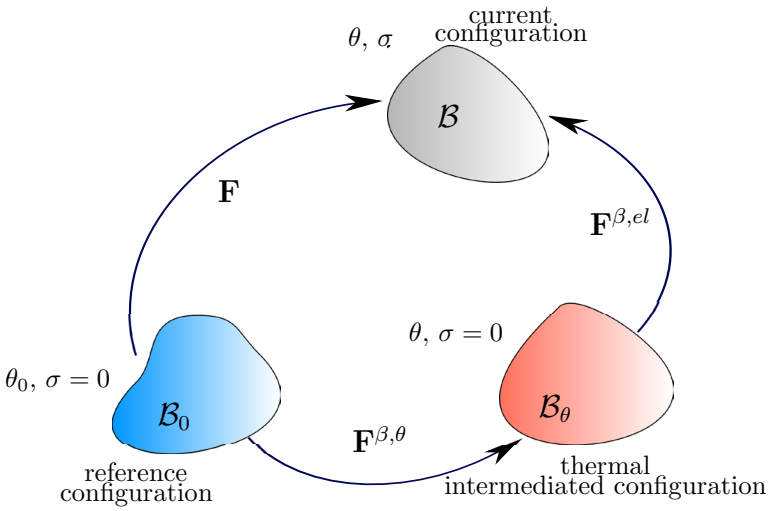


Figure 3: Multiplicative decomposition of the deformation gradient $\mathbf{F}$ in a thermal $\mathbf{F}^{\beta,\theta}$ and an elastic part $\mathbf{F}^{\beta,el}$ for the constituent φ^{β}; $\mathbf{F} = \mathbf{F}^{\beta} = \mathbf{F}^{\beta,el} \cdot \mathbf{F}^{\beta,\theta}$.

configuration $\mathcal{B}$ by isothermal destressing to zero stress. The deformation gradient from initial to deformed configuration $\mathbf{F}$ is decomposed into an elastic part $\mathbf{F}^{el}$ and a thermal part $\mathbf{F}^{\theta}$, such that $\mathbf{F} = \mathbf{F}^{\theta} \cdot \mathbf{F}^{el}$ [6, 7]. Under the assumtion of equal motion functions [A 4] it follows that

$$\mathbf{F} = \mathbf{F}^{\beta} = \mathbf{F}^{\beta,\theta} \cdot \mathbf{F}^{\beta,el} \tag{27}$$

Where $\mathbf{F}^{\beta,\theta}$ describes the thermal deformation and $\mathbf{F}^{\beta,el}$ the elastic deformation of the constituent φ^{β}. Under the assumption of isotropy and pure volumetric temperature evolution, $\mathbf{F}^{\beta,\theta} \cdot \mathbf{F}^{\beta,el}$ leads to the same result as $\mathbf{F}^{\beta,el} \cdot \mathbf{F}^{\beta,\theta}$ [6]. For the thermal part of the deformation gradient under the above-mentioned assumptions, one has

$$\mathbf{F}^{\beta,\theta} = \vartheta_{\beta}(\theta)\, \mathbf{I} \tag{28}$$

where $\vartheta_{\beta}(\theta)$ is the thermal stretch ratio. With the linear coefficient of thermal expansion (CTE) α, the thermal stretch ratio can be written as

$$\vartheta_{\beta}(\theta) = \frac{l}{l_0} = 1 + \alpha(\theta)\, \Delta T \approx 1 + \alpha_0\, (\theta - \theta_0) \tag{29}$$

ΔT describes the temperatue difference between the temperature θ_0 in the reference configuration and the temperature θ of the thermal intermediate configuration. Based on the different CTEs of carbon and magnesia ($\alpha_C \neq \alpha_{MgO}$), the thermal part of the deformation gradient differs for each phase, whereas the total deformation gradient is equal.
For the Green-Lagrangian strain tensor, reads as follows

$$\mathbf{E} = \frac{1}{2} \left(\mathbf{F}^{\mathbf{T}} \cdot \mathbf{F} - \mathbf{I} \right) = \mathbf{E}^{\beta} = \frac{1}{2} \left((\mathbf{F}^{\beta})^{\mathbf{T}} \cdot \mathbf{F}^{\beta} - \mathbf{I} \right) \tag{30}$$

The thermal strain of the mixture on the reference configuration is defined by

$$\mathbf{E}^{\theta} = \frac{1}{2}\left((\mathbf{F}^{\theta})^{T} \cdot \mathbf{F}^{\theta} - \mathbf{I}\right) \tag{31}$$

The thermal strains of the various constituents are defined by

$$\mathbf{E}^{\beta,\theta} = \frac{1}{2}\left((\mathbf{F}^{\beta,\theta})^{T} \cdot \mathbf{F}^{\beta,\theta} - \mathbf{I}\right) \tag{32}$$

The Green-Lagrangian strain can be additively splitted into

$$\mathbf{E} = \mathbf{E}^{\theta} + \mathbf{E}^{el} \tag{33}$$

The elastic Lagrangian strain $\mathbf{E}^{el} = \mathbf{E} - \mathbf{E}^{\theta}$ defined relative to $\mathcal{B}$ is a mixed term of $\mathbf{E}^{el}$ and $\mathbf{E}^{\theta}$. With the push-forward operation, the Lagrangian strain $\hat{\mathbf{\Gamma}}^{el}$ relative to the intermediate configuration $\mathcal{B}_{\theta}$ at non-uniform temperature θ relates to

$$\hat{\mathbf{\Gamma}}^{el} = (\mathbf{F}^{\theta})^{-T}\left(\mathbf{E} - \mathbf{E}^{\theta}\right)(\mathbf{F}^{\theta})^{-1} = \frac{1}{\vartheta^{2}}\left(\mathbf{E} - \mathbf{E}^{\theta}\right) \tag{34}$$

Using Eq. 28 and the symmetry of the thermal deformation gradient, the thermal strain in Eq. 31 becomes to

$$\mathbf{E}^{\theta} = \frac{1}{2}\left(\vartheta^{2} - 1\right)\mathbf{I} \tag{35}$$

By substituting $\mathbf{E}^{\theta}$ in equation Eq. 34 in relation to $\mathcal{B}_{\theta}$, the elastic strain on the intermediate configuration can be expressed by the total Lagrangian strain $\mathbf{E}$ of the reference configuration and the thermal stretch ratio ϑ as

$$\hat{\mathbf{\Gamma}}^{el} = \frac{1}{2}\left(\frac{1}{\vartheta^{2}} - 1\right)\mathbf{I} + \frac{1}{\vartheta^{2}}\mathbf{E}. \tag{36}$$

Since the overall strains $\mathbf{E}$ for all phases are equal (see Eq. 30), the elastic strain for each phase is distinguished by the thermal stretch ratio ϑ_{β} and hence, by the linear coefficient of thermal expansion α_{β}. The elastic strain for the constituent φ^{β} becomes to

$$\hat{\mathbf{\Gamma}}^{el,\beta} = \frac{1}{2}\left(\frac{1}{\vartheta_{\beta}^{2}} - 1\right)\mathbf{I} + \frac{1}{\vartheta_{\beta}^{2}}\mathbf{E}. \tag{37}$$

4.2 Stresses based on linear thermoelasticity

Using the concept of multiplicative decomposition of the deformation by the introduction of a thermal intermediate configuration, the Helmholtz free energy for the various constituents φ^{β} can be additively split into an isotropic function $\psi^{el,\beta}$ of the elastic strain

$\hat{\mathbf{\Gamma}}^{el,\beta}$ and the temperature θ and a part $\psi^{\theta,\beta}$ only dependent on the temperature, according to [7]

$$\psi^{\beta} \;=\; \psi^{el,\beta}\left(\hat{\mathbf{\Gamma}}^{el,\beta},\theta\right) \;+\; \psi^{\theta,\beta} \tag{38}$$

The partial stress $\mathbf{T}^{\beta}$ relies only on the elastic strains and the elastic part of the free energy, but it is weighted by the thermal stretch ratio [7],

$$\mathbf{T}^{\beta} \;=\; \frac{\rho_{\beta,0}}{\vartheta_{\beta}^{2}}\,\frac{\partial \psi^{el,\beta}}{\partial \hat{\mathbf{\Gamma}}^{el,\beta}} \tag{39}$$

With the relationship $\rho_{\beta,0} \;=\; \vartheta_{\beta}^{3}\,\rho_{\beta,\theta}$, between the densities $\rho_{\beta,0}$ of the reference configuration and $\rho_{\beta,\theta}$ in the intermediate configuration, the stress response can also be written as

$$\mathbf{T}^{\beta} \;=\; \vartheta_{\beta}\,\mathbf{T}^{el,\beta}, \qquad \mathbf{T}^{el,\beta} \;=\; \rho_{\beta,0}\,\frac{\partial \psi^{el,\beta}}{\partial \hat{\mathbf{\Gamma}}^{el,\beta}} \tag{40}$$

For the suggestion, that the Helmholtz free energy $\psi^{el,\beta}$ is a quadratic function of the elastic strain components, such that

$$\rho_{\beta,\theta}\,\psi^{el,\beta} \;=\; \frac{1}{2}\,\lambda_{\beta}(\theta)\left(\mathrm{tr}\left(\hat{\mathbf{\Gamma}}^{el,\beta}\right)\right)^{2} \;+\; \mu_{\beta}(\theta)\,\hat{\mathbf{\Gamma}}^{el,\beta} : \hat{\mathbf{\Gamma}}^{el,\beta} \tag{41}$$

whereas $\lambda_{\beta}(\theta)$ and $\mu_{\beta}(\theta)$ are the temperature-dependent Lamé constants of the various solid phases φ^{β}. It follows, that

$$\mathbf{T}^{el,\beta} \;=\; \mathbb{C}^{el,\beta}(\theta) : \hat{\mathbf{\Gamma}}^{el,\beta}, \qquad \text{with} \quad \mathbb{C}^{el,\beta}(\theta) \;=\; \lambda_{\beta}(\theta)\,\mathbf{I}\otimes\mathbf{I} \;+\; 2\,\mu_{\beta}(\theta)\,\mathbb{I} \tag{42}$$

Where $\mathbb{I}$ is the fourth-oder unit tensor

$$\mathbb{I}_{ijkl} \;=\; \frac{1}{2}\,\left(\delta_{ik}\,\delta_{jl} \;+\; \delta_{il}\,\delta_{jk}\right) \tag{43}$$

According to the assumptions for the motion function [A 4] and the deformation gradient of the various constituents, the overall strain of the mixture is equal to the partial overall strains, $\mathbf{E} \;=\; \mathbf{E}^{\beta} \;=\; \mathbf{E}^{1} \;=\; \mathbf{E}^{2}$ since $\mathbf{F} \;=\; \mathbf{F}^{\beta} \;=\; \mathbf{F}^{1} \;=\; \mathbf{F}^{2}$. By substituting Eq. 42 and Eq. 36 into Eq. 40, based on the isotropic linear thermoelasticity, the partial stress $\mathbf{T}^{\beta}$ for the various constituents φ^{β} is

$$\mathbf{T}^{\beta} \;=\; \frac{1}{\vartheta_{\beta}}\,\left[\lambda_{\beta}(\theta)\,\mathrm{tr}(\mathbf{E})\,\mathbf{I} \;+\; 2\,\mu_{\beta}(\theta)\,\mathbf{E}\right] \;+\; \frac{1}{2}\,\mathbf{I}\,\left(\frac{1}{\vartheta_{\beta}} \;-\; \vartheta_{\beta}\right)\,\left[3\,\lambda_{\beta}(\theta) \;+\; 2\,\mu_{\beta}(\theta)\right] \tag{44}$$

Notice, the thermal stretch ratio $\vartheta_{\beta}(\theta)$ depends also on the temperature. Replacing the Lamé constant λ_{β} by the compression modulus κ_{β} with

$$\kappa_{\beta}(\theta) \;=\; \lambda_{\beta} \;+\; \frac{2\,\mu_{\beta}}{3} \quad \Leftrightarrow \quad \lambda_{\beta} \;=\; \kappa_{\beta} \;-\; \frac{2\,\mu_{\beta}}{3} \tag{45}$$

the partial stress can be written as

$$\mathbf{T}^{\beta} = \frac{1}{\vartheta_{\beta}(\theta)} \left[\lambda_{\beta}(\theta)\,\mathrm{tr}(\mathbf{E})\,\mathbf{I} + 2\,\mu_{\beta}(\theta)\,\mathbf{E}\right] + \frac{3}{2}\left(\frac{1}{\vartheta_{\beta}(\theta)} - \vartheta_{\beta}(\theta)\right)\kappa_{\beta}(\theta)\,\mathbf{I} \tag{46}$$

If the Lamé constants are taken to be temperature-independent, and if for the thermal stretch ratio, the approximation $\vartheta_{\beta}(\theta) \approx 1 + \alpha_{\beta,0}\,(\theta - \theta_0)$ is used, the partial Cauchy stress $\mathbf{T}^{\beta}$ can be expressed by the infinitesimal strain $\mathbf{E}$ in form of the *Duhamel-Neumann* expression of isotropic linear thermoelasticity [8, 9].

$$\mathbf{T}^{\beta} = \lambda_{\beta,0}\,\mathrm{tr}(\mathbf{E})\,\mathbf{I} + 2\,\mu_{\beta,0}\,\mathbf{E} - 3\,\alpha_{\beta,0}\,(\theta - \theta_{\beta,0})\,\kappa_{\beta,0}\,\mathbf{I} \tag{47}$$

Hence, the partial stresses of the phases (1), C and (2), MgO depend on the same strain $\mathbf{E}$ but differ according to the different Lamé constants and CTEs.

5 CONSTITUTIVE EQUATIONS FOR MgO-C FOAMS

The constitutive equation for the MgO-C foams is evaluated using the rule of mixture for the partial stresses $\mathbf{T}^{\beta}$, whereas the partial stresses are weighted by their volume fraction n^{β}. The pore pressure p^3 is caused by the compression of the mixture, leading to the compression of the pore gas (phase (3), air). According to Newton's law 'actio = reactio', the pore pressure p^3 acts on the both other phases and hence contributes to 100% to the stress of the mixture. For the overall stress of the MgO-C foams, made of the three phases, one has

$$\mathbf{T} = -\,p^3\,\mathbf{I} + n^1\,\mathbf{T}^1 + n^2\,\mathbf{T}^2 \tag{48}$$

To describe the evaluation of the pore pressure p under deformation of the mixture, the ideal gas law was used as equation of state for the pore gas.

$$p\,V = n\,R\,\theta \qquad \text{with} \quad \rho = \frac{m}{V} \quad \text{and} \quad M = \frac{m}{n} \tag{49}$$

V denotes to the gas volume, n is the amount of substance, R the ideal gas constant, m the mass and M the molar mass of the gas. Finally, using the mass balance for the pore gas (Eq. 23), the pore pressure p^3 can be determined by

$$p^3 = \rho_0^{3R}\,\frac{R\,\theta}{M^{air}}\,\frac{1 - n_0^1 - n_0^2}{\det\mathbf{F} - n_0^1 - n_0^2} \tag{50}$$

6 CONCLUSIONS

In this contribution, we presented a new multiphysics approach for porous media, especially cellular MgO-C refractories. The key assumption is the use of the theory of porous media (TPM) with a kinematic coupling of the displacement and temperature fields of all constituents. It was assumed, that there is no interdiffusion of the three constituents

(MgO, C and pore gas) with the result that the motion functions and hence the deformation gradients of all phases are equal. Linear thermoelasticity with a multiplicative decomposition of the deformation gradient into an elastic and a thermal part for isotropic materials according to the Duhamel-Neumann relation was extended to the mixture of MgO and C phase. The difference in the coefficient of thermal expansion (CTE) leads to different thermal induced stresses in both phases. For the pore gas, the ideal gas law was used as equation of state for the evolution of pressure as function of the density and the temperature. The total macroscopic stress was calculated using the theory of mixture, inclusing the contributions from the pore pressure and the two solid phases.

REFERENCES

[1] Falk G, Jung A, Da Silveira W, Diebels S. Experimental characterization and thermo-mechanical modelling of micro structure interactions in cellular carbon magnesia refractories. *J. Ceram. Sci. Tech.* (2014) **5(2)**:101–114.

[2] Jung A, Falk G, Petri D, Diebels S. Thermal shock resistivity of hybrid carbon foam materials: Experiments and model predictions. *Mech. Mater.* DOI: 10.1016/j.mechmat.2014.12.007

[3] Truesdell, C. and Toupin, A. *The classical field theories. Handbuch der Physik, Band III/1* , Springer-Verlag, (1960).

[4] Bowen, R. M. Incompressible porous media models by use of the theory of mixtures. *Int. J. Engng. Sci.* (1980) **18**:1129–1148.

[5] Ehlers, W. Grundlegende Konzepte in der Theorie Poröser Medien. *Tech. Mech.* (1996) **16**:63–76.

[6] Hartmann, S. Comparison of the multiplicative decomposition $F=F_\theta F_M$ and $F=F_M F_\theta$ in finite strain thermoelasticity. *Technical Report Series TU Clausthal, Clausthal University of Technology* (2001):Fac3-12-01.

[7] Lubarda, V.A. Constitutive theories based on the multiplicative decomposition of deformation gradient: Thermoelasticity, elastoplasticity, and bimechanics. *Appl. Mech. Rev.* (2004) **57(2)**:95–108.

[8] Carlson, D.E. *Linear thermoelasticity, Handbuch der Physik, Band VIa/2.* Springer-Verlag, (1972).

[9] Nowacki, W. *Thermoelasticity.* Pergamon Press, (1986).

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis (Eds)

TWO SIDES OF A COIN:

A CRITICAL REVIEW, AND MATHEMATICAL AND PHENOMENOLOGICAL STUDY OF WHAT WE CALL HYDRO-MECHANICAL COUPLING

J.M.HUYGHE*, E. NIKOOEE†, T. SWEIJEN† AND S. MAJID HASSANIZADEH†

*Department of Mechanical Engineering, TU Eindhoven, The Netherlands

e-mail: j.m.r.huyghe@tue.nl

† Earth Sciences Department, Utrecht University, The Netherlands

E-mail: e.nikooee@uu.nl, T.Sweijen@uu.nl, s.m.hassanizadeh@uu.nl

Key words: Unsaturated soils, Coupled phenomena, Capillary-pressure saturation curve, Discrete Element Modeling (DEM), Two phase flow

Abstract. In this paper a brief and critical review of the current literature on hydro-mechanical coupling is presented. Furthermore, an enhanced discrete element model is used to investigate the mutual relationship of soil water retention curve and suction stress curves and how the two are affected as a result of change in the initial porosity of the soil sample. The model revealed the suction stress values in wetting were less affected as in drying branch as a result of the change in the initial porosity of the soil sample.

1 INTRODUCTION

Unsaturated soil is a complex medium composed of different phases and interfaces. The volume fraction of different phases influences the overall macroscopic behavior of unsaturated soil. On the other hand the porous network of unsaturated soil changes under external loading and is stress level dependent. Therefore, the volume fraction of different phases changes with the stress level and accordingly the topology, connectivity of the porous network, and fluid paths change. This means the mechanical behavior of system is saturation dependent and the hydraulic behaviour of the system is stress level dependent. This has been referred to as hydro-mechanical coupling in the literature of unsaturated soil mechanics. There are various studies in the literature which investigate each of these phenomena alone and independently and the connection of the two is not comprehensively looked into [1]. The soil water retention curves for deformable porous media have been explored by various researchers [e.g., 2-4] and the dependency of stress measures on the saturation level has been investigated in some others [5]. However, these two phenomena are mutually inter-related and it is also of practical importance to look into saturation dependent stress measures and stress dependent retention properties in a single framework, and not only that, different micro scale and macro scale phenomena are also

worth being investigated from this perspective. In this study, a thorough and critical review of current literature on the hydro-mechanical coupling is presented and different phenomenological and mathematical aspects of hydro-mechanical coupling are discussed.

2 A BRIEF AND CRITICAL DISCUSSION OF THE LITERATURE

Noting the points raised in the introduction section, one will find the following two issues have faced less attention in the current literature:

1. Is there any mathematical, and or physical framework to demonstrate the mutual relationship discussed above? In other words, such coupled behavior which can be empirically and intuitively understood and is widely acknowledged among researchers, should also be established and demonstrated via a mathematical and/or physical approach. The benefit of any further elaboration on this is better understanding of such coupling and arriving at the better options to incorporate it in the current constitutive models and computational tools instead of the ad hoc, heuristic and empirical ways of proposing equations for it.

2. One of the important factors influencing the hydro-mechanical coupling is the internal structure of the porous medium (i.e., the porous structure of the medium). In other words, the effect of stress level on the hydraulic behaviour of the porous media can be different for different porous network structures. It can vary in different porous media with various porous networks (in networks with mono-modal or bimodal pore size distribution, dissimilar pore shapes, different distribution of pores and throats, and coordination number distributions (or porous network connectivity). This means a pore and/or grain scale study of the coupled phenomena is important to shed light into such coupling. We hereafter refer to this part as phenomenological study as with such a tool we can better study different processes and physical effects associated with the coupled phenomena.

This paper briefly reports some of the recent works of the authors on these two important aspects of coupled phenomena in unsaturated poromechanics. In the following section, we first quickly illustrate how such mutual relationship is expected to exist due to the inherent mathematical symmetry in the unsaturated system. Afterwards, the DEM approach which is employed by the authors to look at hydro-mechanical behaviour of unsaturated soils is introduced. Finally concluding remarks are given and further research directions are discussed.

3 A MATHEMATICAL VIEW POINT: INHERENT SYMMETRY

Let W be a strain energy function which represents the free energy of the deformable unsaturated medium per a unit initial volume of it. Huyghe et al. (2015) have revealed that the following results can be obtained through a systematic formulation of the thermodynamics of such media [6]:

$$\boldsymbol{\sigma}_{eff} = \frac{1}{J}\boldsymbol{F}.\frac{\partial W}{\partial \boldsymbol{E}}.\boldsymbol{F}^{c} \tag{1}$$

$$P_C = -\rho_i^w \frac{\partial W}{\partial R^w} \tag{2}$$

In Equation 1, $\boldsymbol{\sigma_{eff}}$is the so-called effective stress tensor. **E** specifies the Green strain, and J denotes the Jacobian of the deformation gradient tensor (**F**).

In Equation 2, P_c is capillary pressure which is the pressure difference between air and water. ρ_i^w stands for the intrinsic density of the wetting fluid and R^w is Lagrangian apparent density of wetting fluid. These two equations are two of the most important equations which can define mechanics and hydraulics of the unsaturated mixture. Having a look at the symmetry involved in the following equation:

$$\frac{\partial W}{\partial R^w \partial \boldsymbol{E}} = \frac{\partial W}{\partial \boldsymbol{E} \partial R^w} \tag{3}$$

If we note that the left hand side is connected to the deformation effect on the values of capillary pressure and the right hand side is the effect of change of saturation on effective stress relationship, then this equation and the symmetry involved there basically states from an energetic view point these two phenomena are mutually interrelated. In fact such relationship can be used to establish a formulation for the effective stress in unsaturated soil which is fully discussed in Huyghe et al. (2015) and interested reader can refer to that [6].

We now add two other important equations to the set of equations presented above. The soil water retention curve (SWRC), defined as follows:

$$P_C = F(S^W) \tag{4}$$

It is worth mentioning that the F function (appearing in Equation 4) could be directly thermodynamically/ mathematically obtained, if the exact form of dependencies of W (mixture strain energy) was known. In the absence of such knowledge, the numerical and experimental approaches are utilized to get insights on the form of F function. Therefore the F function is usually determined empirically, most well-known forms of it are Brooks-Corey (BC), Van Genuchten (VG), Fredlund and Xing (FX) [7-9] but of course it can also be obtained from modeling tools such as discrete element model provided that the model can hydraulically represent the unsaturated soil well enough to give results which are within a reasonable/acceptable range of accuracy. For this purpose, calibration of the model with experimental data at the first stage is pursuit. In general, the pore and grain characteristics of the model and the medium (and their degree of similarity) can be considered as two most influential factors in arriving at a proper estimation of SWRC. Geometry of the grain and the intra and inter-aggregate porosities are also of importance.

Not surprisingly there is the same problem with the effective stress defined in Equation (1) based on the strain energy function of the mixture which leaves us with the option of empirical equation for it (of course the effective stress formulation can still be obtained through a thermodynamic framework but we need to either assume some dependences for strain energy function or for the soil water retention relationship and plug it into the Equation (1) or (2), respectively). This is of course subject of a separate study (see for instance Huyghe et al. 2015). For the current study, let us consider one of the most famous empirical equations for the effective stress in unsaturated soils, known as Bishop Equation [10]:

$$\boldsymbol{\sigma_{eff}} = \boldsymbol{\sigma} - p^a \boldsymbol{I} + \chi(p^a - p^w)\boldsymbol{I} \tag{5}$$

where $\boldsymbol{\sigma}$ is total stress. Hereafter, the first term on the right hand side (the difference between total stress and pore air pressure) is called net stress. p^w and p^a are wetting phase and nonwetting phase pressures, respectively (in our case, water and air). χ is the so-called effective stress parameter.

The difference between the effective stress and net stress has been coined by Lu and Likos as suction stress [11].

Lu et al. (2010) proposed the following equation for the suction stress [12]:

$$\boldsymbol{SS} = \boldsymbol{\sigma_{eff}} - (\boldsymbol{\sigma} - p^a \boldsymbol{I}) = S^e(p^a - p^w)\boldsymbol{I} \tag{6}$$

where $\boldsymbol{S^e}$ is the effective degree of saturation. Here, we have changed their formulation regarding the sign convention and we have also rewritten it in a tensorial format.

Therefore, one straight forward way to look into the coupled phenomena and mutual relationship stated by Equation 3 is first investigating the effect of stress level on the soil water retention curve. The values of effective saturation at each capillary pressure can then be employed to look into the effect of stress level or deformation on the stress measures of the medium (effective stress or suction stress). This can be done by investigating the suction stress at different stress levels or porosity levels, or investigating it, at least (in a most simple way), at different initial porosities. For this purpose experimental data or numerical models can be employed. Numerical models have this merit that they are not as time consuming as the experimental tests and a better control on the physical factors can be achieved but on the other hand, at the present, there are lots of simplifying assumptions involved in the numerical modeling. That said, due to considerable merits of the numerical model, we have focused on a numerical model (discrete element model) here.

One can notice the mathematical analysis presented above did not elaborate on the hysteresis phenomenon (i.e., the difference in the soil water retention curves in drying and wetting as well as the other hydraulic paths). However, a possible approach to account for the hysteresis phenomenon is considering different parameters for Equation (4) for different hydraulic paths. Nonetheless, in order to consider the effect of deformation, similar hydraulic paths would need to be compared at different states of stress (or measures of deformation). The numerical simulation presented in this study, thus, includes the variation of suction stress for both drying and wetting at different porosity values.

In what follows a modeling tool (discrete element technique) which is capable of capturing the main features of unsaturated porous media is employed to model the two sides of hydro-mechanical coupling. As we are in the preliminary stages of this research, the study presented here is limited to first model the effect of different initial porosities on the soil water retention curve and then to investigate the suction stress values at different levels of capillary pressure for different values of initial porosity (for drying and wetting branches). Therefore, in the current study basically the change of initial porosity has been considered as a measure of stress level and then drying and wetting paths have been modeled. This means highly collapsible soils and/or highly swelling soils are not subject of this study. In order to have a precise model to study them, the instantaneous change of porosity as a result of suction change needs to be accounted for.

4 THE PHENOMENOLOGICAL STUDY: DISCRETE ELEMENT MODELING OF THE COUPLED PHENOMENA

The phenomenological study contains two parts. First simulating the soil water retention curves using a Discrete Element Method (DEM). The discrete elements model is calibrated to be utilized as a virtual laboratory. It is then employed to obtain the soil water retention curve for different initial porosities of the same soil type to account for the effect of stress level. To study the other direction of this mutual relationship (the effect of change of porosity on soil water retention curve and then consequently on the effective stress state of the medium), the suction stress curves under different porosities are obtained using the results of DEM.

The simulation of SWRCs are done in two subsequent stages. First a packing was generated in particle model Yade-DEM, and second the SWRC was determined by an implemented quasi-static two-phase flow module.

Particle model Yade-DEM is based on the Discrete Element Method (DEM), where particles are considered as discrete elements. Each particle is simulated as an elastic sphere. At a contact between two particles a force is calculated which is a consequence of elasticity, this is calculated using the physical based Hertz-Mindlin mechanics. The elastic force at a contact between two particles is given by [13]:

$$f_n = -k_n \delta_n^{\,3/2} \tag{7}$$

f_n is the inter-particle force at a contact, δ_n is the normal displacement. k_n is the contact stiffness in the normal direction and is given by:

$$k_n = \frac{4}{3} E^* \sqrt{R^*} \tag{8}$$

Where $E^* = \left(\frac{1-\nu_1{}^2}{E_1} + \frac{1-\nu_2{}^2}{E_2}\right)^{-1}$, representing the average of Young's moduli of the two particles, R^* is the reduced radius and given by $R^* = \frac{R_1 R_2}{R_1 + R_2}$

For each particle all forces arising from contacts, gravity, and boundary conditions are summed and transferred to particle motions (using Newton's second law). Using Yade-DEM, a packing of 4000 spheres was randomly generated with the particle size distribution for Hostun sand [14]. The particles were packed by an all-round stress, before a vertical confinement was applied. The vertical net-stress was increased until a target porosity was achieved, this was 0.39 [14]. For the inter-particle friction we used that of Labenne sand (17°) and a Poisson ratio of 0.5 [15]. The particle stiffness was set at 1MPa. However as we used porosity as a measure for state-of-stress, the exact particle stiffness is not required to arrive at a certain initial porosity, as the confining stress would capture the exact state-of-stress.

After the target porosity was achieved, an implemented quasi-static two-phase flow algorithm was used to obtain points on the SWRC for both drainage and imbibition. First the pore geometry was meshed into tetrahedra using Delauney triangulation [16]. Each tetrahedron represents one pore-unit and has at each corner one particle. The resulting network of pore-units was used to simulate the SWRC experiments. The top of the box generated in Yade-DEM represented an air reservoir and had an air-pressure (P^a) and at the bottom a water reservoir with a water pressure (P^w). As mentioned before, the capillary pressure is defined as $P_C = P^a - P^w$. The water saturation in each pore-unit was either one or zero (i.e., small partial water volumes in pore units were considered negligible). Initially the packing was saturated with water. At the pore-scale the capillary pressure was used to establish whether air will invade during drainage or water during imbibition, using the following criteria. Drainage: the invasion criterion was based on the entry pressure (P_e) associated with the smallest transect in between two pore units. Air invaded if $P_c > P_e$. Imbibition: the invasion criterion was based on the critical pressure (P) corresponding to the largest curvature found inside a pore-unit, which is that of an inscribed sphere and was calculated following [17]. Water invaded if $P_c < P_e$. For each invasion of a pore-unit the connectivity of both water and air phase with their reservoirs was checked, if a residual phase was present invasion of that pore-unit did not occur.

5 RESULTS AND DISCUSSION: SIMULATED SWRCS AND SSCCS FOR HOSTUN SANDS

Most granular materials do not contain perfect spherical particles. Therefore, a slightly more angular material, sand, is tested. Lins and Schanz (2005) have studied the soil-water retention curve in Hostun dune sand, at a porosity of 0.39 [14] The results (Fig. 1) indicate that the capillary pressure for both drainage and imbibition is in good agreement with the experimental data. We simulated spherical particles in contrast to sand, which typically is more angular, although the effect of particle angularity on SWRC is limited [18]. During drainage, the irreducible water saturation in angular sands can be present in pendular rings, corners, and as disconnected blobs, from which the later one is only accounted for in these simulations. During imbibition corner flow can strongly influence snap-off, thus affecting the amount of disconnected air blobs. To improve the magnitude of residual air saturation and the irreducible water saturation, effects of angularity such as corner flow, should be accounted for. Our above-mentioned relatively basic simulation of imbibition does, however, still capture the main effects.

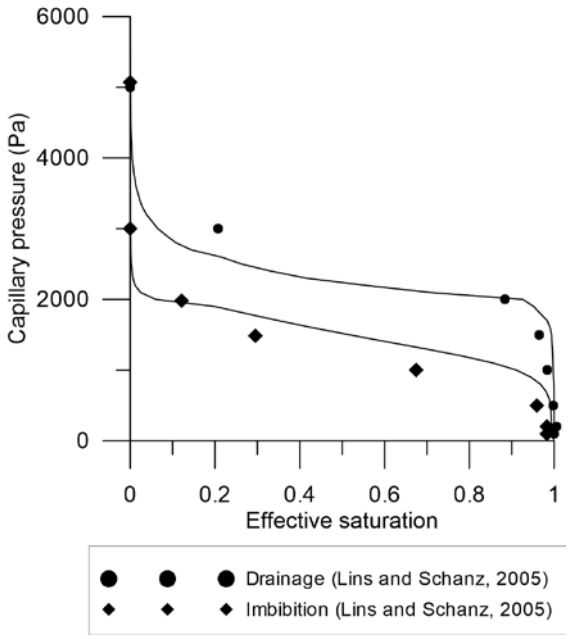


Fig. 1: Capillary-pressure saturation behavior for Hostun sand, with a porosity of 0.39. Experimental data by Lins and Schanz (2005), and model results by quasi-static two-phase flow within Yade-DEM.

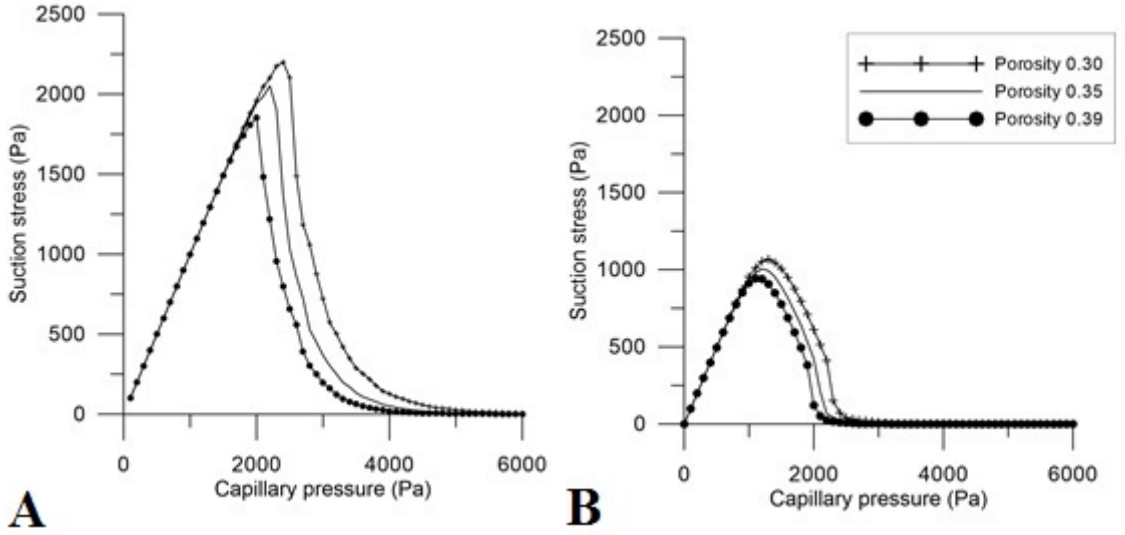


Fig 2: Suction stress as function of capillary pressure for Hostun sand at different porosities A) drainage and B) imbibition.

To investigate the hysteresis and stress level effects on suction stress, the capillary pressure - saturation curve of Huston sand at a porosity of 0.39 (Fig. 1) is transferred to suction stress – capillary pressure curve for drainage (Fig. 2A) and imbibition (Fig. 2B).

It is apparent that larger states of suction stresses are found during drainage than imbibition. The effect of porosity on suction stress is investigated by decreasing the porosity from 0.39 to 0.35 and 0.30. Simulations indicate that a decrease in porosity causes an increase in suction stress, as pore sizes become smaller. Such effects are more pronounced in the drying curves as compared to the imbibition ones. That can root in the fact that throat radii control the drying process and they are more easily affected by a change in the porosity and/or an increase in the stress level than the radii of pores. The radii of pores control the imbibition and therefore, imbibition branches of suction stress are not affected as much as drying branches as a result of porosity change.

6 CONCLUDING REMARKS

- Critical evaluation of the literature showed the mutual relationship between stress level effects on the capillary pressure and effective stress measure has not yet been well studied in the literature. To better investigate such mutual interrelationship, some insights based on mathematical symmetry were presented. This issue has extensively and comprehensively been explained in a forthcoming publication of the authors [6].
- A discrete element model was employed to elucidate how such mutual relationship holds. In particular, we investigated the effect of deformation (change in the initial porosity of the soil sample) on the suction stress and soil water retention curve in both drying and wetting paths. The utilized discrete element model was first calibrated with soil water retention data of one of the samples and was then used to predict the effect of porosity change on the suction stress and soil water retention curves.

ACKNOWLEDGEMENT

The research of Thomas Sweijen has been made possible thanks to grant 12538 of the Technology Foundation STW, the technological branch of the Netherlands Organization of Scientific Research NWO, and the Ministry of Economic Affairs. The second author would like to thank the European Research Council (ERC) and the Department of Earth Sciences of Utrecht University for hosting and financing his research as a Post-Doctoral Researcher. We sincerely acknowledge the valuable intellectual contributions of our best friend and colleague, Dr. Chareyre of Grenoble Institute of Technology who is the YADE project leader. Without his help, this study could not reach the current state.

REFERENCES

[1] Mašín, D. Predicting the dependency of a degree of saturation on void ratio and suction using effective stress principle for unsaturated soils. *International Journal for Numerical and Analytical Methods in Geomechanics* (2010) **34(**1):73-90.

[2] Tarantino, A. A water retention model for deformable soils. *Géotechnique*, (2009) **59** (9): 751 – 762.

[3] Nuth, M. and Laloui L. Advances in modelling hysteretic water retention curve in deformable soils. *Computers and Geotechnics* (2008) **35**(6): 835-844.

[4] Mbonimpa, M., Aubertin, M., Maqsoud, A., and Bussière B. Predictive Model for the Water Retention Curve of Deformable Clayey Soils. *J. Geotech. Geoenviron. Eng.* (2006)**132**(9):1121–1132.

[5] Nuth, M. and Laloui. L. Effective stress concept in unsaturated soils: clarification and validation of a unified framework, *International Journal for Numerical and Analytical Methods in Geomechanics*, (2008) **32**(7):771-801.

[6] Huyghe, J., Nikooee, E., Hassanizadeh, S. Majid. A thermodynamical analysis of deforming unsaturated media, *Trasnport in porous media,* to be submitted soon, (2015).

[7] Brooks, R. H. and Corey, A. T. Hydraulic Properties of Porous Media. Hydrology Paper No. 3. Colorado State University, Fort Collins, CO, (1964).

[8] Van Genuchten, M. T. A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. *Soil science society of America journal* (1980) **44**(5): 892-898.

[9] Fredlund, D. G., and Xing, A., and Huang, S. Predicting the permeability function for unsaturated soils using the soil-water characteristic curve. *Canadian Geotechnical Journal* (1994) **31** (4): 533-546.

[10] Bishop, A. W. The effective stress principle. *Teknisk Ukeblad* (1959) **39**: 859-863.

[11] Lu, N., and Likos, W.J. *Unsaturated soil mechanics*. J. Wiley, (2004).

[12] Lu, N., Godt, J.w., and Wu, D.T. A closed-form equation for effective stress in unsaturated soil. *Water Resources Researc*h (2010), **46**(5).

[13] Johnson KL. *Contact mechanics*. Cambridge university press; (1987).

[14] Lins Y, Schanz T. Determination of hydro-mechanical properties of sand. In: *Unsaturated soils: Experimental studie*s. Springer (2005):15-32.

[15] Belheine N, Plassiard J, Donzé F, Darve F, Seridi A. Numerical simulation of drained triaxial test using 3D discrete element modeling. *Comput Geotech*. (2009)**36**(1):320-331.

[16] Chareyre B, Cortis A, Catalano E, Barthélemy E. Pore-scale modeling of viscous flow and induced forces in dense sphere packings. *Transp Porous Media*. (2012)**94**(2):595-615.

[17] MacKay A. To find the largest sphere which can be inscribed between four others. *Acta Crystallographica Section A: Crystal Physics, Diffraction, Theoretical and General Crystallography*. (1973)**29**(3):308-309.

[18] Torskaya T, Shabro V, Torres-Verdín C, Salazar-Tio R, Revil A. Grain shape effects on permeability, formation factor, and capillary pressure from pore-scale modeling. *Transp Porous Media*. (2014) **102**(1):71-90.

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis(Eds)

A PARALLEL FULLY-COUPLED FLUID-STRUCTURE INTERACTION SIMULATION OF A CEREBRAL ANEURYSM

ALI EKEN† AND MEHMET SAHIN†

†Astronautical Engineering Department
Faculty of Aeronautics and Astronautics
Istanbul Technical University
Maslak, 34469 Istanbul, Turkey
e-mail: msahin.ae00@gtalumni.org, web page: http://web.itu.edu.tr/~msahin/

Key words: Fluid-Structure Interaction, Monolithic Methods, ALE Methods, Cerebral Aneurysm

Abstract. A parallel fully-coupled approach has been developed for the fluid-structure interaction problem in a cerebral artery with aneurysm. An Arbitrary Lagrangian-Eulerian formulation based on the side-centered unstructured finite volume method [2] is employed for the governing incompressible Navier-Stokes equations and the classical Galerkin finite element formulation is used to discretize the constitutive law for the Saint Venant-Kirchhoff material in a Lagrangian frame for the solid domain. A special attention is given to construct an algorithm with exact fluid mass/volume conservation while obeying the global discrete geometric conservation law (DGCL). The resulting large-scale algebraic linear equations are solved using a one-level restricted additive Schwarz preconditioner with a block-incomplete factorization within each partitioned sub-domains. The parallel implementation of the present fully coupled unstructured fluid-structure solver is based on the PETSc library for improving the efficiency of the parallel algorithm. The proposed numerical algorithm is applied to a complicated problem involving unsteady pulsatile blood flow in a cerebral artery with aneurysm as a realistic fluid-structure interaction problem encountered in biomechanics.

1 INTRODUCTION

The development of increasingly accurate and reliable simulation tools in hemodynamics allows to better understand development of cardiovascular diseases, design and evaluation of medical devices and prediction of surgical outcomes. It is known that hemodynamic factors like the wall shear stress (WSS) play a major role in cardiovascular diseases. In particular, areas of turbulence, flow recirculation or places where the artery

wall is subject to low or oscillating shear stress are at higher risk for plaque formation and disease.

To investigate the influence of hemodynamic factors in blood vessels, we have developed a parallel fluid-structure interaction (FSI) analysis technique [1] and applied to a complicated problem involving unsteady pulsatile blood flow in a cerebral artery with aneurysm as a realistic fluid-structure interaction problem encountered in biomechanics. The side-centered unstructured finite volume method based on an arbitrary Lagrangian-Eulerian formulation [2] is employed for the governing incompressible Navier-Stokes equations and the classical Galerkin finite element formulation is used to discretize the constitutive law for the compressible Saint Venant-Kirchhoff material in a Lagrangian frame. The present arrangement of the primitive variables leads to a stable numerical scheme and it does not require any *ad-hoc* modifications in order to avoid odd-even pressure decoupling or spurious pressure modes on unstructured meshes [3]. To the authors' best knowledge, the present arrangement of the primitive variables is not considered for the FSI problems. In the present work, a special attention will be given to satisfy the continuity equation exactly within each element and the summation of the continuity equations can be exactly reduced to the domain boundary. In addition, a special attention is given to construct a second-order ALE algorithm obeying the local DGCL [4]. Furthermore, a more compatible discrete kinematic boundary condition is enforced at the common fluid-structure interface in order to satisfy the global DGCL, which is required in order to conserve total fluid volume/mass at machine precision.

FSI simulations in general can be solved in a monolithic or partitioned way. In the partitioned approach, separate solvers are utilized for the fluid and structure subproblems. The main advantage of the partitioned approach is the ability to reuse existing solvers which allows the application of different, possibly more efficient, computational methods specifically developed for either the fluid or the structure subproblem. Although the implementation of this approach is relatively easy, it does, however, suffer some serious drawbacks. The fixed point iterations tend to converge slowly and the iterations may diverge in the presence of strong fluid-structure coupling due to the high fluid/structure density ratio which causes to the so-called artificial added mass effect [5]. In a fully coupled (monolithic) approach, the fluid and structure equations are discretized and solved simultaneously as a single equation system for the entire problem. However, this requires an efficient numerical technique for the solution of a large system of coupled nonlinear algebraic equations, which poses the major challenge of monolithic FSI approaches, especially in large scale problems. Although monolithic solvers are believed to be too expensive for use in large-scale problems, more recent studies [6, 7] demonstrate that monolithic solvers are competitive even in the case of weak fluid-structure interactions problems. In the present paper, the original system of equations is preconditioned with an upper triangular right preconditioner which results in a scaled discrete Laplacian instead of a zero block in the original system due to the divergence-free constraint. Then a one-level restricted additive Schwarz preconditioner with a block-incomplete factorization within each parti-

tioned sub-domains is utilized for the resulting fully coupled system. The implementation of the preconditioned Krylov subspace algorithm, matrix-matrix multiplication and the restricted additive Schwarz preconditioner are carried out using the PETSc [8] software package developed at the Argonne National Laboratories.

2 MATHEMATICAL AND NUMERICAL FORMULATION

The integral form of the incompressible Navier-Stokes equations over an arbitrary moving control volume are discretized using the arbitrary Lagrangian-Eulerian (ALE) based side-centered finite volume method [2]. In this approach, the velocity vector components are defined at the mid-point of each cell face, while the pressure term is defined at element centroids. The present arrangement of the primitive variables leads to a stable numerical scheme and it does not require any *ad-hoc* modifications in order to enhance the pressure-velocity coupling. In the current discretization, the continuity equation is satisfied within each hexahedral elements at machine precision and the summation of the discrete equations can be exactly reduced to the domain boundary, which is important for the global mass conservation. In addition, a special attention is given to construct a second-order accurate arbitrary Lagrangian-Eulerian algorithm obeying the discrete geometric conservation law (DGCL) [4]. The classical Galerkin finite element formulation is used to discretize the governing solid equations for the compressible Saint Venant-Kirchhoff material. In here, the displacements at any point in the isoparametric hexahedral element are approximated by a linear combination of the displacements at the nodal points of the element. The numerical simulation of fluidstructure interaction problems requires to fulfill two coupling conditions: the kinematic and the dynamic continuity across the fluid-solid interface at all times. The kinematic boundary condition on the fluid-structure interface is driven by requiring continuity of the velocity while the dynamic condition means that the equilibrium equation holds for the surface traction at the common fluid-structure interaction boundary. However, the problem with the classical application of kinematic boundary condition is that the discrete equations are not compatible with the global DGCL and the total volume/mass of the fluid domain will not conserved at machine precision if an incompressible fluid is fully enclosed in the solid domain. A more compatible application of the kinematic boundary condition is given in [1] in order to conserve total fluid volume/mass at machine precision. In the present paper, the original system of equations is preconditioned with an upper triangular right preconditioner which results in a scaled discrete Laplacian instead of a zero block in the original system due to the divergence-free constraint. Then a one-level restricted additive Schwarz preconditioner with a block-incomplete factorization within each partitioned sub-domains is utilized for the resulting fully coupled system. The implementation of the preconditioned Krylov subspace algorithm, matrix-matrix multiplication and the restricted additive Schwarz preconditioner are carried out using the PETSc [8] software package developed at the Argonne National Laboratories. The METIS library [9] is used to partition the computational domain for a balanced domain decomposition.

3 NUMERICAL RESULTS

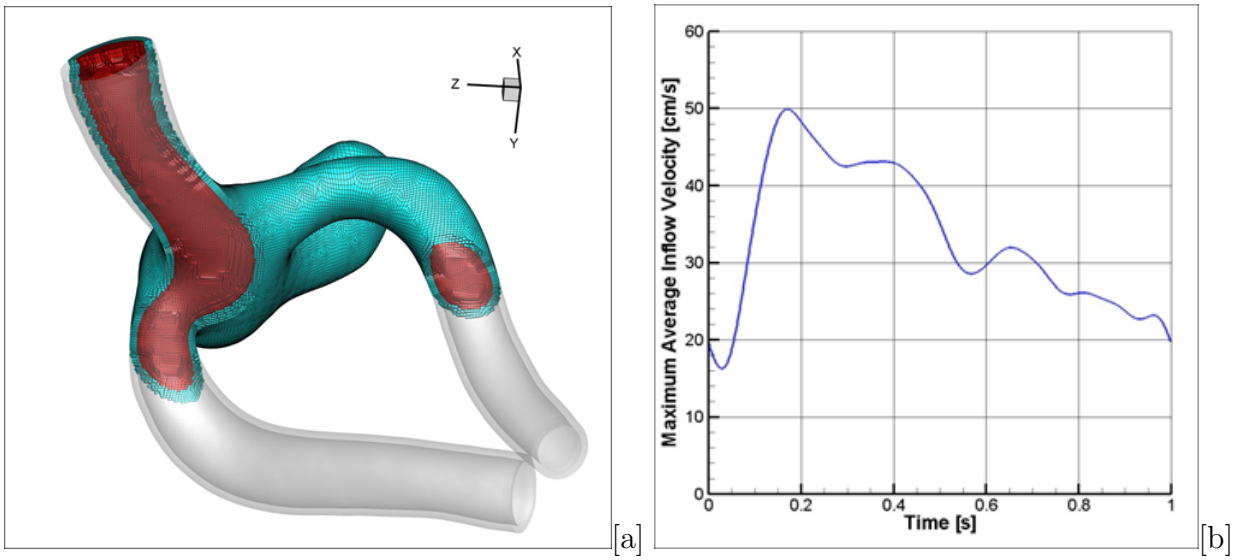
In this section, the numerical method is applied to a more complicated problem involving unsteady pulsatile blood flow in a cerebral artery with aneurysm. The geometry of the arterial lumen is taken from the work of Marchandise et al. [10]. The average radius of the arterial lumen is $0.1075cm$ at the inlet and the values are $0.0855cm$ and $0.0729cm$ at the two outflow ends. All hexahedral conforming mesh generation for the present geometry is rather challenging and the DISTENE MeshGems-Hexa algorithm based on the octree method is used to generate the initial all hexahedral conforming coarse mesh. Then the CUBIT Geometry and Mesh Generation Toolkit developed at the Sandia National Laboratories [11] is employed to refine the initial mesh with optimization-based smoothing techniques. The boundary layer mesh is constructed by extruding the mesh in the outward normal direction. However, this approach is not suitable for the construction of the mesh for the vascular wall due its relatively large thickness which leads to untangled elements. For this purpose, we use the radial basis function (RBF) based mesh deformation algorithm given in [2] which moves the coarsened number of mesh vertices on the lumen surface in the outward normal direction exactly with a distance of $0.02cm$, meanwhile the rest of vertices on the lumen surface vertices are moved using the RBF interpolation. In addition, we added two layers of hexahedral meshes at the inflow and outflow ends in order to make these surfaces planar. The computational mesh shown in Figure 1-a consists of 501,065 vertices and 472,794 hexahedral elements for the fluid domain. The solid domain is constructed using 5 layers of hexahedral elements across the arterial wall and consists of 262,870 vertices and 262,275 elements leading to 7,364,244 DOF for the whole domain.

The boundary condition at the inflow is set to the periodic velocity boundary condition across a single cell distance with the time-dependent prescribed mass flow rate given in [12]. The time variation of the cross-sectional average velocity at the inlet is provided in Figure 1-b. At the outflow the natural (traction-free) boundary condition is imposed. Although more realistic outflow conditions such as the resistive boundary condition [13] may be applied, neither the resistance parameter nor the physical properties of the arterial system is available to us. For the solid domain, the arterial walls are clamped near to the inflow and a Neumann homogeneous boundary condition is applied on the other ends of the structure as in the work of Crosetto et al. [14]. On the exterior surface of the solid domain, we impose the following condition due to the support of exterior tissue [14]

$$\boldsymbol{\sigma}_s \mathbf{n} + \alpha_s \mathbf{d} = 0 \tag{1}$$

with $\alpha_s = 1 \times 10^4$. Rayleigh damping $[C] = \alpha[M] + \beta[K]$ with $\alpha = 6 \times 10^3$ and $\beta = 0$ is also introduced in order to model the damping effect of surrounding tissue. The present simulations are carried out with a constant time step of $0.004s$. In these simulations, the blood is assumed to be a Newtonian fluid and the material properties are provided in Table 1.

The numerical simulation of the blood flow within the present saccular aneurysm is presented by Salmon et al. [15] by assuming that the vessel wall is rigid. The non-dimensional Reynolds number based on the peak cross-sectional average velocity and the trunk radius at the entrance cross section $Re = \rho_f \bar{U}_{max} \bar{R}/\mu_f$ is equal to 159.38. The Womersley number, based on the trunk radius at the entrance cross section $Wo = \bar{R}\sqrt{\omega/\nu_f}$ equals to 1.6 with an angular frequency of $\omega = 2\pi\ rad/s$. The Strouhal number $St = \omega \bar{R}/\bar{U}_{max}$ is 0.016. In the current numerical simulations, the vessel wall is modelled as an isotropic hyperelastic material using the compressible Saint Venant-Kirchhoff model. The present numerical calculations are started impulsively and the numerical solutions presented in here correspond to third cardiac cycle. The velocity vector magnitude contours are presented in several planes normal to $y-$axis along with the stream traces in Figure 2-a at

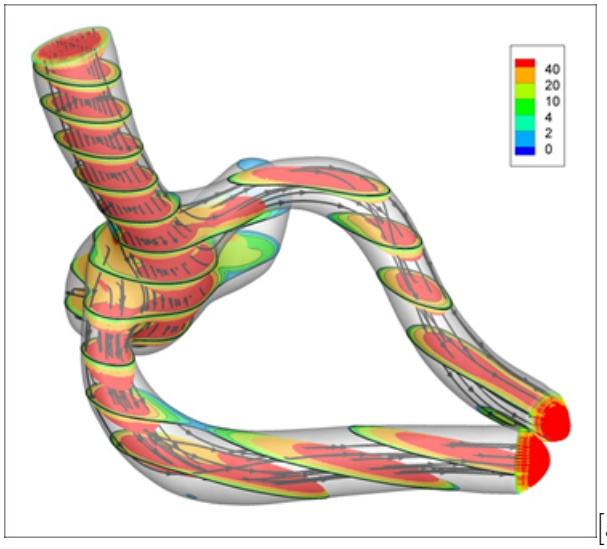


[a] [b]

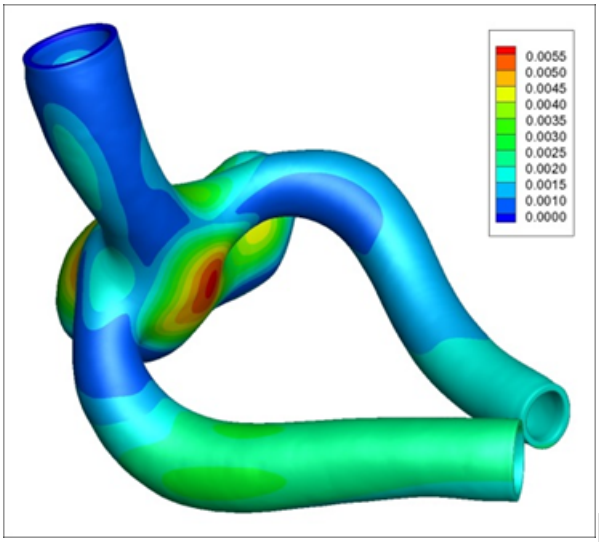
Figure 1: The computational all hexahedral conformal mesh with 735,069 hexahedral elements and 763,935 nodes (7,364,244 DOF) for fluid and solid domains [a] and the time variation of the total mass flow rate at inlet [b].

Table 1: Material properties.

Fluid	Density ρ_f $[g/cm^3]$	1
	Dynamic viscosity μ_f $[g/cm\dot{s}]$	0.04
Solid	Density ρ_s $[g/cm^3]$	1.2
	Poisson ratio ν_s	0.45
	Young modulus E_s $[dynes/cm^2]$	6×10^6



[a]



[b]

Figure 2: The computed velocity vector magnitude contours at several planes normal to $y-$axis with streamtraces [a] and the arterial wall displacement [b] at $t = 2.2s$

$t = 2.2s$. The velocity contours indicate that a low velocity zone exists towards to the end of the saccular aneurysm. In addition, a large swirl is observed within the saccular aneurysm. Furthermore, three-dimensional Dean vortices are observed within the two curved branches of the saccular aneurysm. These three-dimensional flow structures may be seen more clearly from Figure 3-b. The simulations also indicate that the initial artery wall deformation at the inlet section is not radially symmetric due to the elliptical geometry of the inlet cross section. The maximum deformation is observed at the middle of the bifurcation of the branch where the lumen surface curvature is very small, as seen from Figure 2-b. In these simulations, it is important to impose the support due to exterior tissue since the branches tend to move away from each other due to the employed Neumann homogeneous boundary condition at the outlets. The computed wall shear stress (WSS) lines in Figure 3-a indicate separation line just before the saccular aneurysm. Large shear stresses are also observed close to the shoulder of the branch bifurcation. However wall shear stress is rather low towards to the end of the saccular aneurysm where the velocity magnitude is relatively low. It should be noted that the use of present octree based all hexahedral meshes leads to a more effective solution technique since the background mesh consists of mostly uniform Cartesian meshes. Therefore, the combination of the present algorithm with the octree based mesh generation is very efficient for treating complex FSI problems. However, we should mention that there are still open issues related to all hexahedral conformal mesh generation [16].

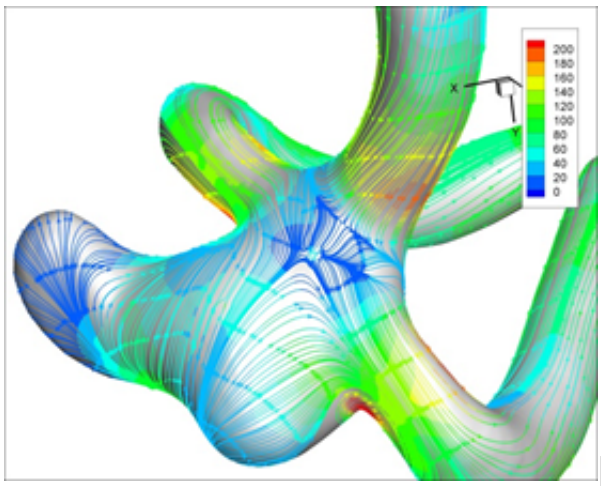
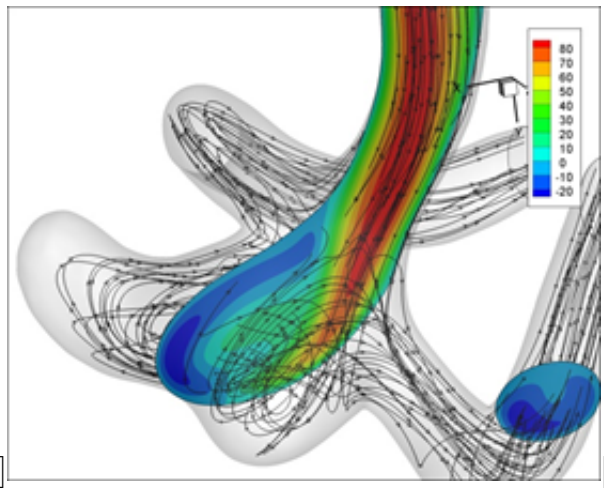

[a]

[b]

Figure 3: The wall shear stress lines [a] and the vertical velocity component contours with the 3d streamtraces [b] at $t = 2.16s$.

4 CONCLUSIONS

The parallel monolithic approach [1] has been applied to the unsteady pulsatile blood flow in a cerebral artery with aneurysm. The governing equations for the incompressible Navier-Stokes equations for the fluid dopmain is discretized using the side-centered finite volume method based on an Arbitrary Lagrangian Eulerian formulation meanwhile the nonlinear Saint Venant-Kirchhoff equations for the solid domain is discretized using the classical Galerkin finite element. The continuity equation is satisfied within each element exactly and the summation of the continuity equations can be exactly reduced to the domain boundary, which is important for the global mass conservation. In addition, a special attention is given to construct a second-order ALE algorithm obeying the DGCL. Furthermore, a more compatible application of kinematic boundary condition is introduced at the common fluid-structure interface in order to conserve total fluid volume/mass at machine precision. The resulting large-scale nonlinear equations from the discretization of fluid and solid domains are solved in a fully coupled form using a monolithic approach based on a one-level restricted additive Schwarz preconditioner with a block-incomplete factorization within each partitioned sub-domains.

5 ACKNOWLEDGMENTS

The authors acknowledge financial support from Turkish National Scientific and Technical Research Council (TUBITAK) through project number 112M107. The authors also gratefully acknowledge the use of the computing resources provided by the National Center for High Performance Computing of Turkey (UYBHM) under grant number 10752009 and the computing facilities at TUBITAK ULAKBIM, High Performance and Grid Computing Center.

REFERENCES

[1] A. Eken and M. Sahin, A parallel monolithic algorithm for the numerical simulation of large-scale fluid structure interaction problems. (submitted).

[2] B. Erzincanli and M. Sahin, An arbitrary Lagrangian-Eulerian formulation for solving moving boundary problems with large displacement and rotations. *Journal of Computational Physics* (2013) **255**:660–679.

[3] Y. H. Hwang, Calculations of incompressible flow on a staggered triangle grid, Part I: Mathematical formulation. *Numer. Heat Transfer B*(1995) **27**:323–1995.

[4] P. D. Thomas and C. K. Lombard, Geometric conservation law and its application to flow computations on moving grids. *AIAA J.* (1979) **17**:1030–1037.

[5] C. Förster and W. A. Wall and E. Ramm, Artificial added mass instabilities in sequential staggered coupling of nonlinear structures and incompressible viscous flows. *Comput. Meth. Appl. Mech. Engrg.* (2007) **196**:1278-1293.

[6] M. Heil and A. L. Hazel and J. Boyle, Solvers for large-displacement fluid structure interaction problems: segregated versus monolithic approaches. *Comput. Mech.* (2008) **43**:91–101.

[7] U. Küttler, M. Gee, Ch. Förster, A Comerford and W. A. Wall, Coupling strategies for biomedical fluid-structure interaction problems. *Int. J. Numer. Meth. Bio.* (2010) **26**:305–321.

[8] S. Balay, J. Brown, K. Buschelman, V. Eijkhout, W. D. Gropp, D. Kaushik, M. G. Knepley, L. C. McInnes, B. F. Smith and H. Zhang, PETSc Users Manual. *ANL-95/11 - Revision 3.3* Argonne National Laboratory (2012).

[9] G. Karypis and V. Kumar, A fast and high quality multilevel scheme for partitioning irregular graphs. SIAM *J. Sci. Comput.* (1998), **20**:359–392.

[10] E. Marchandise, P. Crosetto, C. Geuzaine, J-F, Remacle and E. Sauvage, Quality open source mesh generation for cardiovascular flow simulations. *MS&A Modeling, Simulation and Applications*. Springer Milan, (2012) 395–414.

[11] T. D. Blacker, S. Benzley, S. Jankovich, R. Kerr, J. Kraftcheck, R. Kerr, P. Knupp, R. Leland, D. Melander, R. Meyers, S. Mitchell, J. Shepard, T. Tautges and D. White, CUBIT Mesh Generation Enviroment Users Manual Volume 1. Sandia National Laboratories: Albuquerque, NM (1999).

[12] Thiriet, M. *Biology and Mechanics of Blood Flow. Part II: Mechanics and Medical Aspects*. Springer, (2008).

[13] Y. Wua and X.-C. Cai, A parallel two-level method for simulating blood flows in branching arteries with the resistive boundary condition. *Comput. Fluids* (2011) **45**:92–102.

[14] P. Crosetto, P. Reymond, S. Deparis, D. Kontaxakis, N. Stergiopulos and A. Quarteroni, Fluidstructure interaction simulation of aortic blood flow. *Computers & Fluids* (2011) **43**:46–57.

[15] S. Salmon, M. Thiriet and J.-F. Gerbeau, Medical image-based computational model of pulsatile flow in saccular aneurisms. *ESAIM: Mathematical Modelling and Numerical Analysis* (2003) **37**:663–679.

[16] R. Schneiders, Octree based hexahedral mesh generation. *Int. Journal of Comp. and Applications* (2000) **10**:383–398.

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis (Eds)

ADVANCED IMAGE PROCESSING METHODS FOR AUTOMATIC LIVER SEGMENTATION

PETR STRAKOS[*], MILAN JAROS[†], TOMAS KARASEK[†], TOMAS KOZUBEK[†], PETR VAVRA[††] AND TOMAS JONSZTA[††]

[*†] IT4I, VSB – Technical University of Ostrava
17. listopadu 15, 708 33 Ostrava
Czech Republic
[*] e-mail: petr.strakos@vsb.cz

[††] Faculty of Medicine - University of Ostrava
Syllabova 19, 703 00 Ostrava
Czech Republic

Key words: Image Processing, Automatic Segmentation, Liver, Computed Tomography, Magnetic Resonance Imaging.

Summary. This paper presents advanced methods of image segmentation suitable for automatic recognition of the human liver and its vessel system, but in general could be used to segment any organ or body tissue. The comparison of studied methods is being made in terms of segmentation quality and algorithm speed. The main criterion for quality evaluation of each selected method is the level of conformity between the automatically recognized boundary and the reference boundary specified by experienced user. For all the tests sequences of CT and MRI images were used.

1 INTRODUCTION

Nowadays digital image processing and analysis is extensively used in different areas of human activities. One of these fields is medicine, where the enormous amount of data has to be processed. In diagnostic medicine, radiologists use for example computed tomography (CT) or magnetic resonance imaging (MRI) to help diagnose diseases or to perform surgical procedures. Although those high-end technologies are very sophisticated and developed, there is still room for improvements especially in the area of post-processing. For example to plan the liver resection, surgeon will need an accurate 3D model of a liver with its vessel system. To visualize 3D data, volume rendering method is available and can be used directly. Disadvantage of this technique is that it does not provide any other information than 3D models for visualization itself. To obtain models for hemodynamics simulations or even models where volume of certain part of the model could be measured, methods of advanced image processing have to be used.

The keystone of the digital image processing methods is an image segmentation which works with pixel intensity levels. Segmentation methods like those based on regions (region growing) [1, 2, 3], thresholding (Otsu's methods) [1, 4], clustering (k-means clustering) [1, 5, 6] and others are available. Those techniques used solely without any further processing are

usually sufficient in cases where different parts of the image have significant differences in intensities of pixels like in cases of bones segmentation. The soft tissues like liver and its vessel system are much more challenging due to low contrast and direct influence of other organs which have the same intensity levels of the pixels.

In the paper three main image segmentation algorithms with other pre- and post-processing procedures are used for automatic image segmentation. Segmentation is used to recognize liver tissue and liver vessel system. The first segmentation algorithm is based on region growing [1, 2, 3], the second one works with multi Otsu's algorithm [1, 4] and the third one is based on k-means clustering [1, 5, 6]. In the paper comparison of above mentioned techniques and procedures in terms of segmentation quality and speed is provided. The main criterion for quality evaluation of each segmentation method is the level of conformity between the automatically recognized boundary and the reference boundary made by specialist. For all the tests sequences of CT and MRI images were used. Each modality of the input data had at least two different sources. One served as training data in the preparation phase and the other one was used for testing of the proposed methods. Combining three main segmentation algorithms and selected pre- and post-processing procedures five different methods for liver segmentation were used in total. Multi Otsu algorithm and k-means algorithm were then used to create two methods for segmentation of the liver vessel system.

Paper is organized in the following manner: in section 2 we present detailed description of each automatic segmentation method used for liver and the liver vessel system; in section 3 description of the image data selected for the tests is provided; section 4 describes evaluation method; in section 5 results of the evaluation are shown; section 6 discusses the results; section 7 brings the conclusion.

2 AUTOMATIC SEGMENTATION METHODS

In this section description of the main segmentation methods is provided in more detail. All necessary steps for pre- and post-processing of the data are also described.

Introduced methods are considered as automatic in case we know at least one point within the segmented object. In case of liver segmentation it is a point inside the liver. For segmentation of vessel system we have to define one point inside the vessel system. Methods for determination of such point or points are described in more details in [7, 8].

2.1 Region growing algorithm

In general region growing algorithm is one of the region based techniques. This algorithm partitions an image into regions that are similar according to a set of predefined criteria [1, 2, 3].

In our case we use a region growing algorithm that groups every pixel around the seed point to one group. To add new pixel to the group, intensity of that particular pixel is evaluated and it is added to the group only if its intensity does not exceed the mean intensity of the group more than the tolerance. The region is grown in 4-neighbourhood manner (north, south, east and west) around the seed pixel, while in total 8-neoghbours are available around the pixel.

2.2 Multi Otsu algorithm

Multi Otsu algorithm [1, 4] uses multilevel thresholding to segment the image. It works with pixel intensity levels. Histogram of an image is divided into *M* classes. The number of classes is user defined parameter. To divide the histogram into *M* classes *M*-1 thresholds are necessary. Optimal threshold values are chosen is such a way that between-class variance is maximized. Every pixel in the image is then classified into one of the *M* classes based on the rule described in (1).

$$\begin{aligned} g(i,j) &= 1 \quad for \quad f(i,j) < T_1 \\ &= 2 \quad for \quad T_1 \le f(i,j) < T_2 \\ &\vdots \quad \vdots \quad \vdots \quad \vdots \\ &= M \quad for \quad T_{M-1} \le f(i,j) \end{aligned} \tag{1}$$

where *f(i,j)* is image pixel and *g(i,j)* is pixel after classification into one of the *M* classes.

2.3 K-means clustering

K-means clustering is image segmentation algorithm based on clustering. This algorithm partitions *n* pixels into *k* clusters, where *k* is integer value that holds $k<n$. K-means algorithm classifies pixels in an image into *k* number of clusters according to some similarity feature like grey level intensity of pixels and distance of pixel intensities from centroid pixel intensity [1, 5, 6]. The algorithm works in the following way:

(i) Selection of *k* clusters (it is a user defined parameter)

(ii) Calculation of the total number of image pixels *N*

(iii) Random selection of *k* initial pixel intensity centroids μ_j

(iv) Calculation of distances D_{ij} between pixel x_i and each centroid μ_j as in (2). Particular pixel x_i is then classified to cluster c_j to which centroid it has the smallest distance

$$D_{ij} = (x_i - \mu_j)^2 \tag{2}$$

where $i=1 \div N$ and $j=1 \div k$.

(v) Re-calculation of centroid positions μ_j as a mean value from all pixel intensities which belong to cluster c_j.

$$\mu_j = \frac{1}{l_j} \sum_{i=1}^{l_j} x_i \tag{3}$$

where l_j is a number of pixels that belong to cluster c_j.

(vi) Steps (iv) and (v) are repeated until classification of image pixels does not change or equivalently centroids do not move.

2.4 Pre-processing procedures

Before image segmentation could be performed images have to be pre-processed to reduce noise. Different noise reducing filters are available. In this paper adaptive Wiener filter was used for this purpose. Adaptive Wiener filter is based on statistics estimated from a local

neighborhood of each pixel [9].

Determination of necessary parameters for the three segmentation algorithms to run without further user intervention could be also considered as one of the steps of the pre-processing stage. In case of region growing algorithm optimal intensity tolerance was determined. In case of multi Otsu algorithm optimal number of dividing classes was determined and in case of k-means algorithm optimal number of segmenting clusters was determined. During the preparation phase series of ten images was segmented. Optimized parameter of each segmentation algorithm has varied in particular user defined band over each image and volumetric overlap error (VOE) (see section 4) was evaluated for every segmentation. Lowest sum of VOE values over all images specified the optimal value of the parameter to be used in the test phase. This part of pre-processing can be thus considered as a reference model for the segmentation. The reference model supplies necessary parameters so the segmentation of test images can run automatically.

2.5 Post-processing procedures

To obtain boundaries of segmented objects flooding of segmented regions is performed. This is the main post-processing operation. Extracted boundaries can then be approximated by closed cubic spline. This is the case of liver segmentation where such smoothing of the boundary suits the expected shape of the liver. In case of vessel system segmentation only area flooding with the boundary extraction without spline approximation was used.

For area flooding with boundary extraction we used modification of the basic 4-directional flood-fill algorithm [10]. Modification resides in marking the boundary pixel every time the flooding boundary is met and not only flooding the area without creating the boundary. User can also size the flooding element of square shape above the 1 pixel minimum. This modification is used to minimize the flooding of the areas with narrow connections to the main area. These areas usually occur in over-segmented images (more area is being segmented as part of the segmented object).

For the spline approximation we used periodic cubic spline [11] to smooth the segmented boundary of the liver.

Based on the specified post-processing procedures five different methods to segment the liver and two different methods to segment the vessel system have been evaluated in this paper. For liver segmentation it was:

(i) Multi Otsu method with 8 or 5px flooding element (Multi Otsu 8(5)px flood)
(ii) Multi Otsu method with 1px flooding element (Multi Otsu)
(iii) K-means method with 8 or 5px flooding element (K-means 8(5)px flood)
(iv) K-means method with 1px flooding element (K-means)
(v) Region growing method (Region growing)

For vessel system segmentation it was:

(i) Multi Otsu 1px flooding element without spline approximation (Multi Otsu)
(ii) K-means 1px flooding element without spline approximation (K-means)

3 SELECTED IMAGE DATA

All evaluated methods were tested on two different modalities of the image data. We used data from the computed tomography (CT) and from the magnetic resonance imaging (MRI). All CT data had resolution of 512x512 pixels while MRI data had resolution 256x256 pixels. Publicly available DICOM database [12] was used as a source of data. Presented methods were used to segment the liver and the liver vessel system. Evaluated methods first proceeded through the preparation phase to establish settings which were then used in the test phase. Beside the different modalities, images from different patients were used as well. In preparation phase methods worked with series of ten images while the test phase performed on two images. Images in preparation phase always differ from those in the test phase. In Figure 1 example of two different modalities of the image data is shown.

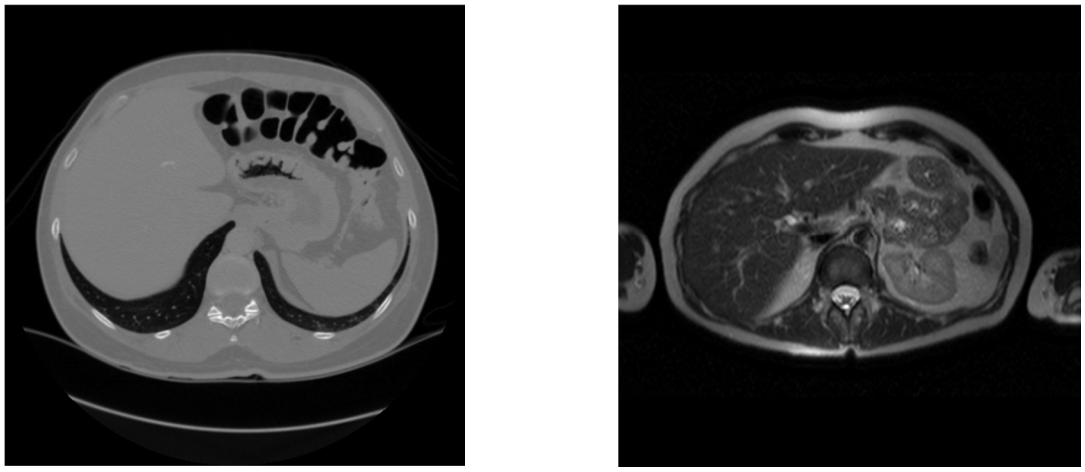


Figure 1: Example of the CT (left) and MRI (right) modality of the used image data

4 EVALUATION METHOD

Comparison of presented segmentation algorithms is based on computational time and segmentation quality.

Since each method ran in the preparation phase to establish important setting values which are then used in the subsequent test phase, two different runtime values were measured. First was the preparation phase runtime and the second was the test phase runtime.

Concerning the segmentation quality, several metrics are available [13, 14] which compare the segmentations with the manually segmented reference made by trained specialist. In this paper we use two metrics based on volumetric overlap. It is volumetric overlap error (VOE) and relative volume difference (RVD). As the main criterion for quality evaluation of each selected technique VOE is used. In section 5 where results are presented we use the mean VOE of two test images for the quality evaluation.

4.1 Volumetric overlap error

The volumetric overlap error (VOE) is given in percent and it calculates the error between two sets of pixels A and B as described in (4).

$$VOE = 100 \cdot \left(1 - \frac{|A \cap B|}{|A \cup B|}\right) \tag{4}$$

One of the pixel sets stands for the reference and the other one for the segmented set. Which set is which, whether A or B, does not matter as we calculate union and intersection between the sets. The 0 value of the volumetric overlap error means that there is a perfect match between the sets A and B. Value of 100 means that the pixel sets A and B do not overlap at all. The reason why this metric was used is that it is one of the most popular method for evaluation of the segmentation accuracy.

4.2 Relative volume difference

The relative volume difference (RVD) is also given in percent and it is calculated between two sets of pixels A and B by the equation (5).

$$RVD = 100 \cdot \left(\frac{|A| - |B|}{|B|}\right), \tag{5}$$

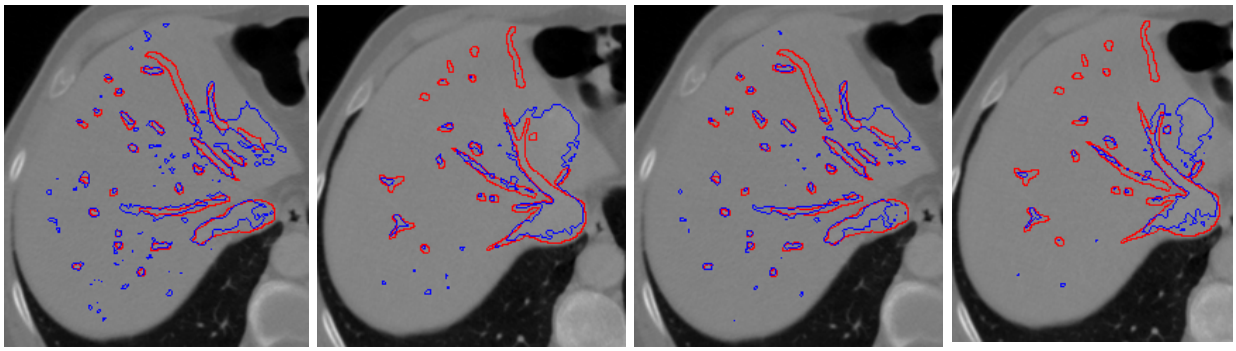
In equation (5) A stands for the segmented pixel set and B stands for the reference pixel set. The 0 value of equation (5) means both pixel sets have the same volume and in this sense they are identical. It has to be noted that this does not imply that A and B are identical, or overlap with each other. This is the drawback of the RVD method in comparison with VOE. On the other hand the RVD metric gives direct volumetric information. The RVD is a signed value and thus it also gives a good insight whether the method tends to over- or under estimate total volume. This is the main reason why RVD metric was used in this paper.

5 RESULTS

As mentioned earlier, tests were performed on two different images in each modality (CT and MRI) for liver segmentation and also on two images in each modality for vessel system segmentation. Results in Table 1 show numerical values for CT vessels segmentation. Images after segmentation are depicted in Figure 2. In Table 2 and Figure 3 results of CT liver segmentation are provided. Results for MRI modality are listed in Table 3 and Figure 4 in case of vessel segmentation. Results for MRI liver segmentation are listed in Table 4 and Figure 5. All results were obtained on PC with Intel Core i3, 1.9 GHz, with 4GB of RAM and MATLAB R2014a.

Table 1: Vessel segmentation results on two test CT images for evaluated methods. Mean VOE value calculated as mean from VOE values on each image.

	Prep. Runtime (~100 iterations) [s]	Runtime Image 1 [s]	VOE Image 1 [%]	RVD Image 1 [%]	Runtime Image 2 [s]	VOE Image 2 [%]	RVD Image 2 [%]	Mean VOE [%]
K-means	1174	6.1	57.959	11.640	5.3	62.014	7.931	59.987
Multi Otsu	28	2.9	58.576	-14.768	2.3	65.825	-27.500	62.201



a) b) c) d)

Figure 2: Vessel segmentation results on two CT test images for evaluated methods. Red line is reference, blue line is segmentation. a) K-means, image 1; b) K-means, image 2; c) Multi Otsu, image 1;d) Multi Otsu, image 2

Table 2: Liver segmentation results on two test CT images for evaluated methods. Mean VOE value calculated as mean from VOE values on each image.

	Prep. Runtime (~100 iterations) [s]	Runtime Image 1 [s]	VOE Image 1 [%]	RVD Image 1 [%]	Runtime Image 2 [s]	VOE Image 2 [%]	RVD Image 2 [%]	Mean VOE [%]
Multi Otsu, (8px flood)	355	4.8	4.623	-2.966	3.9	4.937	2.158	4.780
K-means, (8px flood)	1190	10.5	4.860	-3.770	9.9	5.407	-2.162	5.134
K-means	1190	13.0	6.044	-2.077	12.5	7.811	7.765	6.928
Multi Otsu	355	5.0	6.959	3.131	5.4	7.943	7.912	7.451
Region growing	499	5.3	6.186	-5.675	5.7	9.613	-8.493	7.900

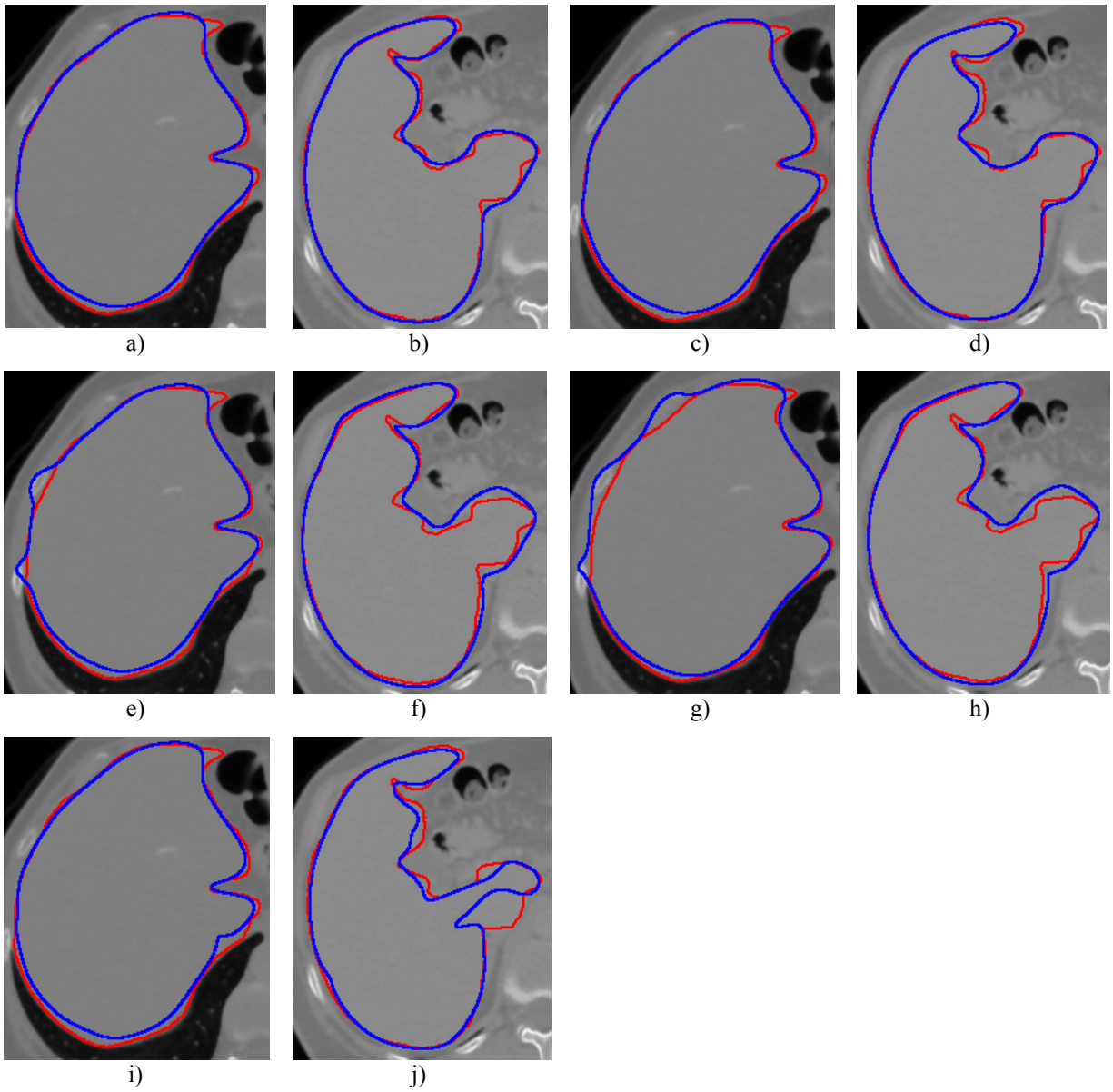
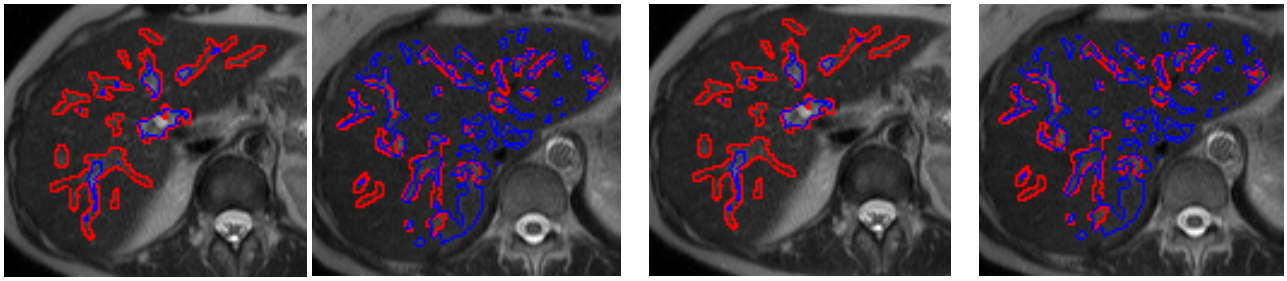


Figure 3: Liver segmentation results on two CT test images for evaluated methods. Red line is reference, blue line is segmentation. a) Multi Otsu (8px flood), image 1; b) Multi Otsu (8px flood), image 2; c) K-means (8px flood), image 1; d) K-means (8px flood), image 2; e) K-means, image 1; f) K-means, image 2; g) Multi Otsu, image 1; h) Multi Otsu, image 2; i) Region growing, image 1; j) Region growing, image 2

Table 3: Vessel segmentation results on two test MRI images for evaluated methods. Mean VOE value calculated as mean from VOE values on each image.

	Prep. Runtime (~100 iterations) [s]	Runtime Image 1 [s]	VOE Image 1 [%]	RVD Image 1 [%]	Runtime Image 2 [s]	VOE Image 2 [%]	RVD Image 2 [%]	Mean VOE [%]
Multi Otsu	12	1.2	78.681	-78.681	0.9	62.334	27.632	70.508
K-means	230	1.7	79.147	-79.147	0.9	62.589	25.804	70.868



a) b) c) d)

Figure 4: Vessel segmentation results on two MRI test images for evaluated methods. Red line is reference, blue line is segmentation. a) Multi Otsu, image 1; b) Multi Otsu, image 2; c) K-means, image 1;d) K-means, image 2

Table 4: Liver segmentation results on two test MRI images for evaluated methods. Mean VOE value calculated as mean from VOE values on each image.

	Prep. Runtime (~100 iterations) [s]	Runtime Image 1 [s]	VOE Image 1 [%]	RVD Image 1 [%]	Runtime Image 2 [s]	VOE Image 2 [%]	RVD Image 2 [%]	Mean VOE [%]
Multi Otsu, (5px flood)	72	1.3	11.538	2.035	1.3	18.585	3.350	15.062
K-means	321	3.1	17.940	-10.940	3.4	22.167	2.849	20.054
K-means, (5px flood)	321	2.6	32.661	-31.220	3.0	12.308	7.278	22.485
Region growing	96	1.9	26.490	-20.432	1.5	18.632	3.530	22.561
Multi Otsu	72	3.0	22.176	21.340	1.8	23.337	19.497	22.757

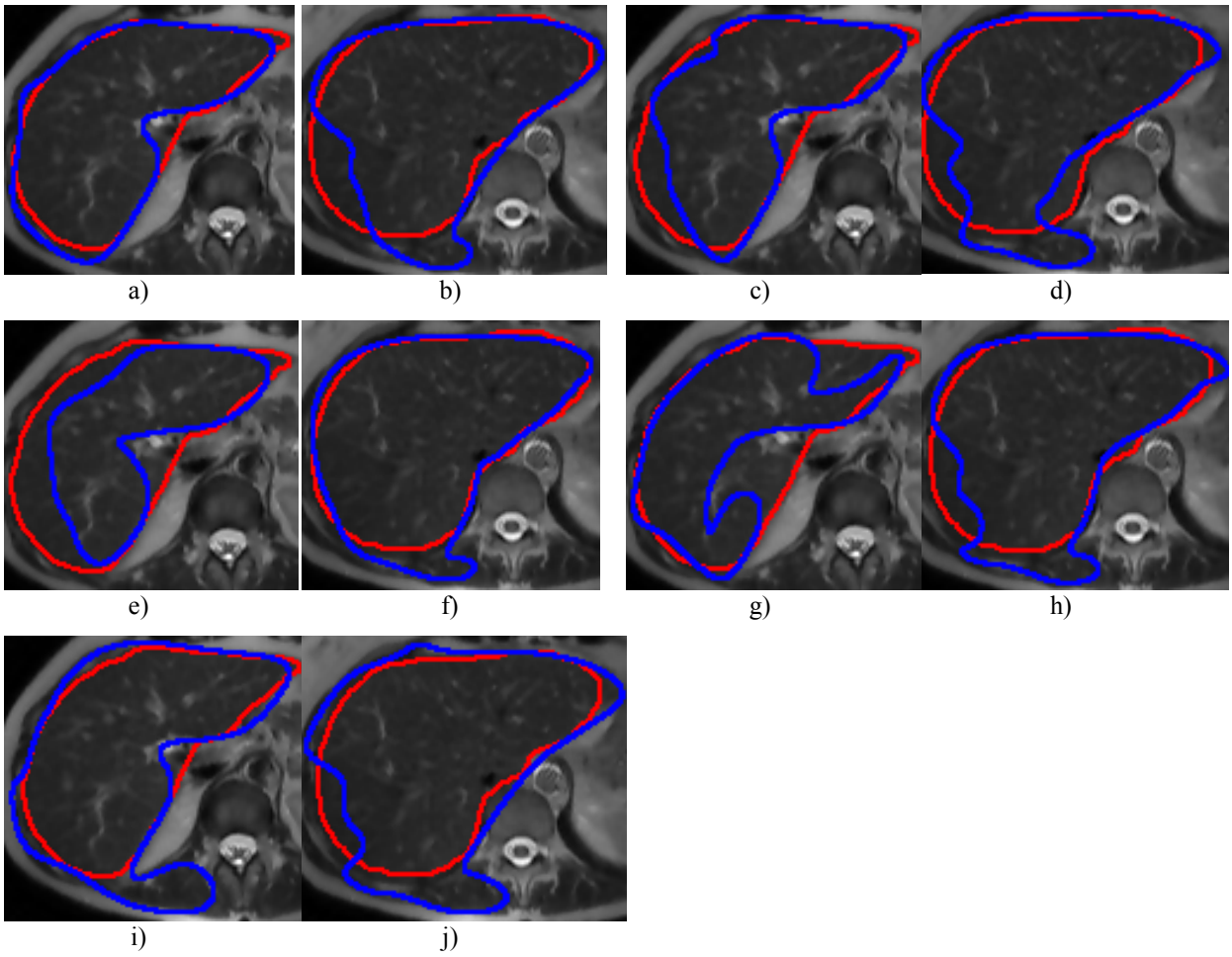


Figure 5: Liver segmentation results on two MRI test images for evaluated methods. Red line is reference, blue line is segmentation. a) Multi Otsu (5px flood), image 1; b) Multi Otsu (5px flood), image 2; c) K-means, image 1; d) K-means, image 2; e) K-means (5px flood), image 1; f) K-means (5px flood), image 2; g) Region growing, image 1; h) Region growing, image 2; i) Multi Otsu, image 1; j) Multi Otsu, image 2

6 DISCUSSION

By comparing automatic segmentation methods we could observe that the best results are obtained on CT modality when segmenting the liver tissue itself. CT modality brings two times higher resolution in comparison with MRI (see section 3). The lowest VOE value of 4.780 % is obtained by multi Otsu (8px flood) method. Possibility to restrict the flood fill algorithm to 8 pixels helps a lot since the plain method with 1 pixel flooding tends to over-segment. This is true also for k-means method in CT liver segmentation, see Table 2 and Figure 3. On the other hand region growing method tends to under-segment as can be seen in Table 2. This is why restriction of flood fill algorithm to higher pixel values has not been used for region growing method.

For the MRI modality of the liver tissue segmentation, multi Otsu (5px flood) is a best

performing method with VOE value of 15.062 %. Restriction of the flood fill algorithm does bring improvements only in the case of multi Otsu method which tends to over –segment in its non-restricted version.

From Table 1 and Table 3 can be observed that segmentation of the liver vessel system performs better on CT than MRI images. In case of CT the best VOE value of 59,987 % is produced by k-means method as shown in Table 1. In case of MRI images segmentation methods show almost identical results with VOE around 71 % as visible from Table 3.

Concerning the runtimes, long runtimes in preparation phase can be explained by high number of evaluations (10 evaluated images, each image processed approx. 10 times). Highest runtimes occur for k-means method. This is caused by the fact that k-means method execution time depends on initial setting of cluster centroids. If initial setting is far away from optimum, k-means takes long.

7 CONCLUSION

It has been shown that best performing segmentation method for liver segmentation is multi Otsu method with restriction of flood fill algorithm.

There is significant difference in segmentation quality when segmenting CT or MRI modality. This is due to two times less image information in MRI images in comparison to CT images. Therefore CT images appear as more suitable for the used segmentation techniques.

Segmentation of the liver vessel system does not show any particular advantage of any used methods. Slightly better results are obtained in case of CT data, which can be again explained by the fact that CT data contain two times more image information than MRI data. Problems in vessel segmentation are caused mainly by the presence of noise. It happens that small vessels simply disappear in the noise.

For a future work tested methods need to be effectively applied to the sequence of consecutive images covering the whole liver area. For segmentation of the liver vessel system different and more proper methods have to be searched since the presented methods are not very effective.

8 ACKNOWLEDGMENT

This paper has been elaborated in the framework of the project New creative teams in priorities of scientific research, reg. no. CZ.1.07/2.3.00/30.0055, supported by Operational Programme Education for Competitiveness and co-financed by the European Social Fund and the state budget of the Czech Republic.

REFERENCES

[1] Sonka, M., Hlavac, V., Boyle, R. *Image Processing, Analysis and Machine Vision*. Thomson, (2006). ISBN 978-0-495-24428-7.

[2] Neri, E., Caramella, D., Bartolozzi, C. *Image Processing in Radiology*. Springer, (2008). ISBN 978-3-540-49830-8.

[3] Terry Y. *Insight into Images: Principles and Practice for Segmentation, Registration, and Image Analysis*. A K Peters/CRC Press, (2004), ISBN 978-1568812175.

[4] Liao, P-S., Chung, P-C. A fast algorithm for multilevel thresholding, *Journal of*

Information Science and Engineering (2001) **17** (5):713-727.

[5] Dass, R., Priyanka, Devi, S. Image Segmentation Techniques. *IJECT* (2012) **3** (1):2230-7109.

[6] Luo, S., Li, X., Li, J. Review on the Methods of Automatic Liver Segmentation from Abdominal Images. *Journal of Computer and Communications* (2014) **2**, pp 1-7. DOI: 10.4236/jcc.2014.22001.

[7] Ruskó L., Bekes G., Németh G., Fidrich M. Fully automatic liver segmentation for contrast- enhanced CT images. In: *Proceedings of MICCAI workshop on 3D segmentation in the clinic: a grand challenge*, pp 143–150.

[8] Juszczyk, J., Piętka, E. Automatic generation of initial points for CT abdominal organ segmentation. *International Journal of Computer Assisted Radiology and Surgery* (2013) **8** (1 Supplement):263-336. DOI 10.1007/s11548-013-0880-0.

[9] Jin, F., Fieguth, P., Winger, L.; Jernigan, E. Adaptive Wiener filtering of noisy images and image sequences. *ICIP* (2003) **3**:1522-4880. DOI:10.1109/ICIP.2003.1247253.

[10] Flood fill, 2014. Available from: <http://en.wikipedia.org/wiki/Flood_fill>. [19 January 2015].

[11] Graham, N. Y. Smoothing With Periodic Cubic Splines. *Bell System Technical Journal* (1983) **62** (1), pp 101–110.

[12] DICOM files, 2014. Available from: <http://www.osirix-viewer.com/datasets/>. [16 January 2014].

[13] Heimann, T. et al. Comparison and evaluation of methods for liver segmentation from CT datasets. *IEEE Transactions on Medical Imaging* (2009) **28** (8):1251–1265.

[14] Mharib, A., Ramli, A., Mashohor, S., Mahmood, R. Survey on liver CT image segmentation methods. Springer Netherlands. In: *Artificial Intelligence Review* (2012) **37** (2):83-95. DOI: 10.1007/s10462-011-9220-3.

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis (Eds)

VALIDATION OF HIGH FIDELITY CFD/FE FSI FOR FULL-SCALE HIGH-SPEED PLANING HULL WITH COMPOSITE BOTTOM PANELS SLAMMING

SILVIA VOLPI[*], HAMID SADAT-HOSSEINI[*], MATTEO DIEZ[*§], DONG-HWAN KIM[*], FREDERICK STERN[*#], ROBERT S. THODAL[†], AND JOACHIM L. GRENESTEDT[†]

[*]IIHR-Hydroscience & Engineering
The University of Iowa, Iowa City, IA 52242, USA

[§]National Research Council-Marine Technology Research Institute (CNR-INSEAN)
Rome, 00128, Italy

[†]Department of Mechanical Engineering and Mechanics
Lehigh University, Bethlehem, PA 18015, USA

[#]Corresponding author. Email: frederick-stern@uiowa.edu

Key words: slamming, fluid structure interaction, computational fluid dynamics, finite elements

Abstract. High fidelity CFD/FE FSI (Computational Fluid Dynamics/Finite Element Fluid-Structure Interaction) code development and validation by full-scale experiments is presented, for the analysis of hydrodynamic and structural slamming responses. A fully instrumented 9 meter high speed-planing hull with sterndrive is used. Starboard and port bottom panels are constructed with different composite materials and fiber orientations, allowing for study of the relation between structural properties and slamming. The code CFDShip-Iowa is employed for CFD simulations and the commercial FE code ANSYS is used as structural solver. The hydrodynamic simulations include captive (2DOF without sterndrive) and 6DOF free running conditions for various Froude numbers in calm water and regular waves. Calm water simulations compares well with the experimental data and 1D empirical data provided by the sterndrive manufacturer for resistance, heave, pitch and roll motions. Numerical one-way coupling FSI is performed in head and following regular waves representative of sea-trial conditions, using FE models for two bottom panels. The resulting strains are compared with experimental data showing a good qualitative and quantitative agreement.

1 INTRODUCTION

Slamming impact loads are a critical factor in the structural design, performance and safety of ships, especially for high speed planing hulls. The complex physics of the fluid-structure interactions are not well understood. Experimental studies have primarily involved wedge drop tests, while model or full-scale ship test data is limited. USNA model planing hull slamming pressures and accelerations are reported in [1]. Current prediction methods are

largely empirical or use analytical [2] or potential flow-FE methods, often for 2D sections or idealized geometries. Studies have demonstrated the effectiveness of CFD in slamming analysis, including uncertainty quantification for regular/irregular waves using the Delft catamaran [3] and validation for regular/irregular waves for Fridsma model [4] and USNA model including slamming pressures [5]. Most slams show both primary re-entering (the bow enters the wave face) and secondary emerging (the bow impacts the wave crest) pressure peaks, whereas some show only re-entering pressure peaks, which is more typical of wedge drop and full-scale displacement ship test data. Extreme event slams (about twice standard deviation) correlate with three consecutive incoming wave lengths close to ship length with large steepness. Accelerations and pressure display Froude scaling.

The present collaborative research utilizes an instrumented slamming load test facility (high-speed planing hull - Numerette) for full-scale experimental validation of high-fidelity CFD/FE fluid-structure interaction.

The simulation environmental conditions model the experiments for head and following waves. Hydrodynamic calm water and seakeeping validation uses limited Numerette data along with 1D Mecury Marine system based predictions and other planing hull data, respectively. Initial one-way coupling fluid-structure interaction validation uses Numerette strain data from strain gages embedded in bottom composite sandwich panels, collected by an onboard data acquisition system.

Figure 1: Slamming load test facility

2 COMPUTATIONAL METHODS

The FSI study is performed by means of CFD/FE coupling routines. One-way coupling is realized by application of the hydrodynamic loads on the structure. CFDShip-Iowa V4.5 [6] is used as high-fidelity solver for the flow field, whereas ANSYS Mechanical APDL V14.5 is used to solve the structural displacements and strains.

The CFDShip-Iowa is an overset, block structured CFD solver designed for ship applications. Absolute inertial earth-fixed coordinates are employed with turbulence model k-ε/k-ω based isotropic and anisotropic RANS. A single-phase level-set method is used for free-surface capturing. Dynamic overset grids use SUGGAR to compute the domain connectivity.

ANSYS Mechanical is a comprehensive commercial code for structural FE analysis. A fully transient dynamic analysis is used to determine the dynamic response of the structure under the action of time-dependent loads. It includes structural nonlinearities and utilizes the Newmark time integration method to solve the FE equations.

The one-way coupling method consists of computing the forces acting on the structure using CFD, assuming rigid-body motion of the entire ship, and then applying the forces on the elastic model of the panels. The response is determined in one way, since the deformed geometry is not fed back into the CFD solver. In a two-way coupling approach, the flow field and the elastic deformations are computed by feeding back the elastic motions of the structure into the CFD solver. A tradeoff between one- and two-way coupling methods consists in extending the former, using the wet elastic modes of the structure. This requires the modeling of added mass and damping due to the elastic deformation of the body in water. The

acceleration of the water due to the body deformation is not taken into account in the CFD solver. In general, the use of a feedback (two-way coupling) is required when large deformation significantly affects the flow field. In this work, the one-way coupling method is used for preliminary qualitative/quantitative analysis and comparison with experimental data.

Specifically, CFDShip-Iowa provides the hydrodynamic loads in terms of distribution of force per unit area over the ship hull surface. The force distribution is given for the CFD grid points and in the CFD coordinate system. A coordinate transformation is applied in order to provide the force distribution on the FE model, which has its own coordinate system. The interpolation of the loads on the FE grid points is carried out using Gaussian quadrature. The structural problem is solved by ANSYS for displacements, strains and stresses. CFD/FE numerical results are validated by comparison with experimental strain data.

3 EXPERIMENTAL SETUP

The slamming load test facility is a 9 meter long 1.9 meter wide steel/composite boat designed and manufactured by Grenestedt [7]. The boat structure consists of a welded AL-6XN stainless steel frame and composite sandwich panels. The boat has a top speed of approximately 27 m/s and a full load displacement of 2450kg.

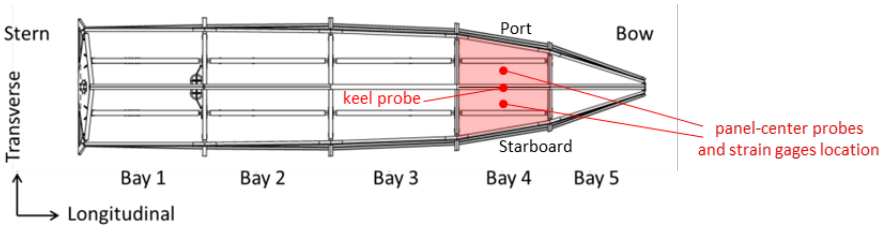


Figure 2: Slamming load test facility layup

To facilitate comparison of different panel constructions, the 10 bottom panels have varied composite layups. All bottom panels are vacuum infused with vinyl ester resin and use a Divinycell H250 foam core but vary in both reinforcement types and fiber direction. The results presented will focus on the behavior of panels in bay 4. The layup of these panels is given in Table 1.

Table 1: Slamming load test facility bay 4 bottom panel layup

	Bay 4 Port Panel	**Bay 4 Starboard Panel**
Top	2 layers DBL700 (0°, ±45°)*	2 layers DBL700 (0°, ±45°)**
↓	Divinycell H250 Foam core	Divinycell H250 Foam core
	3 layers DBL700 (0°, ±45°)*	3 layers DBL700 (0°, ±45°)**
Bottom	1 layer L(X) 440-C10 (0°)*	1 layer L(X) 440-C10 (0°)*

* 0° parallel to keel
** 0° perpendicular to keel

Devold AMT DBL700 triaxial carbon and L(X) 440-C10 unidirectional carbon reinforcements are used in both port and starboard bay 4 panels, but the orientation of the DBL700 differs resulting in a large difference in stiffness.

The slamming load test facility is instrumented with strain gages on both inner and outer skins of the bay 4 bottom panels. Vishay CEA-06-250UN-350/P2 and CEA-06-250UT-

350/P2 gages were used in quarter bridge configuration to measure strain parallel and perpendicular to the keel in the center of each panel. National Instruments NI-9237 signal conditioning and ADC modules were used to acquire 24 bit strain data at 50 kHz per channel. This data was filtered to 5 kHz in post processing.

Modal tests of the dry structure were conducted by exciting the panels at a number of grid points using an instrumented impact hammer and measuring the response with an accelerometer. The least squares complex exponential method was used to extract modal parameters. A National Instruments NI-9234 signal conditioner was used with the PCB Piezotronics 086c03 modal analysis impact hammer and 352c04 accelerometer.

Sea trials were conducted in the Atlantic Ocean near the Barnegat Inlet in Barnegat Light, NJ. Multiple tests, each with duration of 5-10 minutes, were performed. The strain gages were zeroed when the craft was at rest before each test sequence. The vessel was then accelerated to the maximum speed allowable in the conditions. Steering input was used to achieve as close to neutral roll angle as possible. The vessel has since been outfitted with a trim tab to control roll angle. Test segments consisted of a single loop. Data acquisition was stopped when the vessel returned to the approximate starting position.

4 COMPUTATIONAL SETUP

The total number of grid points for CFD simulations with sterndrive is 18.2 M (Figure 3). For bare hull simulations, symmetry with respect to the longitudinal plane is imposed; accordingly, the grid includes only the starboard side of bare hull and half-domain background with 6.94 M grid points.

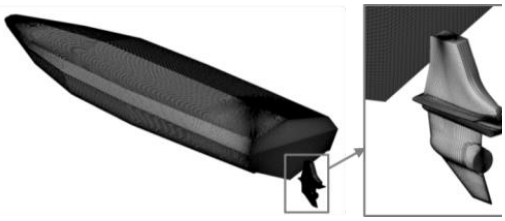


Figure 3: CFD model of Numerette with detailed view of the sterndrive

During the experimental tests, the ship experiences irregular wave, variable heading, and variable speed. A CFD captive regular-wave simulation is used to model the irregular wave pertaining to real-sea conditions [3]. Available information about test conditions includes: sea state 3 conditions; nearly head waves in the first segment of the trial; ship trajectory and speed.

Since the sea trial consists of a single loop, two segments are selected that present alignment between the ship trajectory and the wave direction. The segments are used to model a regular head wave (S1) and following wave (S2) simulations and they are taken as a benchmark for validation. The speed used in S1 and S2 is the average speed $\bar{V}$ of the trial within the selected segments. The wave height is defined as the most probable condition associated to the Bretschneider spectrum, representing a fully-formed sea state 3 (see, e.g., [3]). The wave angular frequency ω is derived by $\omega_e = \omega - (\omega^2 \bar{V}/g) cos\bar{\theta}$, where ω_e is the encounter angular frequency determined as the frequency

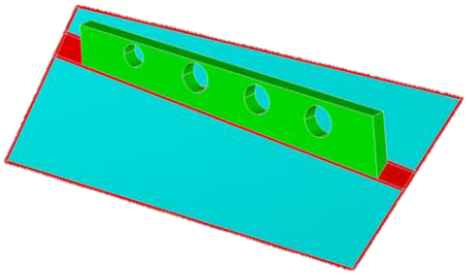


Figure 4: FE model of a Bay 4 panel showing cored sandwich areas (blue), single skin areas (red) and hollow steel longeron (green)

associated with the FFT-peak of the experimental strain, g is the gravity acceleration, and $\bar{\theta}$ is the heading angle ($\bar{\theta} = 0$ for following seas), which is approximated assuming the ship longitudinal axis always aligned with the trajectory.

Finite element models were developed for the slamming load test facility port and starboard bay 4 composite sandwich bottom panels extending from the keel to the chine and from the aft vertical bulkhead in bay 4 to the fore vertical bulkhead in bay 4 (Figure 4). The panel model consists of a sandwich cored region, a perimeter with only composite skins and the stainless steel longeron. All areas were modelled with Shell99 elements in ANSYS. The model is constrained in X,Y,Z displacement at the keel and chines, Y,Z displacement at the bulkheads and Y,Z displacement at the ends of the longerons. The total number of grid points is 51,648. The model was validated by comparison with experimental modal tests and static displacement tests.

5 EXPERIMENTAL ANALYSIS

Experimental data collected from operation of the slamming load test facility during a 400 second duration test is presented here. The position track and speed over the duration of the tests are shown in Figure 5. Head (S1) and following (S2) wave segments are also identified.

The transverse and longitudinal strains measured at the center of the port panel are shown in Figure 6. The port panel strain waveforms for a typical single slamming event are shown in Figure 7. Also shown are the maximum strains for the 100 highest slamming events, used for comparison with CFD/FE results from regular wave simulations.

The highest magnitude strains are seen on the inner skins transverse to the keel. Strains on the inner skins are primarily in tension indicated by positive strain, while the outer skins are under compression indicated by negative strain. The mean value of the highest 1/3 strain peak strains identified during the S1 head wave segment, S2 following wave segment and full test duration are indicated in Table 2.

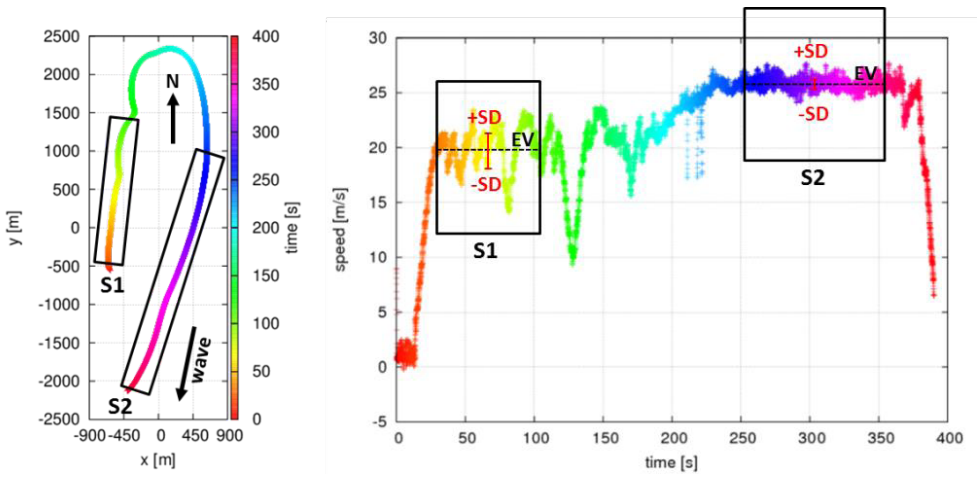


Figure 5: Sea trial trajectory and speed with color mapping for time

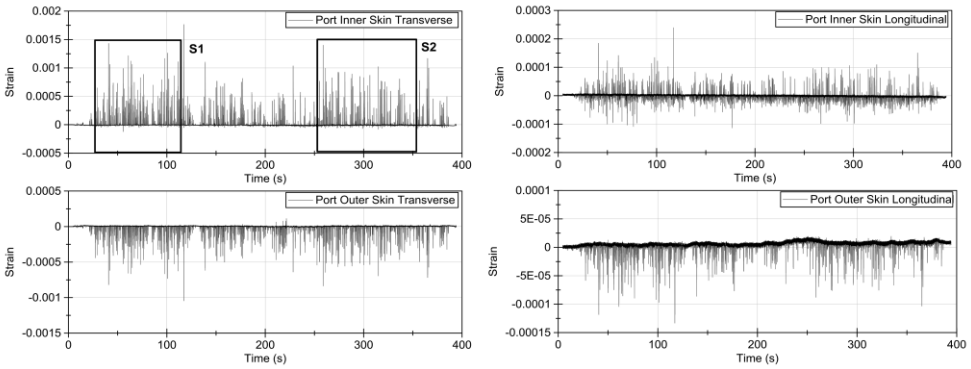


Figure 6: Port bay 4 panel transverse (left) and longitudinal (right) strains

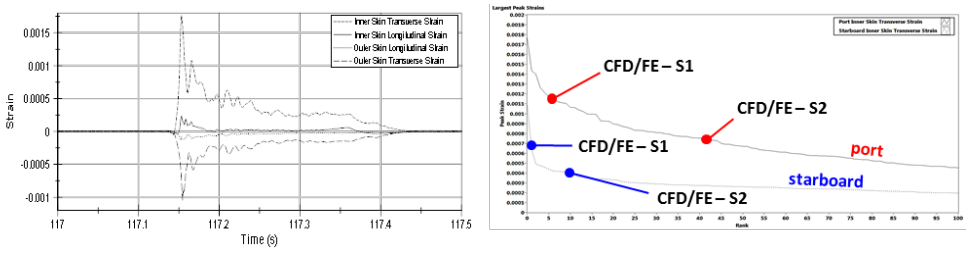


Figure 7: Typical slamming event strain waveforms and maximum strains for the 100 most severe slamming events (note that regular wave CFD/FE gives one value per simulation).

Table 2: Average of highest 1/3 peak strains in port and starboard bay 4 panels for full test duration and S1, S2 wave segments

	Port Transverse Strain	**Starboard Transverse Strain**
Full Test Duration	8.04×10^{-4}	3.11×10^{-4}
Head Wave Segment S1	8.89×10^{-4}	2.70×10^{-4}
Following Wave Segment S2	8.40×10^{-4}	3.83×10^{-4}

6 HYDRODYNAMIC ANALYSIS

The calm water simulations are conducted for both a captive and a free running model. The captive simulations are conducted for a wide range of Fr for the bare hull model free to heave and pitch. The free running simulations are conducted at Fr=1.1, 1.9 and 2.7 for the model appended with sterndrive and body force propeller. The free running model has 6DOF.

Figure 8 shows the comparison of steady state values for both captive and free running simulations, compared with the experimental data. Heave and pitch motions are slightly larger for free running simulation, but the trends versus Fr are similar for both captive and free running simulations. The maximum pitch is for Fr=1.1 (3.6 and 4.2 deg for captive and free running simulations, respectively). Compared to the available experimental pitch data, the comparison errors E=(D-S)%D (D and S are the experimental and simulation values, respectively) for captive and free running simulations are E=6.4 and -8.9%D, respectively. Roll motion is only predicted for the free running model. The roll angle increases by Fr and it is about 2.5 deg for the highest speed, very close to the available experimental value at Fr=2.7 (E=0.8%D). The propeller RPS shows the same trend as EFD, however, it is over predicted for high Fr (E=-12%D). Since the experimental resistance could not be measured for the full

scale Numerette, it is estimated from the propeller input power computed using both the engine curve and propeller open water torque curve. The estimated experimental resistance based on engine curve shows similar resistance for low and high Fr, while open water curve estimates very small resistance at high speeds. The captive CFD simulations show E=58%, 49%, and 13% for Fr=1.1, 1.9, and 2.7, respectively. Corresponding errors for free running simulation are 49%, 36%, and -9%. The study of the free running results show that the pressure resistance of the sterndrive is comparable with the resistance of the bare hull. Therefore, captive simulations for the bare hull geometry with no sterndrive under predicted the resistance significantly. Figure 8 also shows the comparison of CFD results with 1D simulation results, provided by Mercury Marine. The propeller RPS, sterndrive resistance and total resistance show fairly good agreement with CFD free running simulations. However, the 1D simulation predicts larger trim angle as no model was used for the stepped bottom of the boat. CFD free surface and pressure distribution on the hull are shown in Figure 9 for Fr = 2.1 and 2.7, i.e. same as used later for regular wave simulations (S1 and S2, respectively).

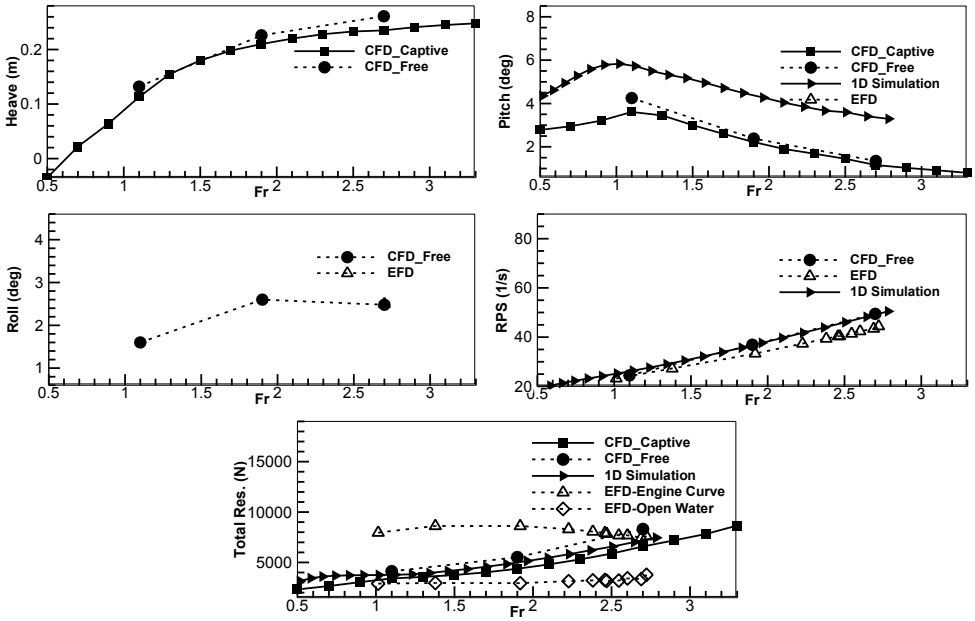


Figure 8: Comparison of CFD and EFD results in calm water

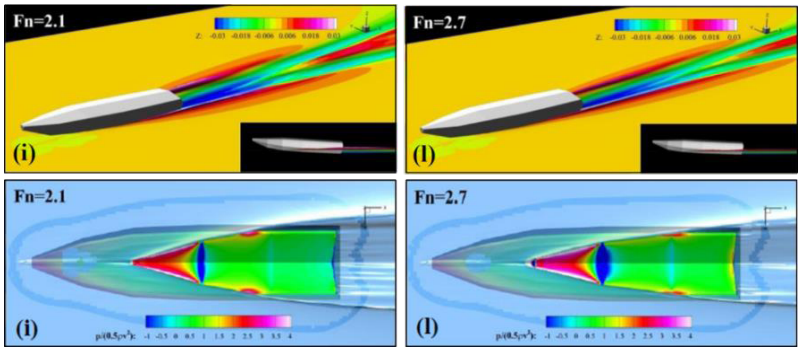


Figure 9: Free surface and pressure distribution for calm water simulation at Fr = 2.1 and 2.7

Figure 10 shows comparison of the captive simulations with the results for other ship hulls including USNA, USCG and Fridsma model. All geometries show similar non-dimensional resistance at high speed. The non-dimensional heave motion is smaller for Numerette, but it follows the same trend as for other geometries. The largest heave motion is for Fridsma, which has the shortest length among all geometries (L=4.5 ft). For pitch motion, Numerette and USNA show similar values for high speed and both show smaller values compared to those for USCG and Fridsma. The trends for pitch motion are the same for all geometries, i.e. there is a peak for pitch motion around Fv=$V/\sqrt{\Delta^{1/3}g}$=2.5-3.0.

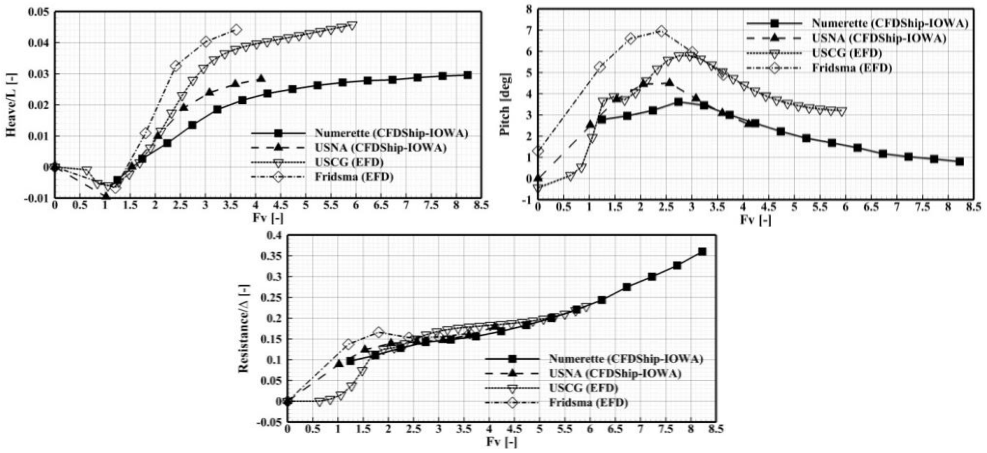


Figure 10: Comparison of motions and resistance against USNA (CFD), USCG (EFD) and Fridsma (EFD)

The speed values used for S1 and S2 are based on Figure 11, which is a close-up of Figure 5, including running mean and RMS. The corresponding inputs for S1 are: Fr equal to 2.15, wave height equal to 0.587 m, with a frequency of the encounter equal to 0.9625 Hz (corresponding to a wave frequency equal to 0.2380 Hz, λ/L equal to 3.120 and H/λ = 1/47). For S2, Fr equals 2.77, the wave height is 0.587 m, and the (negative) frequency of the encounter is 0.9331 (corresponding to a wave frequency equal to 0.2700 Hz, λ/L equal to 2.424 and H/λ = 1/37).

The resistance coefficient, heave and pitch motions indicate satisfactory convergence, as shown in Figure 12 and Figure 13, for S1 and S2 respectively. The response is highly nonlinear. Figure 14 shows the slamming pressures for the keel and panel center, as shown in Figure 2. Note that the experimental strains are measured at the panel center. The slamming pressures for S1 panel center and S2 keel are similar to those described earlier. For S1 keel, the emerging phase has two peaks, which requires more study. For S2, the panel center is not wetted. The re-entering peak is in correspondence with minimum pitch, whereas the emerging peak(s) occurs when the heave starts rising, as shown in Figure 15. The re-entering peaks on the keel for S1 and S2 are comparable in magnitude, and close to 140 kPa.

The results of current regular wave simulations are compared to other planing hulls. Specifically, the 1st harmonic amplitude of heave and pitch motions is compared to Fridsma and USNA models in Figure 16. For Fridsma, x_3/A increases gradually by decreasing the encounter frequency and reaches 1 for long waves. The peak is found near a wavelength corresponding to its resonance condition. Fridsma results also show that x_5/Ak increases with decreasing encounter frequency and reaches nearly 1 for long waves, presenting a small peak. The non-dimensional heave and pitch resonance frequencies are nearly 0.6 for Fridsma. Both

USNA and Numerette geometries have data for limited encounter frequencies and thus the peaks of x_3/A and x_5/Ak are not identified. USNA show similar trend i.e. small motions at short wavelength, increasing to $x_3/A=1$ and $x_5/Ak=1$ at long wavelength. However, Numerette data shows small values at long wavelength, which requires more study.

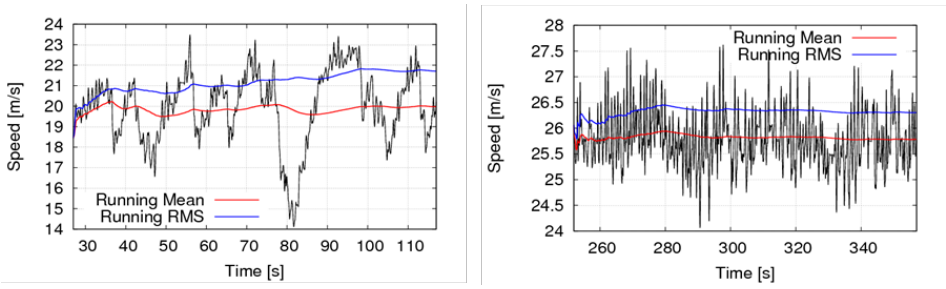


Figure 11: S1 (left) and S2 (right) speed from experiments including running mean and RMS

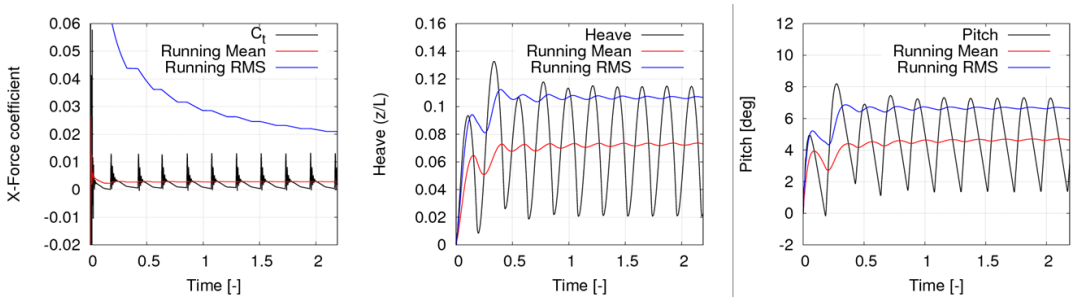


Figure 12: CFD-predicted forces and motions for S1 (time scale=tU/L)

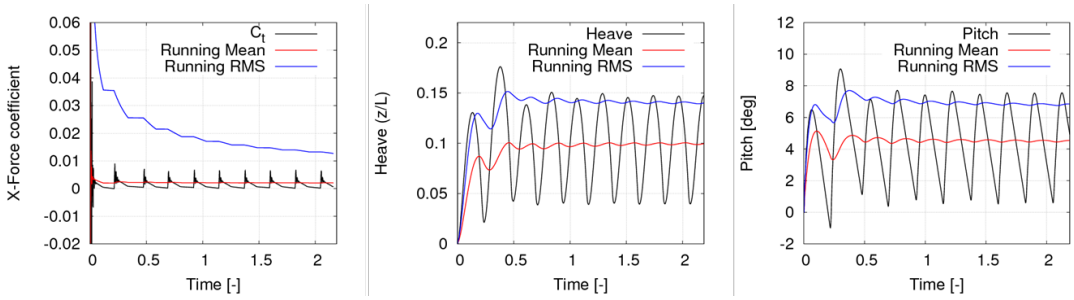


Figure 13: CFD-predicted forces and motions for S2 (time scale=tU/L)

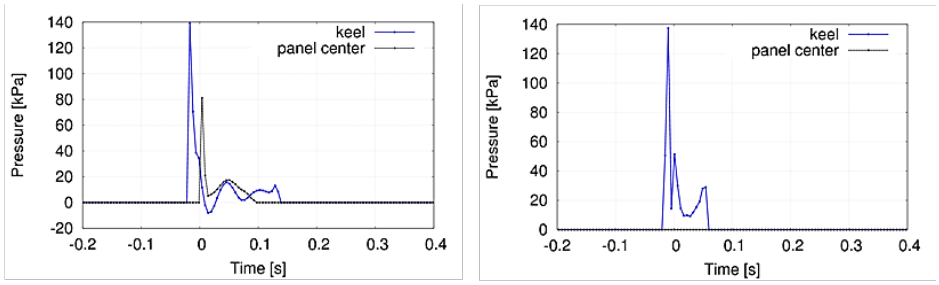


Figure 14: CFD pressure probes location and pressure history for S1 (left) and S2 (right) (time scale=tU/L)

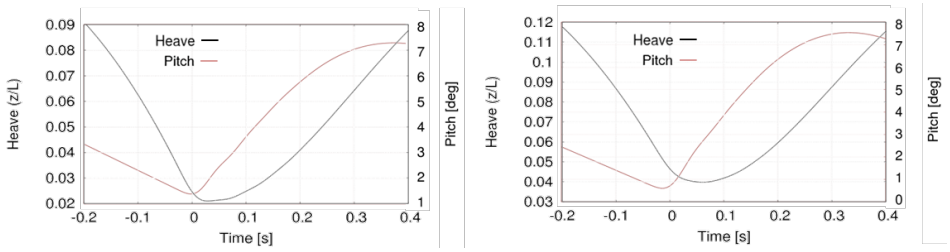


Figure 15: Heave and pitch associated with S1 (left) and S2 (right) simulations

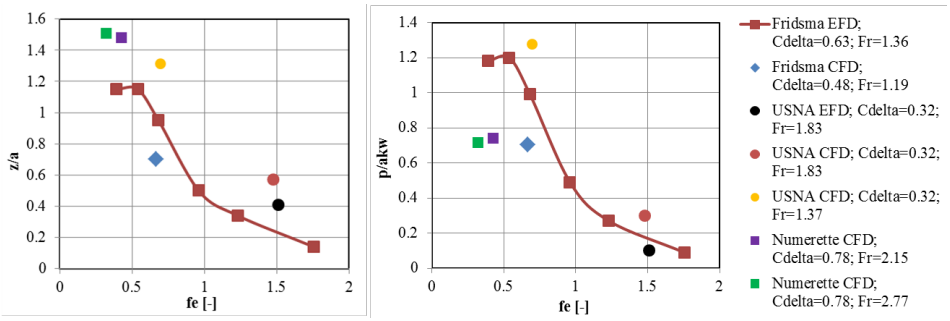


Figure 16: Heave and pitch motions for different geometries

7. FSI ANALYSIS

The CFD/FE slamming strains are validated for S1 and S2, following the approach used for validation of slamming pressures in [5]. Slamming strain events are aligned in time by their re-entering peaks, which provide the expected value EV and standard deviation SD for the peak value and duration, and mean slamming strain. The event duration is defined by the re-entering peak and the signal drop below a given threshold (5% of the peak value). Only strains exceeding the 30% of the highest peak are considered, which is reasonable to detect actual slams from the strain signal. The inner skin transverse strains at the panel center are used for validation, since more severe. The strains are very irregular, however the mean strain is smooth and has a trend similar to typical slamming pressures, as shown in Figure 17.

The CFD/FE slamming strains are also shown in Figure 17, which shows a reasonable agreement with the experiments, especially in consideration of the simulation input approximation to the experimental conditions. The trend of port versus starboard panel strains is well captured by CFD/FE. For S2, peak value and duration are validated since the comparison error E%D is less than the SD%D, as shown in Table 3. The duration is found 0.14 times the encounter period. For S1, the CFD/FE shape is similar to the keel slamming pressure in showing a large re-entering peak and two emerging peaks. The peak E%D is about 3 times larger than SD%D, whereas the duration E%D is less than the SD%D, as shown in Table 3. The duration is found 0.08 times the encounter period.

The comparison of CFD/FE peaks to 1/3 highest experimental peaks average is also provided in Table 3. The average error for the 1/3 highest peaks average is larger for S1 than S2. Overall, the average error is slightly smaller for the 1/3 highest peaks average than that for EV using 30% maximum peak threshold, nevertheless the trend is similar. The identification of the best statistical approach for validation requires irregular wave free-running CFD/FE simulations consistent with the actual experimental conditions. Moreover, the trend of S1 versus S2 depends on head versus following waves, which may be affected by surge and

propeller controller, not considered herein (the CFD/FE model is towed at constant speed). S1 and S2 CFD/FE simulation peaks are included in Figure 7, for comparison with the 100 most severe slams from experiments.

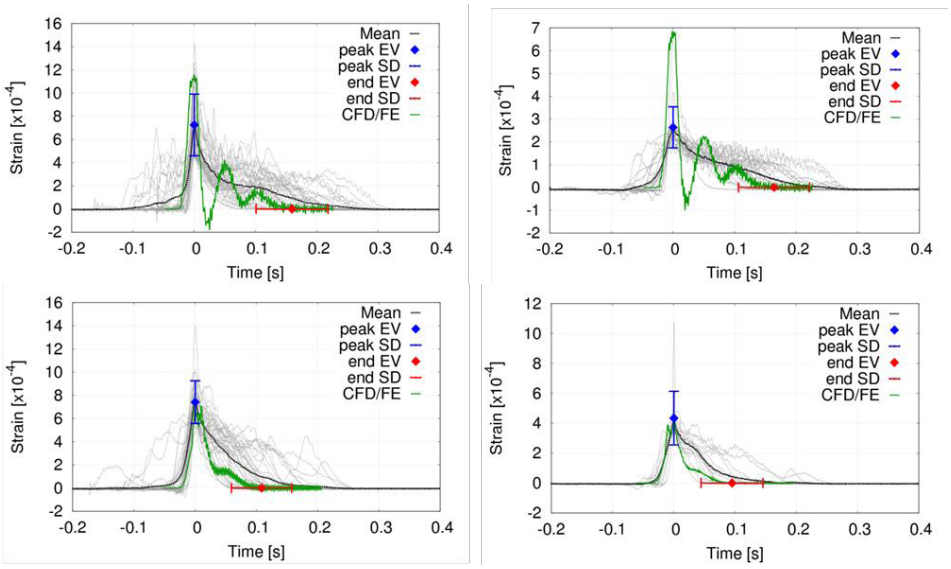


Figure 17: Inner skin transverse strains for port (left) and starboard (right), and S1 (top) and S2 (bottom)

Table 3: Comparison of CFD/FE and EFD results for inner skin transverse strains at panel center

		S1					S2				
		S	EV (D)	E%D		SD%EV (D)	S	EV (D)	E%D		SD%EV (D)
				EV	Ave. 1/3 hst				EV	Ave. 1/3 hst	
Port	Peak	1.19E-03	7.25E-04	-64.3	-34.0	38.0	7.46E-04	7.36E-04	-1.39	11.2	25.5
	Duration	0.15	0.16	6.25		36.8	0.09	0.11	18.2		44.7
Starboard	Peak	6.96E-04	2.56E-04	-172.	-157.	36.0	4.44E-04	4.34E-04	-2.38	-16.0	42.7
	Duration	0.14	0.16	12.5		36.1	0.09	0.10	5.26		52.9
Average absolute	Peak	9.44E-04	4.91E-04	118.	95.9	37.03	5.95E-04	5.85E-04	1.89	13.6	34.15
	Duration	0.15	0.16	9.38		36.48	0.09	0.11	11.72		48.80

7 CONCLUSIONS AND FUTURE RESEARCH

Hydrodynamic slamming on the bottom of a high-speed planing hull was studied experimentally and numerically (CFD). A highly instrumented 9 meter long hull developed for slamming research was used. The bottom of this craft consists of ten separate carbon and glass fiber skin / foam core sandwich panels, each with its unique set of material combinations and fiber layup angles. This allows for study of the influence of bottom stiffness on slamming pressures and deformations. The code CFDShip-Iowa was employed for CFD simulations and the commercial FE code ANSYS was used as structural solver. The hydrodynamic simulations included captive (2DOF without sterndrive) and 6DOF free running conditions for various Froude numbers in calm water and regular waves. Resistance, heave, pitch and roll motions correlated well between experimental operation and numerical simulations for calm water. In offshore sea trials operating in head (S1) and following (S2) waves, strains in two different bottom panels were measured experimentally as well as calculated numerically using one-way coupling (pressures from CFD of rigid hull, applied on dynamic FE model). A few simple parameters such as average peak strains in starboard and

port panels were compared; initial indications are that the numerical procedure correctly predicts which panel is straining more, although the error may be on the order of 30-50%.

Future research will focus on: grid and time-step verification for slamming pressure and strains; semi-coupled two-phase free running irregular wave hydrodynamics simulations including sterndrive/propeller/controller and superstructure; and trim tab and asymmetric pressure distribution effect on the slamming strains. Experimentally, bottom pressures will be measured with piezoresistive thick film high-speed sensors at over 100 locations and correlated with numerical analyses. The influence of bottom stiffness on slamming pressures will be studied experimentally and numerically; in particular, an attempt will be made to answer questions such as whether a more compliant bottom leads to lower slamming pressures. Two-way fluid-structure interaction analyses are required for studying such effects, which is of top priority.

ACKNOWLEDGEMENTS

Research was sponsored by the Office of Naval Research (ONR) grants No. N00014-13-1-0616 and N00014-13-1-0617 administered by Dr. Roshdy Barsoum. The authors thank Dr. Maysam Mousaviraad for assisting the comparison with other geometries and John Scherer, technical advisor at Mercury Marine, for providing the required information for CFD simulations and conducting 1D simulation.

REFERENCES

[1] Ikeda, C. M., Judge, C. Q. Slamming Impacts of Hydrodynamically–Supported Craft. *ASME 2014 33rd International Conference on Ocean, Offshore and Arctic Engineering*. American Society of Mechanical Engineers (2014)

[2] Lv, J., Grenestedt, J. L. Analytical study of the responses of bottom panels to slamming loads. *Ocean Engineering* (2015) **94**:116-125.

[3] He, W., Diez, M., Zou, Z., Campana, E. F., and Stern, F. URANS study of Delft catamaran total/added resistance, motions and slamming loads in head sea including irregular wave and uncertainty quantification for variable regular wave and geometry. *Ocean Engineering* (2013) **74**:189-217

[4] Mousaviraad, S. M., Wang, Z. and Stern, F. URANS Studies of Hydrodynamic Performance & Slamming Loads On High–Speed Planing Hulls in Calm Water & Waves For Deep & Shallow Conditions. *Proceedings of the 3rd International Conference on Ship Manoeuvring in Shallow & Confined Water: Ship Behaviour in Locks*, 3–5 June (2013) Ghent, Belgium.

[5] Fu, T.C., Brucker, K.A., Mousaviraad, S.M., Ikeda, C.M., Lee, E.J., O'Shea T.T., Wang, Z., Stern, F., and Judge C.Q. An Assessment of Computational Fluid Dynamics Predictions of the Hydrodynamics of High-Speed Planing Craft in Calm Water and Waves. *30th Symposium on Naval Hydrodynamics,* Hobart, Tasmania, Australia, 2−7 November (2014)

[6] Huang, J., Carrica, P., Stern, F. Semi-coupled air/water immersed boundary approach for curvilinear dynamic overset grids with application to ship hydrodynamics. *International Journal Numerical Methods Fluids* (2008) **58**(6):591-624.

[7] Thodal R, Grenestedt J, Lv J, Experimental Study of Structural Responses in High Speed Craft subjected to Atlantic Slamming Loads, *In Preparation*

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis(Eds)

ON THE TERM AND CONCEPTS OF NUMERICAL MODEL VALIDATION IN GEOSCIENTIFIC APPLICATIONS

Uwe-Jens Görke*, Thomas Nagel* and Olaf Kolditz*,†

* Department of Environmental Informatics
Helmholtz Centre for Environmental Research - UFZ
Permoserstr. 15, D-04318 Leipzig, Germany
e-mail: uwe-jens.goerke@ufz.de - Web page: http://www.ufz.de

† University of Technology Dresden, Germany

Key words: Coupled Problems, Model Development, Numerical Methods, Model Confirmation, Inverse Problems

Abstract. Modeling and numerical simulation of the coupled physical and chemical processes observed in the subsurface are the only options for long-term analyses of complex geological systems. This contribution discusses some more general aspects of the (dynamic) process modeling for geoscientific applications including reflections about the slightly different understanding of the terms *model* and *model validation* in different scientific communities, and about the term and methods of *model calibration* in the geoscientifc context. Starting from the analysis of observations of a certain part of the perceived reality, the process of model development comprises the establishment of the physical model characterizing relevant processes in a problem-oriented manner, and subsequently the mathematical and numerical models. Considering the steps of idealization and approximation in the course of model development, Oreskes et al. [1] state that process and numerical models can neither be verified nor validated in general. Rather the adequacy of models with specific assumptions and parameterizations made during model set-up can be confirmed. If the adequacy of process models with observations can be confirmed using lab as well as field tests and process monitoring, the adequacy of numerical models can be confirmed using numerical benchmarking and code comparison. Model parameters are intrinsic elements of process and numerical models, in particular constitutive parameters. As they are often not directly measurable, they have to be established by solving inverse problems based on an optimal numerical adaptation of observation results. In addition, numerical uncertainty analyses should be an obligatory part of numerical studies for critical real world applications.

1 INTRODUCTION

Modeling and numerical simulation of the coupled physical and chemical processes observed in the subsurface are the only options for long-term analyses of complex geological systems. The numerical simulation of geotechnological processes (*dynamic modeling*) requires the existence of models that describe the considered problems as adequately as possible. Within this context, it has to be considered that models represent approximations and idealizations of the considered part of the perceived reality, and will map the real world by approximation. During the process of model development, certain relevant aspects will be covered sufficiently accurate, whereas other details, which can be considered as irrelevant for the specific problem definition, can be neglected. In this approximate sense, models are qualified to enable predictions of the behavior of real-world processes, which are not (yet) open for efficient empirical measuring procedures due to exceptional long time scales to be considered (e.g., within the context of geological waste deposition) or due to the general difficult access to local *in situ* measurements in the subsurface. However, absolute exact predictions can never be expected based on process modeling. The model quality, i.e., the degree of conformance of models with the part of the perceived reality they describe, and thus the reliability of model predictions, depends on many factors. Within this context, the quality as well as the spatial and time density of measured data characterizing the observation area plays a crucial role. These data are necessary for parameterization (i.e., calibration; parameter identification; inverse modeling) and confirmation of the developed models.

2 MODELS WITHIN A GEOSCIENTIFIC CONTEXT – AN ATTEMPT OF DISAMBIGUATION

Discussing about *models* in a geoscientifc context, different scientific communities (e.g., geologists; geophysicists; experts in continuum mechanics; engineers) sometimes use this term based on slightly different associations. Thus, in the following we attempt to provide clear definitions of the term *model*:

- **Spatial and/or structural models** usually will be summarized using the term *(static) reservoir model*, and comprise all empirically ascertainable information regarding geometry, geological structure, basic material characteristics (e.g., porosity, intrinsic permeability, thermal and mechanical parameters), and reservoir conditions (e.g., pressure, temperature) of the considered area.

- **Process models** represent *mathematical models* based on physically founded assumptions about coupled physical, geochemical, geoelectrical, and (micro-)biological processes observed in real-world applications. Basically, process models include balance laws and constitutive relations. More detailed reflections regarding process modeling will be discussed in Sec. 3 of this contribution.

- **Numerical models:** The field equations based on balance laws including relevant constitutive relations can be solved analytically just in exceptional cases. Corresponding specific problems frequently will be defined in terms of benchmarks for model and software confirmation. Usually, the solution of field equations describing a real-world problem requires the use of numerical methods. For that, the local formulation of balance laws in terms of differential equations has to be transformed into a global integral formulation. In general, the resulting system of equations is a nonlinear one and contains time derivatives of the primary variables to be calculated. As numerical methods do not provide spatially and temporally continuous solutions, but rather solutions in discrete points of the given solution space, the system of field equations including constitutive relations has to be discretized in space and time within the context of incremental-iterative approximation procedures. In addition, nonlinear systems have to be linearized. The resulting so-called *initial-boundary value problem* can be solved numerically in discrete locations at discrete points of time. As spatial discretization approaches, the methods of finite differences (FDM), finite volumes (FVM), and finite elements (FEM) are widely excepted means of choice. Within this context, the term *numerical model* characterizes either
  - numerical methods and algorithms necessary for computer-based simulations of the process model, or
  - the entire data set necessary for the numerical simulation of a specific problem (parameterized and spatially discretized structural model; boundary and initial conditions; parameters controlling the simulation procedure).

3 PROCESS MODELING

The process modeling of physical, geochemical, and (micro-)biological processes in the subsurface is based on mathematical theories enabling the description of spatially and temporally evolving processes of the perceived reality in terms of differential equations or systems of differential equations. Developing the relevant mathematical apparatus, we make use of physically, chemically, and biologically founded assumptions.

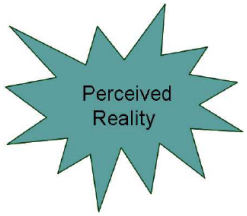


Figure 1: Origin of the development of process models

Thus, based on reasonable abstractions, the process model to be defined can be formulated to be sufficiently complex in order to cover all relevant aspects of the specific

real-world application but at the same time to be sufficiently simple in order to enable its numerical solution. Observations of the interesting, problem-specific part of the perceived reality serve as the nucleus of the development of process models (cf. Fig. 1). A first step in this procedure is the establishment of physical models adequately mapping the perceived reality based on decisions to neglect those processes that are not relevant for the considered application (cf. Fig. 2). This kind of idealization is problem-dependent, and thus, different problems related to the same observation area may result in different physical models (e.g., if a shallow subsurface area is designated for groundwater remediation activities or for use as geothermal reservoir).

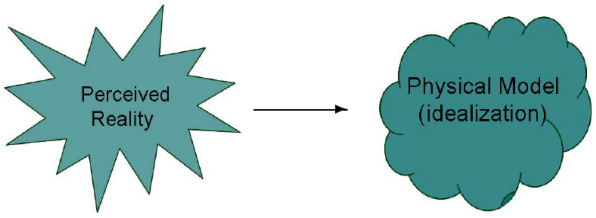


Figure 2: Generation of a physical model

The second step in the definition of process models is the development of a mathematical framework describing the physical model in an approximate manner, and consisting of balance laws and constitutive relations in terms of algebraic, differential or integral equations (cf. Fig. 3).

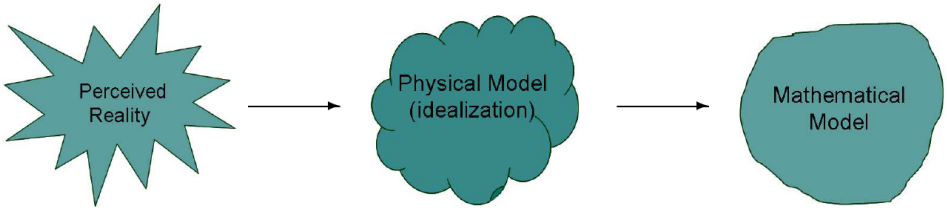


Figure 3: Generation of a mathematical model

As a matter of principle, processes of the nature can be characterized using balance laws related to fundamental state variables. Those are primarily the balances of mass, linear momentum, moment of momentum, and energy. Balance laws are independent from the real materials under consideration, and can be formulated locally in terms of problem-oriented (systems of) differential equations based on convenient assumptions. Within this context, they are valid for each location of the observation area, and are usually called *field equations*.

Observations of the nature show that certain physical variables may evolve in different manner dependent on the specific materials involved in the considered application (e.g., various pictures of deformation of a solid body of the same geometry and subjected to the same external loading, but consisting of different materials). This causes the necessity of the definition of additional equations, the so-called *constitutive relations* characterizing material-dependent dependencies between state variables. Equations of state for fluids, stress-strain relations for solid bodies, and kinetic (reaction) models for chemical reactions are typical examples for constitutive relations. For mathematical reasons, the definition of constitutive relations is necessary in order to define a closed system of equations including balance laws and constitutive relations. The pure system of field equations resulting exclusively from the balance laws contains more unknowns as equations, and can thus not be solved in a unique manner.

4 MODEL CONFIRMATION

The procedure of the development of process models illustrated in Sec. 3 shows that models are always idealizations and approximations of a considered part of the perceived reality reflecting specific assumptions and parameterizations made for the model set-up. In order to assure the usability of a process model and the corresponding numerical model, it has to be shown that it adequately reflects real-world processes. In the literature, this process is called *model validation* or even *model verification* – but are these terms justified in their absolute, literal meaning? According to Oreskes et al. [1]

- **Model verification** "... means an assertion or establishment of truth." and "To say that a model is verified is to say that its truth has been demonstrated, which implies its reliability as a basis for decision-making."

- **Model validation** "... does not necessarily denote an establishment of truth... Rather, it denotes the establishment of legitimacy..." and "... a model that does not contain known or detectable flaws and is internally consistent can be said to be valid."

Considering the steps of idealization and approximation in the course of model development, and the open character of models, Oreskes et al. [1] state that process and numerical models can neither be verified (establishment of the truth of the model) because

- "... models... are never closed systems" (but truth can be demonstrated only for closed systems),

- models "... require input parameters that are incompletely known",

- data sets for the parameterization (i.e., calibration) of process models are always incomplete (e.g., due to the availability of only a few locally determined parameters from core tests considering a model of a whole geological reservoir), and

- often used phenomenological constitutive theories are characterized by loss of information on real physical scale (e.g., use of macroscale models not being based on real microscale behavior),

nor validated (establishing the legitimacy of a model) in general because

- model results depend on assumptions required for model development and on parameterization, and
- models valid for one mapping of the reality may not be valid for another one (e.g., material behavior dependent on temperature).

If the terms model verification and model validation are unfavorable in their literal meaning, how else the usability of process models can be assured? For this purpose, Oreskes et al. [1] introduce the term *model calibration* stating: "...science requires that empirical observations be framed as deductive consequences of a general theory... If these observations can be shown to be true, the theory... is *confirmed*...", and "...confirming observations do not demonstrate the veracity of a model..., they only support its probability." Thus, rather the adequacy of models with specific assumptions and parameterizations made during model set-up can be confirmed, not their (general) truth. Within this context, model parameterization is performed using well-established approaches of model calibration. If the adequacy of process models with observations can be confirmed using lab as well as field tests and process monitoring, the adequacy of numerical models can be confirmed using numerical benchmarking (e.g., providing analytical solutions) and code comparison for more complex systems (see Class et al. [2], Kolditz et al. [3], and others).

Figure 4: Direct modeling: simulation of real-world problems using parameterized process and numerical models

5 MODEL CALIBRATION AS AN OPTIMIZATION PROBLEM

Model parameters, in particular constitutive parameters, are constituent, intrinsic elements of process and numerical models. Parameterized models are used for simulations in order to analyze and/or predict the evolution of various processes in real-world applications, thus, solving a *direct problem* (cf. Fig. 4). Within this context, model parameters

can be considered as *reason* for the specific evolution of relevant primary variables (e.g., fluid pressure, deformation of solids, temperature(s), concentration(s)).

As model parameters are often not directly measurable, their identification is based on the analysis of their *effects* onto measurable field variables. This process of model calibration (i.e., model parameterization) requires the solution of an *inverse problem* based on an optimal numerical adaptation of observation results (cf. Fig. 5)

Due to the lack of information about the perceived reality (i.e., observation results are available only for discrete points of the field problem to be solved for real-world applications), the process of identification of model parameters is an ill-posed problem being source of the non-uniqueness of the solution, and potentially suffering from instabilities of the solution.

Figure 5: Inverse modeling: optimal numerical adaptation of observations from lab and field experiments, and from field exploration

As the operator mapping model parameters to the considered primary field variables is usually of implicit, non-linear character with unknown mathematical structure, its explicit closed-form inversion approves to be impossible in this case. Consequently, the ill-posed inverse problem of model calibration in general results in the solution of a non-linear optimization problem: The parameters have to be estimated in such a way that an appropriately defined objective function approaches its minimum (for an overview see Bruhns and Andig [4], and others).

5.1 Objective Function

A model function will be defined, which characterizes an arbitrary field variable y depending on a vector of variables $\mathbf{x}$ as well as on a set (vector) of model parameters $\mathbf{c}$

$$y \;=\; y(\mathbf{x}, \mathbf{c})\,. \tag{1}$$

In order to analyze a deformation problem, the corresponding model function can be, for instance, constituted by the displacement field depending on stresses, internal variables and material parameters.

The calibration of constitutive models is aimed at the determination of parameters realizing a sufficiently accurate approximation of measured discrete data $\hat{y}_i$ representing defined conditions for variables $\mathbf{x}_i$. The corresponding parameter set, which is in this

narrower sense an optimal one, and thus the best approximation of measured data, can be considered as determined if a least squares norm approaches its minimum

$$\frac{1}{2}\sum_{i=1}^{n}\left[\hat{y}_i - y(\mathbf{x}_i,\mathbf{c})\right]^2 \quad \longrightarrow \quad \min\,. \tag{2}$$

Based on the definition of a vector of residuals $\mathbf{r}$ between measured and calculated values

$$\mathbf{r}(\mathbf{c}) = \left\{r_i(\mathbf{c})\right\} \qquad \text{mit} \qquad r_i(\mathbf{c}) = \hat{y}_i - y(\mathbf{x}_i,\mathbf{c}) \tag{3}$$

the objective function Φ can be formulated as the following least squares norm:

$$\Phi(\mathbf{c}) = \frac{1}{2}\,\mathbf{r}^{\mathrm{T}}(\mathbf{c})\,\mathbf{r}(\mathbf{c}) = \frac{1}{2}\sum_{i=1}^{n} r_i^2(\mathbf{c}) \quad \longrightarrow \quad \min\,. \tag{4}$$

The necessary optimality criterion

$$\nabla\Phi(\mathbf{c}^*) = \mathbf{0} \tag{5}$$

results for models representing linear functions of the parameters $\mathbf{c}$ in a system of linear algebraic equations with the coordinates of the optimal parameter vector $\mathbf{c}^*$ as primary variables. This system is known as *normal equation*

$$\widetilde{\mathbf{J}}^{\mathrm{T}}\,\widetilde{\mathbf{J}}\,\mathbf{c}^* = \widetilde{\mathbf{J}}^{\mathrm{T}}\mathbf{r}\,. \tag{6}$$

Here, $\widetilde{\mathbf{J}}$ represents the Jacobian matrix comprising the first derivatives of the residuals with respect to the material parameters.

5.2 Optimization Procedure

If the identification of model parameters is based on the analysis of homogeneous fields of primary variables, only the initial value problem has to be solved to obtain the value of the objective function to be minimized (e.g., using results from uni- or triaxial compression tests). Apart from deterministic (gradient-based) optimization methods in this case special applications of stochastic and/or evolutional (gradientless) approaches are reported in the literature even for quite complicated material models (see e.g. Harth et al. [5]).

Analyzing experiments on samples with an inhomogeneous distribution of primary variables the amount of information for a reliable identification process is distinctly higher. In this case, the objective function has to be calculated solving a complex initial-boundary value problem, for what usually FEM approaches are used. Thus, due to the time-consuming solution of the direct problem stochastic methods are not suitable for the analysis of inhomogeneous physical, geochemical, and/or (micro-)biological states but deterministic approaches are well-tried for this case in engineering sciences. Regarding

their mathematical basics we exemplarily refer to the monographs of Beck and Arnold [6], Dennis and Schnabel [7], Nocedal and Wright [8] or Rao [9]. A crucial step to be performed using deterministic optimization approaches is the so-called *sensitivity analysis*, i.e., the calculation of the Jacobian matrix of first derivatives of the objective function (cf. Eq. 6). Common methods to perform the sensitivity analysis are analytical calculations if exclusively explicit functional dependencies occur, pure numerical approaches using finite differences, or the direct differentiation (semi-analytical methods).

Considering the modeling of geoscientific real-world problems, the so-called *history matching* is well-established for the parameterization of complete static site models. Within the context, model parameters will be adapted performing simulations of the considered problem with known process parameters (e.g., measured borehole pressures and/or temperatures).

6 CONCLUSIONS

The above illustrated procedure of the development of process and numerical models makes clear that at each stage of model definition certain approximations of the real world are implemented. Thus, modeling and numerical simulation can provide a preview of trends of the behavior of the studied systems in accordance to the defined assumptions and preconditions, but no closed, accurate, deterministic predictions can be expected (modeling is not reality). Consequently, formulations in regulatory documents establishing rules for the implementation of (geo-)technologies have to reflect this character that is inherent to modeling: it is not reasonable to claim exact long-term guarantees, but it is rather useful to claim a statistically founded estimation of trends. Nevertheless, it is out of question that modeling and numerical simulation is the only option for long-term analyses of complex systems, and for providing prognoses in the range of reliable probabilities, if empirical data are difficult to access in lab and/or field experiments. Climate simulations, weather forecast, and geoscientific applications as considered here are exemplary for that. Within this context, it is of vital importance to know what can be expected from modeling, how results of numerical simulations have to be interpreted, and to keep limitations in mind.

The following aspects are of great importance with respect to reliable process simulations:

- Definition of the processes and subprocesses that are relevant for the specific problem to be solved in order to establish a system of field equations, which is complex enough to cover the relevant system behavior but simple enough to ensure an efficient and stable solution process.
- Formulation of problem-dependent specific expressions of the balance laws and the constitutive relations according to the latest state of research.
- Model confirmation using standardized benchmarks and code comparison.

- Extensive provision of data from lab and field tests for an improved process understanding and for the model calibration.
- If measured data are not available for all local points of the spatially discretized observation area (this is the usual case for geoscientific real-world applications), a stochastically based parameter space has to be determined based on a statistically sufficient number of different numerical realizations.
- An uncertainty analysis is advised based on this parameter space, and results of the process simulation will be provided with certain likelihood.

These aspects have to be reconsidered, if the process and numerical models undergo changes in the observation area (e.g., enlargement or reduction in one of the geometric dimensions), and/or in the initial, boundary and/or process conditions, or if the specific problem modifies (e.g., interest in hydro-mechanical coupled processes instead of pure hydrological processes). The motivation behind is that each modeling process includes the definition of problem-oriented input data, and thus, provides specific, problem-oriented results.

As modeling cannot provide closed, deterministic predictions, but rather a preview of trends regarding how the studied systems may behave under the defined assumptions and conditions, the solution of inverse problems in geoscientific applications deserves more attention. Within this context, an extensive provision of lab and field test data for an enhanced process understanding combined with a more reliable and documented model calibration are required to improve the adequacy of static and dynamic models. In addition, numerical uncertainty analyses should be an obligatory part of numerical studies for critical real world applications.

REFERENCES

[1] Oreskes, N., Shrader-Frechette, K. and Belitz, K. Verification, Validation, and Confirmation of Numerical Models in the Earth Sciences. *Science* (1994) **263**:641–646.

[2] Class, H., Ebigbo, A., Helmig, R., Dahle, H.K., Nordbotten, J.M., Celia, M.A. et al. A benchmark study on problems related to CO_2 storage in geologic formations. *Comput. Geosci.* (2009) **13**:409–434.

[3] Kolditz, O., Görke, U.-J., Shao, H. and Wang, W. (eds.). *Thermo-hydro-mechanical-chemical processes in porous media: benchmarks and examples*. Lecture Notes in Computational Science and Engineering 86, Springer (2012).

[4] Bruhns, O. and Anding, D.K. On the simultaneous estimation of model parameters used in constitutive laws for inelastic material behaviour. *Int. J. Plast.* (1999) **15**:1311–1340.

[5] Harth, T., Schwan, S., Lehn, J. and Kollmann F.G. Identification of material parameters for inelastic constitutive models: statistical analysis and design of experiments. *Int. J. Plast.* (2004) **20**:1403–1440.

[6] Beck, J.V. and Arnold, K.J. *Parameter estimation in Engineering and Science*. Wiley, New York (1977).

[7] Dennis, J.E. and Schnabel, R.B. *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*. Prentice-Hall Inc., Englewood Cliffs, New Jersey (1983).

[8] Nocedal, J. and Wright, S.J. *Numerical Optimization*. Springer (1999).

[9] Rao, S.S. *Engineering Optimization: Theory and Practice*. 3rd ed., Wiley, New York (1996).

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis (Eds)

A RBF-BASED LOCAL COLLOCATION METHOD FOR MODELLING THERMOMECHANICAL PHENOMENA DURING DC CASTING OF ALUMINIUM BILLETS

BOŠTJAN MAVRIČ[*] AND BOŽIDAR ŠARLER[*,†]

[*]Laboratory for Simulation of Materials and Processes
Institute of Metals and Technology
Lepi pot 11, SI-1000 Ljubljana, Slovenia
e-mail: bostjan.mavric@imt.si

[†] Laboratory for Multiphase Processes
University of Nova Gorica
Vipavska 13, SI-5000 Nova Gorica, Slovenia
e-mail: bozidar.sarler@ung.si

Key words: meshless methods, local radial basis function collocation method, multiquadrics, thermoelasticity, direct-chill casting.

Abstract. In this work, the local radial basis function collocation method is applied to the thermoelasticity with intention to model the low-frequency electromagnetic direct-chill casting process of aluminium billets. The devised thermoelastic model is coupled with the heat transport model for the DC casting process and preliminary results on the stress state are presented. The effect of the casting speed and the application of the electromagnetic field on the principal stresses is presented.

1 INTRODUCTION

The thermomechanical phenomena that occur during DC casting of aluminium billets can have significant impact on the quality of the cast piece. Under specific stress conditions hot tearing and cracking of the cast piece can occur [1, 2]. Unwanted large deformations of the cast piece lead to its scrapping. The deformations also drastically influence the heat transfer efficiency at the contact with the mold, which can lead to unwanted remelting and melt outbreaks.

Modelling of thermomechanical phenomena during this process is not an easy task. In addition to elastic deformation, the strain field also has contributions from viscoplastic creep, plastic deformation, and thermal expansion. All these phenomena occur in nonhomogeneous material with strong dependence of material properties on the temperature. The importance of the stresses during the DC casting process results in a large number of papers dealing with its predictive modelling. The first simple models were developed quite early [3, 4] and consider heat diffusion and thermomechanics by simple constitutive relations. At present, more sophisticated models, involving constitutive relations coupled with heat and fluid flow are being developed [1, 5].

Recently, the performance of the DC casting process is being upgraded by the application

of the oscillating electromagnetic field, with intention of improving the quality of the casting by stirring the melt [6]. To better understand the process, the numerical models are developed in parallel with industrial applications [7, 8].

Many models describing the DC casting process already exist [1, 5, 9, 10] and can provide accurate results. Existing models mainly employ the Finite Element Method (FEM), which may prove inefficient in some circumstances. The local meshless method used in our work has several advantages over FEM [11, 12]. There is no need for expensive polygonization of the domain, since the only information needed are the positions of the points. The computational points can be easily added or removed to achieve optimal accuracy [13] and complex geometries can be easily described since irregular node arrangements can be used. The local radial basis function collocation method (LRBFCM) has already been successfully applied to many physical and engineering problems: heat and fluid flow with [14] and without [15] the influence of magnetic field, solidification [16], continuous casting of steel [17], and modelling of semiconductors [18] .

The thermomechanics model that is presented, is intended to complement the meshless model of mass and heat transfer, which is developed in our group [4]. In this contribution, first the meshless method formulation for solution of boundary value problem is given, followed by some method benchmarks on simple thermoelasticity problems. The thermoelasticity model is described and the preliminary results for stationary state of DC casting are presented.

2 GOVERNING EQUATION

We consider uncoupled formulation of thermoelasticity with stationary thermal profile supplied by the heat and mass transfer model. The stress equilibrium is written in terms of deformation field $\mathbf{u}$ by employing Hooke's law for an isotropic solid. Resulting governing equation is given as

$$G\nabla^2\mathbf{u}+(G+\lambda)\nabla\nabla\cdot\mathbf{u}+\nabla\lambda\nabla\cdot\mathbf{u}+\nabla G\left(\nabla\mathbf{u}+\left(\nabla\mathbf{u}\right)^T\right)=\nabla\left(\beta\left(T-T_{ref}\right)\right)-\mathbf{f}\,. \tag{1}$$

Here G stands for the shear modulus, λ for Lamé parameter and $\mathbf{f}$ for the body force. The coupling with the temperature field is described by the coefficient β defined as $\beta=(3\lambda+2G)\alpha$, where α is the coefficient of linear thermal expansion. T_{ref} stands for the reference temperature at which the thermal expansion is considered to be zero. All the material properties are allowed to vary over the computational domain in a continuous manner.

For the description of the problem setting, the displacement, symmetry, and traction boundary conditions are needed. The deformation of the top part of the billet is restricted by the mould, while the rest of the outer surface is free.

In the computational domain, the material undergoes solidification and therefore significantly changes the elastic properties. The temperature dependence of alloy properties can be obtained from JMatPro database for each alloy considered [19].

3 METHOD FORMULATION

The aim of our work is to develop a method to model the boundary value problem for linear vector partial differential equation (PDE) in two dimensions. The governing equation equation (1) can be written in the form $D_{\xi\zeta}u_{\zeta} = g_{\xi}$, with unknown solution vector u_{μ} and index μ running over the coordinates of the chosen coordinate system. The formulation of the method incorporates the following steps: construction of local influence domains, local interpolation, calculation of the differential operators, and the construction of the system of linear equations representing the governing PDE. These steps are described in the following subsections.

3.1 Node arrangement and influence domain selection

The first step is setting up an appropriate node arrangement and determining the domains of influence for each node. The determination of influence domains is especially important, since the local interpolation of the solution is constructed on them. In general N_{Γ} nodes are put on the boundary Γ and N_{Ω} are distributed over the interior Ω of the computational domain.

The node arrangement used in this work is obtained of by minimizing an energy function which is the sum of Lennard-Jones-like potentials among a certain number of nearby nodes as described in [20]. The procedure results in a node arrangement that is locally similar to the hexagonal grid and is illustrated Figure 1.

The domain of influence for each node is determined by choosing ${}_{l}N$ nearest neighbors of the node with index $l = 1,\ldots,N$, where $N = N_{\Gamma} + N_{\Omega}$ is the number of all discretization nodes. In this step we determine mapping $s_l(i): i \to l'$ from the index $i = 1,...,{}_{l}N$, which enumerates the nodes in the local influence domain, to the global enumeration index l'.

3.2 Interpolant construction

Radial basis function (RBF) interpolant is constructed over each domain of influence. Modifed multiquadrics (MQ), defined as

$$\Phi_j(\mathbf{r}) = \sqrt{\left(\frac{\epsilon_l}{h_l}\right)^2 |\mathbf{r} - \mathbf{r}_j|^2 + 1}\,, \tag{2}$$

augmented by linear and constant monomials are selected. Quantities used in the MQ definition have the following meaning: $\mathbf{r}_j$ is the position of the node in which the MQ is centered, ϵ_l is the MQ shape parameter, which is allowed to be different for different influence domain, and h_l stands for the influence domain size defined as

$$h_l = \sqrt{\frac{\sum_{i=1}^{{}_{l}N} |\mathbf{r}_{s_l(i)} - \mathbf{r}_l|^2}{{}_{l}N - 1}} \tag{3}$$

Formally, the interpolant is given as

$$u_\xi(\mathbf{r}) = \sum_{i=1}^{{}_lN} {}_l\alpha_{i,\xi}\Phi_{s(i)}(\mathbf{r}) + \sum_{i={}_lN+1}^{{}_lN+{}_aN} {}_l\alpha_{i,\xi}p_i(\mathbf{r}-\mathbf{r}_l) = \sum_{i=1}^{{}_lN+{}_aN} {}_l\alpha_{i,\xi}\psi_i(\mathbf{r}) \tag{4}$$

where ${}_aN$ is the number of augmentation monomials used (in present work fixed to 3) and ${}_l\alpha_{i,\xi}$ are the expansion coefficients, which are determined by the collocation. The collocation equations, which are used to determine the coefficients ${}_l\alpha_{i,\xi}$ are modified so that the resulting interpolant satisfies the appropriate boundary conditions, if any of the nodes in the considered influence domain lies on the boundary. The system of equations obtained in this manner can be compactly written in matrix form as

$$\sum_{i,\zeta} {}_lA_{ji,\xi\zeta}\,{}_l\alpha_{i,\zeta} = {}_l\gamma_{j,\xi} \tag{5}$$

with the interpolation matrix ${}_lA$ and the right-hand side vector ${}_l\gamma$ given as

$${}_lA_{ji,\xi\chi} = \begin{cases} \Psi_i(\mathbf{r}_{{}_ls(j)})\delta_{\xi\chi} & \text{if } \mathbf{r}_{{}_ls(j)} \in \Omega \\ B_{\xi\chi}(\mathbf{r}_{{}_ls(j)})\Psi_i(\mathbf{r}_{{}_ls(j)}) & \text{if } \mathbf{r}_{{}_ls(j)} \in \Gamma \\ \mathrm{p}_j(\mathbf{r}_{{}_ls(\mathrm{i})})\delta_{\xi\chi} & \text{if } j > {}_lN \end{cases}, \tag{6}$$

$${}_l\gamma_{j,\xi} = \begin{cases} \mathrm{u}_\xi(\mathbf{r}_{{}_ls(j)}) & \text{if } \mathbf{r}_{{}_ls(j)} \in \Omega \\ b_\chi(\mathbf{r}_{{}_ls(j)}) & \text{if } \mathbf{r}_{{}_ls(j)} \in \Gamma \\ 0 & \text{if } j > {}_lN \end{cases}. \tag{7}$$

In the expressions (6) and (7) we assume that the boundary conditions are linear and specified by appropriate boundary condition operator as $B^j_{\xi\zeta}u_\zeta(\mathbf{r}_j) = b^j_\xi$.

3.3 Discretization of the governing equation

The interpolation of the field values specified in nodal points can be used to estimate the differential operators. Since the expansion coefficients are assumed to be constant, the differential operators act only on the basis functions. This fact can be used to discretize partial differential equations.

By replacing the unknown solution $\mathbf{u}$ with the interpolation given by equation (4), the governing equation at every interior node can be stated in terms of the interpolation coefficients. The interpolation coefficients are further replaced by the components of vector ${}_l\gamma$, thus expressing the governing equation at every interior node by the unknown solution values and given boundary conditions for nodes that belong to the influence domain, centered on the considered node. The resulting governing equation is for each interior node l stated as

$$\sum_{k,\chi} \Upsilon^{\Omega}_{{}_ls(k),\chi} u_{{}_ls(k),\chi} \sum_{i,\zeta} {}_lA^{-1}{}_{ik,\zeta\chi} D_{\xi\zeta}\psi_i(\mathbf{r}_l) = g_{l,\xi} - \sum_{k,\chi} \Upsilon^{\Gamma}_{{}_ls(k),\chi} b_{{}_ls(k),\chi} \sum_{i,\zeta} {}_lA^{-1}{}_{ik,\zeta\chi} D_{\xi\zeta}\psi_i(\mathbf{r}_l) \tag{8}$$

In this expression the boundary and the domain indicators Υ^{Γ}_j and Υ^{Ω}_j are used to achieve an efficient notation. The indicators evaluate to one, if the corresponding point j belongs to the set under consideration and to zero otherwise.

The set of linear equations for the unknown solution values stated in (8) is sparse and can be solved efficiently by specialized solvers. The numerical performance of the method has been investigated in our recent publications [21–23].

4 PRELIMINARY RESULTS ON LFEMC DC CASTING

4.1 Geometry and Boundary Conditions

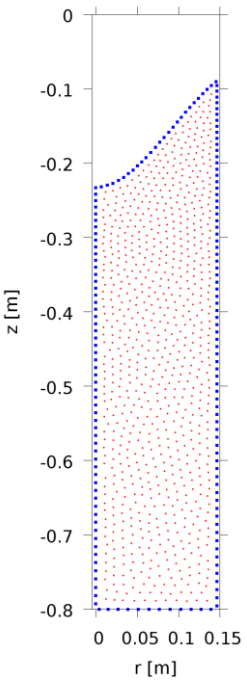


Figure 1 The node arrangement used for the discretization of the stress equilibrium equation. The circles represent interior points and the squares the boundary points.

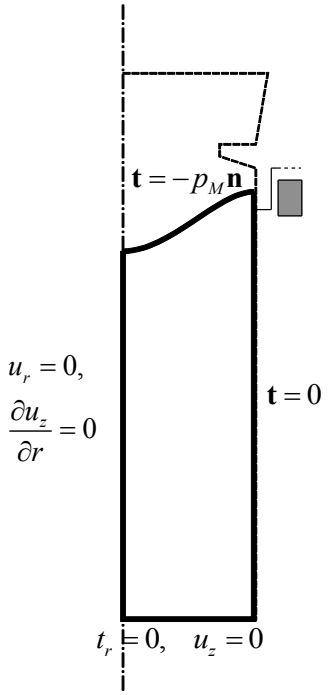


Figure 2 Geometry of the considered DC casting example with scheme of the boundary conditions. The computational domain for the solid mechanics model is denoted by solid line while the computational domain for the heat and mass transfer model additionally includes the part denoted by the dashed line.

In present work we are considering the governing equation (1) applied to the axisymmetric case of the DC casting of round aluminium billets. The temperature field, which has the largest influence on the deformation and stress field during DC casting, is calculated by the heat and fluid flow model. The model itself along with the material properties used is in more detail presented in the accompanying contribution by Šarler et al.. Material properties are obtained from the JMatPro for the alloy AA6082, which is being considered in our case.

Since linear thermoelasticity could not cope with the behavior of fluids, only the part of the billet that is solid is considered for the thermoelasticity model. The top boundary of the computational domain is chosen in such a way that it coincides with the position of the liquidus isotherm. The boundary conditions and the computational domain are shown in Figure 2. On the symmetry axis, the symmetry boundary conditions are assumed. On the bottom the deformation in vertical direction is assumed zero, while zero traction is prescribed in the radial direction. The outer surface is assumed to be free, except for the topmost part, which is constrained by the mould. In this part, the radial deformation is prescribed to be zero, while the zero traction is assumed in the vertical direction. On the top boundary, the

metallostatic pressure of the metal above the liquidus line is applied.

Although in DC casting many process parameters are important, in this work we are considering only the effect of low-frequency electromagnetic stirring (EMS). Some preliminary results regarding the effect of the two parameters are demonstrated in the following two sections. With the reference solution at nominal casting speed 80 mm/min and without the EMS applied is shown in Figure 3.

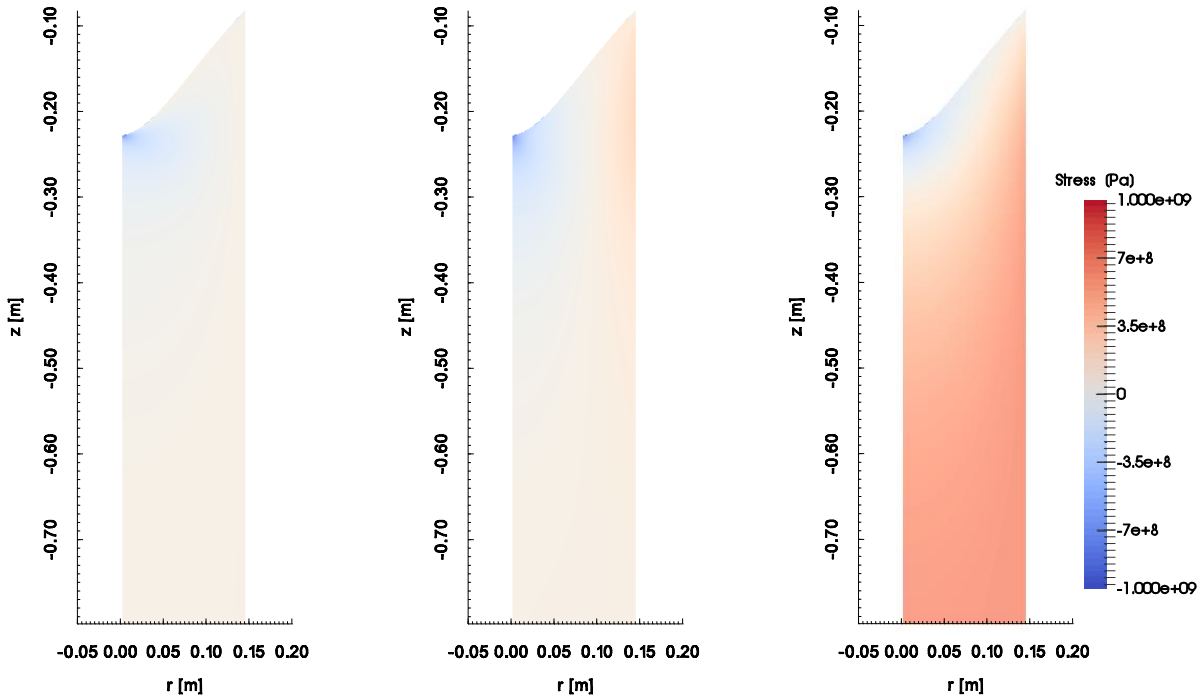


Figure 3 The stress state during DC casting. From left to right the plots show radial, vertical, and circumferential stress.

4.2 Effect of casting speed

Casting speed has important influence on the quality of the cast piece. The effect on the circumferential stress is shown in Figure 4. On all the plots the isoline of zero circumferential stress is shown. We can see that at the nominal casting speed of 80 mm/min the stress isoline touches the liquidus isoline on the surface of the billet. All the solidifying area thus experiences compressive stress. In case the casting speed is either increased or decreased, the zero stress isoline makes contact with the liquidus line in the interior of the billet. The solidifying area on the outside of the point of intersection experiences tensile stress, which increases the possibility of cracking.

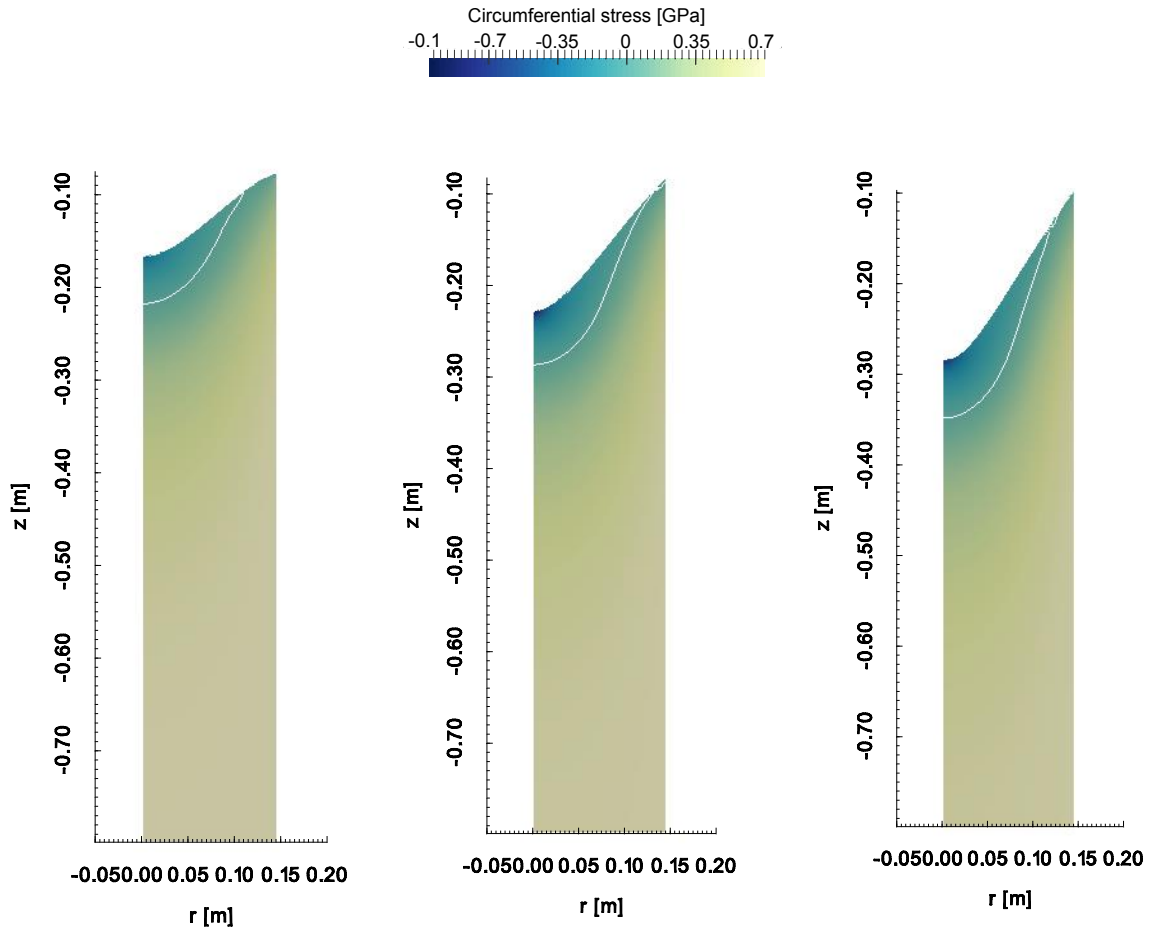


Figure 4 The effect of the casting speed on the circumferential stress. From left to right the plots show circumferential stress at casting speed of 60 mm/min, 80 mm/min and 100 mm/min. On each plot the zero isoline is shown.

4.3 Effect of EMS

In this study the driving current with amplitude 10 A and frequency 20 Hz has been applied to the casting process at the nominal casting speed of 80 mm/min. The difference between the reference stress without EMS and the stress when EMS is applied are shown in Figure 5. We can see that the application of EMS reduces the circumferential stress in the outer region and thus decreases the possibility of cracking. The effect of the EMS is beneficial also to the other two components, since the amplitude of the variations is slightly reduced.

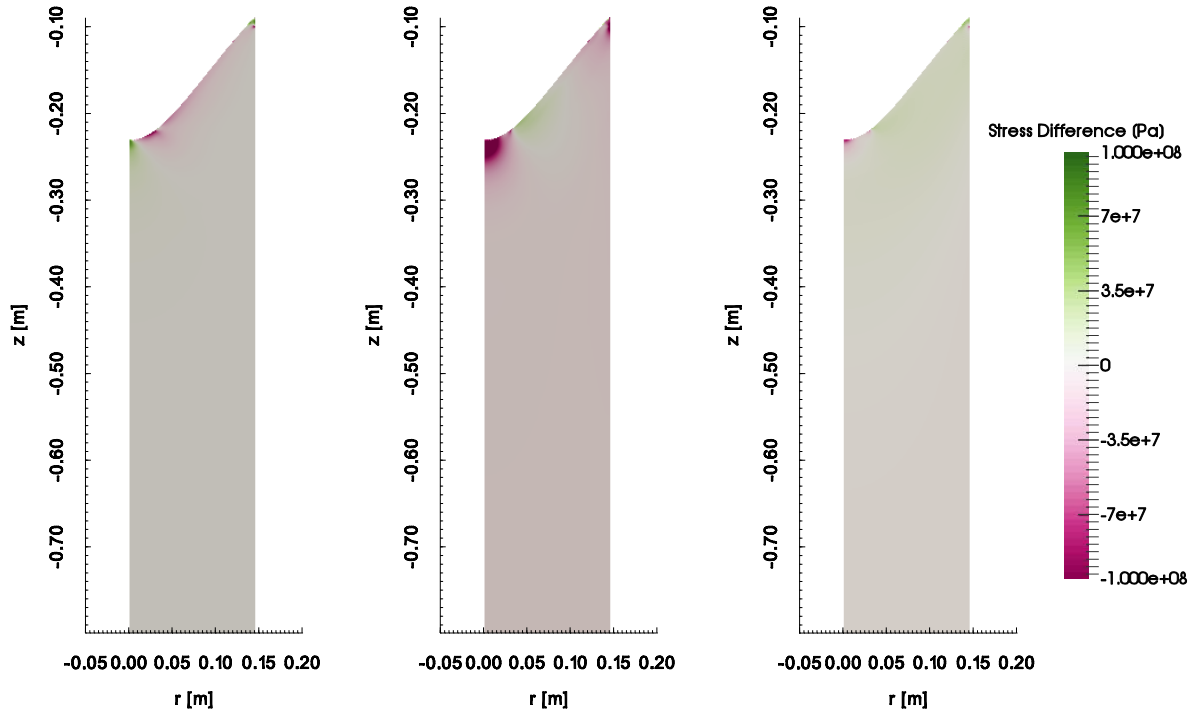


Figure 5 The difference in stress when the EMS is applied. From left to right the plots show the difference in radial, vertical, and circumferential stress.

5 CONCLUSIONS

The meshless method formulated in the paper is applied to the problem of LFEMC DC casting. The preliminary results obtained by the method show great potential in modelling of DC casting.

In this paper only a simple thermoelastic model for the stationary state of the process is considered. In the future, we plan on extending the model to incorporate plastic phenomena and to introduce full coupling between the deformation field and the heat and mass transfer model.

6 ACKNOWLEDGMENTS

Support from the Slovenian Grant Agency in the framework of the project L2-6775 and young researchers grants, Slovenian Ministry for Economy in the framework of the Voucher projects and IMPOL Aluminium Industry is kindly acknowledged.

REFERENCES

[1] Lalpoor, M. et al.: Cold cracking in DC-cast high strength aluminum alloy ingots: An intrinsic problem intensified by casting process parameters. *Mater. Sci. Eng. A*, **528** (6), 2011, p. 2831–2842.
[2] Eskin, D.G. *Physical Metallurgy of Direct Chill Casting of Aluminum Alloys*. Taylor & Francis, (2008).
[3] Fjaer, H.G., Mo, A. ALSPEN-A mathematical model for thermal stresses in direct chill casting of aluminum billets. *Metall. Trans. B* (1990) **21**:1049–1061.

[4] Drezet, J.M. et al. *Thermomechanical effects during direct chill and electromagnetic casting of aluminum alloys. Part II: numerical simulation*. TMS Publ., (1995).
[5] Kumar, P.P. et al. Mechanical Behavior of Mushy Zone in DC casting using a Viscoplastic Material Model. *Tech. Mech.* (2011) **32**:342–357.
[6] Zuo, Y. et al. Study on the sump and temperature field during low frequency electromagnetic casting a superhigh strength Al–Zn–Mg–Cu alloy. *J. Mater. Process. Technol.* (2008) **197:**109–115.
[7] Zhang, H., Cui, J. Production of super-high strength aluminum alloy billets by low frequency electromagnetic casting. *Trans. Nonferrous Met. Soc. China* (2011) **21:**2134–2139.
[8] Zhang, H. et al. Coupled modeling of electromagnetic field, fluid flow, heat transfer and solidification during low frequency electromagnetic casting of 7XXX aluminum alloys. *Mater. Sci. Eng. A* (2007) **448:** 189–203.
[9] M'Hamdi, M. et al. TearSim: A two-phase model addressing hot tearing formation during aluminum direct chill casting. *Metall. Mater. Trans. A* (2006) **37:**3069–3083.
[10] Nallathambi, A.K., Bertram, Albrecht, Specht, Eckehard et al. Thermomechanical simulation of direct chill casting. *Trans. Indian Inst. Met.* (2011) **64:**13–19.
[11] Liu, G.R. *Mesh Free Methods: Moving Beyond the Finite Element Method*. Taylor & Francis, (2010).
[12] Chen, Y. *Meshless methods in solid mechanics*. Springer, (2006).
[13] Kosec, G., Šarler, B. H-adaptive local radial basis function collocation meshless method. *Comput. Mater. Contin.*, (2011) **26:**227.
[14] Mramor, K. et al. Simulation of natural convection influenced by magnetic field with explicit local radial basis function collocation method. *CMES Comput. Model. Eng. Sci.* (2013) **92:**327–352.
[15] Vertnik, R., Šarler, B. Solution of incompressible turbulent flow by a mesh-free method. *Comput. Model. Eng. Sci. CMES* (2009) **44**:65.
[16] Kosec, G., Šarler, B.: Simulation of macrosegregation with mesosegregates in binary metallic casts by a meshless method. *Eng. Anal. Bound. Elem.* (2014) **45:**36–44.
[17] Vertnik, R.: *Heat and fluid flow simulation of the continuous casting of steel by a meshless method*. Nova Gorica: University of Nova Gorica, 2010.
[18] Kosec, G., Trobec, R.: Simulation of semiconductor devices with a local numerical approach. *Eng. Anal. Bound. Elem.* (2015) **50:**69–75.
[19] Saunders, N. et al.: Using JMatPro to model materials properties and behavior. *JOM* (2003) **55**:60–65.
[20] Košnik, N. et al.: Simulation of low frequency electromagnetic DC casting. *Mater. Sci. Forum* (2014) **790:**390–395.

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis (Eds)

MODELLING OF ELECTROMAGNETIC BREAKING AND ELECTROMAGNETIC STIRRING IN THE PROCESS OF CONTINUOUS CASTING OF STEEL

KATARINA MRAMOR[*], ROBERT VERTNIK [†] AND BOŽIDAR ŠARLER [*†]

[*] Laboratory for Multiphase Processes (LMP)
University of Nova Gorica
Vipavska 13, 5000 Nova Gorica, Slovenia
e-mail: katarina.mramor@ung.si, www.ung.si

[†] Institute for Metals and Technology (IMT)
Lepi pot 11, 1000 Ljubljana, Slovenia
e-mail: bozidar.sarler@imt.si, www.imt.si

Key words: Continuous Casting, Electromagnetic Breaking, Electromagnetic Stirring, Meshless Methods.

Abstract. More than 95 % of crude steel is nowadays processed by Continuous Casting (CC) [1]. To further advance the quality of the products and efficiency of the process, electromagnetic (EM) field, which affects the fluid flow as well as the temperature and segregation is added to the CC process. In general, there are two types of electromagnetic devices applicable to the CC process; the electromagnetic breakers (EMBR) which employ the direct current, and the electromagnetic stirrers (EMS), which employ the alternating current. Which of the devices is employed depends on what are the desired effects. Both of the processes are modelled by implementing the Lorentz force into the momentum equation, and if necessary, the Joule heating term into the energy equation. However, the way how these two terms are modelled, depends on the type of the implemented device. In case of EMBRs, the assumption of low magnetic Reynolds number Re_m is made, and consequently, the current density is calculated by solving the Poisson's equation for the electric potential. The EMSs on the other hand, require a low-frequency approximation and the solution of induction equation. The complete set of governing equations for CC process [2] under the influence of magnetic field includes mass, momentum, energy, and species transfer equations, and Maxwell's equations together with Ohm's law and charge conservation equation. Additionally, the turbulent kinetic energy and dissipation rate equations together with Abe-Kondoh-Nagano closures are used to account for the turbulence, the lever rule model is used to model the microsegregation, the mixture continuum model is used to model the macrosegregation, fractional step method is used to model pressure-velocity coupling and the enthalpy-temperature relation is used to calculate the temperature from the enthalpy. The solution is sought for on a five-nodded local subdomains by constructing an approximation with multiquadric radial basis functions as a basis and collocation to find the expansion coefficients [3,4]. Present paper presents the discretization of governing equations, together with boundary conditions for both EMBR and EMS devices with meshless Local Radial Basis Function Collocation Method (LRBFCM) [5].

1 INTRODUCTION

The modelling of application of magnetic field to the continuous casting process started as early as 1982 [6]. In 1986 [7], the 3D flow field with a rotational EMS of round steel strand was calculated, allowing to examine the influence of the stirrers position, the stirring length and EM parameters on the flow field. The first computational 3D study of EMBR systems was performed in 1982 [8]. The modelling of solidification and solute distribution was added to the model in 1998 [9]. The model considers the blockage of fluid flow by columnar dendrites in the mushy zone, the change in liquidus temperature with liquid concentration and the double diffusive convection. Turbulence of molten steel flow was modelled by $k-\varepsilon$ turbulence model in [7, 10], followed by steady Reynolds-Averaged Navier-Stokes (RANS) models [11], filtered unsteady RANS model [11], Large Eddy Simulation (LES) models [11, 12, 13] and RANS Shear Stress (RANS-SST) model [14].

The most common numerical methods used in computational modelling of continuous casting of steel are Finite Difference Method (FDM) [15, 16], Finite Volume Method (FVM) [17, 18] and Finite Element Method (FEM) [19, 20]. As the geometry in the continuous casting process is complex and the physical system requires moving and/or deforming boundaries, the mesh generation can present a substantial problem for the above mentioned numerical methods. To circumvent this problem, also meshless numerical methods have recently been considered for this problem. Among various available meshless methods, such as meshless local Petrov-Galerkin method [21], point interpolation method [22], method of fundamental solutions [23], etc., LRBFCM [5] has been chosen to tackle the problem under consideration. This method has first been applied to the CC problem in 2011 [24]. Lately, the application of magnetic field [3, 4] has been added to the model. The next step will be the application of macrosegregation [25] to the already existing heat transfer and fluid flow model of CC.

The strength of magnetic field is calculated analytically and depends on the strand distance, the number of windings in the coils, the size of the windings, and the electric current. In the present paper the discretization of governing equations or EMBR and EMS devices is presented.

2 GOVERNING EQUATIONS

The continuous casting process with applied magnetic field can be described by five conservation equations; namely the mass, momentum, energy, species and charge conservation equations,

$$\nabla \cdot \mathbf{v} = 0 \tag{1}$$

$$\frac{\partial(\rho \mathbf{v})}{\partial t} + \nabla \cdot (\rho \mathbf{v}\mathbf{v}) = -\nabla p + \nabla \cdot \left[\left(\mu_L + \mu_t \frac{\rho}{\rho_L} \right) \left[\nabla \mathbf{v} + (\nabla \mathbf{v})^T \right] \right] - \frac{2}{3} \nabla(\rho k) - \tag{2}$$

$$- \mu_L \frac{K_0 (1 - f_L)^2}{f_L^3} (\mathbf{v} - \mathbf{v}_S) + \rho \mathbf{g} \left(\beta_T (T - T_{ref}) + \beta_C (C - C_{ref}) \right) + \mathbf{j} \times \mathbf{B}$$

$$\frac{\partial(\rho h)}{\partial t}+\nabla\cdot\left(\rho\mathbf{v}h\right)=\nabla\cdot\left(\lambda\nabla T\right)+\nabla\cdot\left(\rho f_S\left(h_L-h_S\right)(\mathbf{v}-\mathbf{v}_S)\right)+ \\ +\nabla\cdot\left(f_L\frac{\rho_L\nu_t}{\sigma_t}\nabla h_L\right)+\frac{|\mathbf{j}|^2}{\sigma_0} \tag{3}$$

$$\frac{\partial(\rho C)}{\partial t}+\nabla\cdot\left(\rho C\mathbf{v}\right)=\nabla\cdot\left(\rho f_S D_S\nabla C_S+\rho f_L D_L\nabla C_L\right)+ \\ +\nabla\cdot\left(\rho(C_L-C)(\mathbf{v}-\mathbf{v}_S)\right)+\nabla\cdot\left(\frac{f_L\mu_t}{\sigma_C}\nabla C_L\right) \tag{4}$$

$$\nabla\cdot\mathbf{j}=-\frac{\partial\rho_e}{\partial t}=0 \tag{5}$$

where $\mathbf{v}$ stands for the velocity of the mixture, $\rho=\rho_S=\rho_L$ is the density, assumed to be constant and equal in both phases. t stands for time and p for pressure. μ_t is turbulent viscosity and μ_L is the dynamic viscosity, k represents turbulent kinetic energy and K_0 is the permeability constant. $\mathbf{v}_S$, β_T, β_C, $\mathbf{g}$, T, T_{ref}, C, and C_{ref} represent velocity of solid phase, thermal expansion coefficient, solute expansion coefficient, gravitational acceleration, temperature, reference temperature, species concentration and reference species concentration, respectively. $\mathbf{j}\times\mathbf{B}$ is Lorentz force, and is in detail described below. h stands for enthalpy, λ for thermal conductivity, and σ_0 for electrical conductivity. f_S, f_L, h_S, and h_L represent solid volume fraction, liquid volume fraction, enthalpy of the solid phase and enthalpy of the liquid phase, and ν_t is turbulent kinematic viscosity. D_S, and D_L are diffusion coefficients for solid and liquid phase respectively. ρ_e and $\mathbf{j}$ are electric charge density and current density. C_S and C_L are the concentration of solute in solid and liquid phase, respectively. Level rule microsegregation model is used to determine the liquid fraction, structured with

$$f_L=1-\frac{1}{1-k_p}\frac{T-T_L}{T-T_m}, \tag{6}$$

the partition ratio

$$k_p=\frac{C_S}{C_L}, \tag{7}$$

the liquidus temperature

$$T_L=T_m+(T_e-T_m)\frac{C}{C_e}, \tag{8}$$

and the concentration of solute in liquid phase

$$C_L=\frac{C}{1+f_S(k_p-1)}. \tag{9}$$

T_m, T_e, and C_e are the melting temperature, the eutectic temperature and the eutectic solute concentration, respectively.

Additionally, the turbulent kinetic energy and the dissipation rate equations are added to account for the turbulence

$$\frac{\partial(\rho k)}{\partial t}+\nabla\cdot\left(\rho\mathbf{v}k\right)=\nabla\cdot\left[\left(\mu_L\frac{\rho}{\rho_L}+\frac{\mu_t}{\sigma_k}\right)\nabla k\right]+P_k+G_k-\rho\varepsilon \\ -\rho D_{k-\varepsilon}+\mu_L\frac{K_0\left(1-f_L\right)^2}{f_L^3}\frac{\rho}{\rho_L}k \tag{10}$$

$$\frac{\partial(\rho\varepsilon)}{\partial t}+\nabla\cdot\left(\rho\mathbf{v}\varepsilon\right)=\nabla\cdot\left[\left(\mu_L\frac{\rho}{\rho_L}+\frac{\mu_t}{\sigma_\varepsilon}\right)\nabla\varepsilon\right]+\rho E_{k-\varepsilon}- \\ -\mu_L\frac{K_0\left(1-f_L\right)^2}{f_L^3}\varepsilon+\left[c_{1\varepsilon}f_1\left(P_k+c_{3\varepsilon}G_k\right)-c_{2\varepsilon}f_2\rho\right]\frac{\varepsilon}{k} \tag{11}$$

where ε stands for the dissipation rate, σ_t, σ_C, σ_k, σ_ε, $c_{1\varepsilon}$, f_1, $c_{2\varepsilon}$ and f_2 are closure coefficients. P_k, G_k, $D_{k-\varepsilon}$, and $E_{k-\varepsilon}$ are the shear production of turbulent kinetic energy, generation of turbulence due to the buoyancy force, source term in k equation and source term in ε equation, respectively. Abe-Kondoh-Nagano closures are used [26].

Maxwell's equations together with Ohm's law are used to calculate the magnetic field effects

$$\nabla\cdot\mathbf{E}=\frac{\rho_e}{\varepsilon_0}, \tag{12}$$

$$\nabla\cdot\mathbf{B}=0, \tag{13}$$

$$\nabla\times\mathbf{E}=-\frac{\partial\mathbf{B}}{\partial t}, \tag{14}$$

$$\nabla\times\mathbf{B}=\mu_0\left(\mathbf{j}+\varepsilon_0\frac{\partial\mathbf{E}}{\partial t}\right), \tag{15}$$

$$\mathbf{j}=\sigma(-\nabla\phi+\mathbf{v}\times\mathbf{B}), \tag{16}$$

where $\mathbf{E}$, ε_0, $\mathbf{B}$, and μ_0 are electric field, permittivity of free space, magnetic field density, and permeability of free space. The extent of coupling between the magnetic field and velocity, temperature, species concentration or energy, depends on the device (EMBR or EMS) that is used to produce the EM field. The coupling in both cases is done through the Lorentz force

$$\mathbf{F}_m=\mathbf{j}\times\mathbf{B}, \tag{17}$$

which is added to the momentum equation and Joule heating term

$$\frac{|\mathbf{j}|^2}{\sigma_0}, \tag{18}$$

that is added to the energy equation.

3 ELECTROMAGNETIC FIELD EQUATIONS

Both the EMBR and EMS are incorporated into the model through the Lorentz force, and if necessary, the Joule heating term. However, the magnetic field and consequently the Lorentz force for each of the EM devices are calculated in a different way. The calculation procedure is outlined below. In both, EMS and EMBR cases, an assumption of low $\mathrm{Re}_m=vL\mu_0\sigma_0$ is made. v

is the characteristic velocity and L is the characteristic length. The solution of EMS additionally requires a low frequency approximation.

3.1 Electromagnetic braking

The considered device for EMBR consist of two coils facing in the same direction as schematically presented in Figure 1. This coil configuration consists with EMBR ruler. The magnetic field for such coil arrangement can be calculated analytically [27, 28] by first calculating the magnetic field components of each individual winding

$$B_x = \frac{\mu_0 I}{4\pi}\sum_{i=1}^{4}\left[\frac{(-1)^i z_i}{r_i(r_i+(-1)^{i+1}y_i)} - \frac{y_i}{r_i(r_i+z_i)}\right] \tag{19}$$

$$B_y = \frac{\mu_0 I}{4\pi}\sum_{i=1}^{4}\frac{(-1)^{i+1}x}{r_i(r_i+z_i)} \tag{20}$$

$$B_z = \frac{\mu_0 I}{4\pi}\sum_{i=1}^{4}\frac{(-1)^{i+1}x}{r_i(r_i+(-1)^{i+1}y_i)} \tag{21}$$

where x, $y_{1,3} = y + \frac{a}{2}$, $y_{2,4} = y - \frac{a}{2}$, $z_{1,2} = z + \frac{b}{2}$, $z_{3,4} = z - \frac{b}{2}$, $r_i = \sqrt{x^2 + y_i^2 + z_i^2}$. The assumption of tightly wound loops placed next to each other is then made. By substituting $x \to x + n \cdot \phi_w$, $y \to y + m \cdot \phi_w$ and $z \to z + m \cdot \phi_w$, where n, m, and ϕ_w are the number of loops in x direction, the number of loops in y and z directions and the diameter of the wire, the total field of a single solenoid coil is obtained by summing up all of the contributions

$$\mathbf{B} = \mathbf{B}_1 + \ldots + \mathbf{B}_N . \tag{22}$$

The parallel coil configuration is obtained by employing a transformation $x \to x + \frac{d}{2}$ and $x \to x - \frac{d}{2}$, where d is the distance between coils, and the summation of both contributions is $\mathbf{B} = \mathbf{B}_{-\frac{d}{2}} + \mathbf{B}_{+\frac{d}{2}}$.

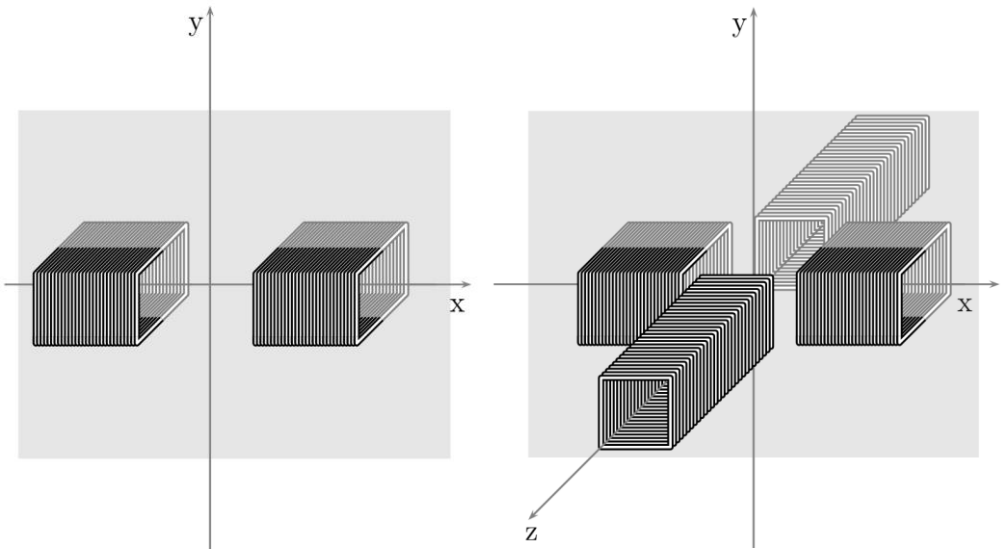


Figure 1: Left: Scheme of EMBR configuration. Right: Scheme of EMS configuration.

The magnetic field is then inserted into the electric potential equation

$$\nabla^2\phi = \nabla\cdot(\mathbf{v}\times\mathbf{B}) \tag{23}$$

which is obtained by inserting Ohm's law (Eq. 9) into the charge conservation equation (Eq. 5). Once the electric potential is calculated, it is inserted into Eq. 9 in order to obtain the current density, which is then used to calculate Lorentz force and, if needed, Joule source term.

3.2 Electromagnetic stirring

The considered device for EMS consists of four coils facing each other. The coils employ alternating current to produce a time varying magnetic field. By introducing the vector $\mathbf{A}$ and scalar ϕ potentials and enforcing the Coulomb gauge condition ($\nabla\cdot\mathbf{A}=0$), the magnetic and the electric fields are rewritten as

$$\mathbf{B} = \nabla\times\mathbf{A}\,, \tag{24}$$

$$\mathbf{E} = -\nabla\phi - \frac{\partial\mathbf{A}}{\partial t}\,. \tag{25}$$

Inserting Eqs. (24) and (25) into Maxwell's equations (Eqs. 12-15), the following equations are obtained

$$\nabla^2\mathbf{A} = \mu_0\sigma_0\frac{\partial\mathbf{A}}{\partial t} - \mu_o\sigma_0\mathbf{v}\times(\nabla\times\mathbf{A}) - \mu_0\sigma_0\nabla\phi\,, \tag{26}$$

$$\nabla^2\phi = 0\,, \tag{27}$$

and consequently, ϕ can be set to $\phi = const.$ By applying the Ohm's law, Eq. 26 can be expressed as

$$\nabla^2 \mathbf{A} = -\mu_0 (\mathbf{j}_{ind} + \mathbf{j}_{ext}), \tag{28}$$

where $\mathbf{j}_{ind}$ and $\mathbf{j}_{ext}$ are induced and imposed currents, respectively. The solution of thus obtained linear system of equations can be further simplified by taking into account that the imposed current density is harmonic and can therefore be written as

$$\mathbf{j}_{ext} = \mathbf{j}_{ext}(\mathbf{r}) e^{i\omega t}, \tag{29}$$

where $\mathbf{j}_{ext}(\mathbf{r})$ is the amplitude of current density in the coils and ω is the source frequency [29, 30]. The Lorentz force and, if necessary, the Joule source terms are calculated as a real part of the complex fields.

4 DISCRETIZATION OF GOVERNING EQUATIONS

The discretization procedure of LRBFCM is discussed in the continuation of the present paper. As all the governing equations follow the conservation principle, they can be described with a general transport equation as

$$\frac{\partial(\rho\Phi)}{\partial t} + \nabla \cdot (\rho \mathbf{v} \Phi) = \nabla \cdot (D \nabla \Phi) + S_\Phi, \tag{30}$$

where D is diffusion coefficient, Φ is general dependent variable (e.g. velocity) and S_Φ a source term. The first term on the left side in the equation is called the transient term, whereas the second term is the convection term. On the right side, the first term is the diffusion term. The description of discretization is demonstrated for a general transport equation, as the discretization of mass, momentum, energy, species conservation, turbulent kinetic energy and dissipation rate governing equations follows the same principle.

4.1 Time discretization

The time discretization is performed by explicit first-order approximation (explicit Euler) scheme. The transient term (first term in Eq. 30) is thus rewritten as

$$\frac{\partial(\rho\Phi)}{\partial t} \approx \frac{\rho(t_0 + \Delta t)\Phi(t_0 + \Delta t) - \rho(t_0)\Phi(t_0)}{\Delta t} = \frac{\rho\Phi - \rho_0\Phi_0}{\Delta t} \approx \rho_0 \frac{\Phi - \Phi_0}{\Delta t} \tag{31}$$

where $\Phi(t_0 + \Delta t)$ and $\Phi(t_0)$ represent the value of sought for variable at time $t_0 + \Delta t$ and t_0, respectively, and ρ_0 represents the value of density at time t_0.

4.2 Space discretization

Space discretization in LRBFCM is done by collocation with RBFs. The general idea behind this method is to construct an approximation function on the local group of nodes, the so-called influence domains, and to apply the PDE on the approximation functions in a strong formulation. The approximation function Φ is constructed on a local subdomain $_l$ and is represented as a linear combination of weighted basis functions $_l\psi_i$

$$ {}_l\Phi(\mathbf{p}) = \sum_{i=1}^{N_{basis}} {}_l\psi_i(\mathbf{p}) {}_l\alpha_i , \tag{32} $$

where ${}_l\alpha_i$, N_{Basis}, ${}_l\psi_i$, and $\mathbf{p} = x\mathbf{i}_x + y\mathbf{i}_y + z\mathbf{i}_z$ are the expansion coefficient, the number of basis functions, the basis function, and the position vector in influence domain ${}_l$, where $l = 1,\ldots,N_{domain}$, respectively. The position vector is expressed in Cartesian coordinate system with coordinates x, y, z and base vectors $\mathbf{i}_x, \mathbf{i}_y, \mathbf{i}_z$. Among the various possible basis functions, such as polynomials, Fourier basis functions, the multiquadric RBFs are chosen

$$ \psi(r) = \sqrt{r^2 + c^2} , \tag{33} $$

where ${}_l r_j = \| \mathbf{p} - {}_l\mathbf{p}_j \| = \sqrt{\left(\frac{x - {}_l x_j}{{}_l x_{max}}\right)^2 + \left(\frac{y - {}_l y_j}{{}_l y_{max}}\right)^2 + \left(\frac{z - {}_l z_j}{{}_l z_{max}}\right)^2}$ is the distance between the central node $\mathbf{p}$ and the support node ${}_l\mathbf{p}_j$ and c is the shape parameter. The distances between the current position $\mathbf{p}$ and the support nodes in the influence domain region are normalized by the maximum lengths in subdomain ${}_l x_{max}$, ${}_l y_{max}$, and ${}_l z_{max}$. This allows to use a constant value of ($c = 32$) in differently arranged subdomains. The expansion coefficients of the approximation function are determined by collocation, which demands that the number of basis functions equals the number of the domain nodes. The collocation condition $\Phi({}_l\mathbf{p}_j) = {}_l\Phi_j$ must hold for all the points in the influence domain. By implementing the collocation condition, a linear system of equations is obtained

$$ {}_l\boldsymbol{\psi}\, {}_l\boldsymbol{\alpha} = {}_l\underline{\Phi} , \tag{34} $$

where ${}_l\boldsymbol{\psi}$, ${}_l\boldsymbol{\alpha}$, and ${}_l\underline{\Phi}$ are the matrix of RBFs, the vector of the expansion coefficients, and the vector of corresponding data values, respectively. If the matrix is non-singular [31], the expansion coefficients can be obtained as

$$ {}_l\boldsymbol{\alpha} = {}_l\boldsymbol{\psi}^{-1}\, {}_l\underline{\Phi} . \tag{35} $$

The approximation function thus becomes

$$ {}_l\Phi(\mathbf{p}) = \sum_{i=1}^{N_{Basis}} {}_l\psi_i(\mathbf{p}) \sum_{j=1}^{N_{Basis}} {}_l\psi_{ij}^{-1}(\mathbf{p})\, {}_l\Phi_j . \tag{36} $$

The convection term is first rewritten as

$$ \nabla\cdot(\rho\mathbf{v}\Phi) = \Phi\nabla\cdot(\rho\mathbf{v}) + \rho\mathbf{v}\nabla\cdot\Phi \tag{37} $$

and spatial discretization is then applied to each of the terms separately. The first term is rewritten as

$$ \left[\Phi\nabla\cdot(\rho\mathbf{v})\right]_l \approx \Phi \sum_{i=1}^{N_{Basis}} \frac{\partial}{\partial x_j} {}_l\psi_i \sum_{n=1}^{N_{Basis}} {}_l\psi_{in}^{-1}\, {}_l(\rho\mathbf{v})_n \tag{38} $$

and the second term as

$$(\rho \mathbf{v} \nabla \cdot \Phi)_l \approx \rho \mathbf{v} (\sum_{i=1}^{N_{domain}} \frac{\partial}{\partial x_j} {}_l\psi_i \sum_{n=1}^{N_{domain}} {}_l\psi_{in}^{-1} {}_l\Phi_n) . \tag{39}$$

In the present case, only the second term (Eq. 39) is calculated, as the density is considered constant. Similarly, the diffusion term is first rewritten as

$$\nabla \cdot (D \nabla \Phi) = \nabla D \cdot \nabla \Phi + D \nabla^2 \Phi , \tag{40}$$

where D is a diffusion coefficient. The first term in Eq. (40) is discretized as

$$(\nabla D \cdot \nabla \Phi)_l \approx (\sum_{i=1}^{N_{domain}} \frac{\partial}{\partial x_j} {}_l\psi_i \sum_{n=1}^{N_{domain}} {}_l\psi_{in}^{-1} {}_lD_n) \cdot (\sum_{i=1}^{N_{domain}} \frac{\partial}{\partial x_j} {}_l\psi_i \sum_{n=1}^{N_{domain}} {}_l\psi_{in}^{-1} {}_l\Phi_n) \tag{41}$$

and the second as

$$(D \nabla^2 \Phi)_l \approx D_l \sum_{i=1}^{N_{Basis}} \frac{\partial^2}{\partial x_j^2} {}_l\psi_i \sum_{n=1}^{N_{Basis}} {}_l\psi_{in}^{-1} {}_l\Phi_n . \tag{42}$$

By using the above described procedure for time and space discretization, a general transport equation (Eq. 31) becomes

$$\begin{aligned} \Phi(t_0 + \Delta t)_l = \Phi(t_0)_l + \frac{\Delta t}{\rho_0} \Big(& \Phi \sum_{i=1}^{N_{Basis}} \frac{\partial}{\partial x_j} {}_l\psi_i \sum_{n=1}^{N_{Basis}} {}_l\psi_{in}^{-1} {}_l(\rho_0 \mathbf{v})_n - \\ & - \rho_0 \mathbf{v} (\sum_{i=1}^{N_{Basis}} \frac{\partial}{\partial x_j} {}_l\psi_i \sum_{n=1}^{N_{Basis}} {}_l\psi_{in}^{-1} {}_l\Phi_n) + \\ & + (\sum_{i=1}^{N_{Basis}} \frac{\partial}{\partial x_j} {}_l\psi_i \sum_{n=1}^{N_{Basis}} {}_l\psi_{in}^{-1} {}_lD_n) \cdot (\sum_{i=1}^{N_{Basis}} \frac{\partial}{\partial x_j} {}_l\psi_i \sum_{n=1}^{N_{Basis}} {}_l\psi_{in}^{-1} {}_l\Phi_n) + \\ & + D_l \sum_{i=1}^{N_{Basis}} \frac{\partial^2}{\partial x_j^2} {}_l\psi_i \sum_{n=1}^{N_{Basis}} {}_l\psi_{in}^{-1} {}_l\Phi_n + {}_l S_\Phi \Big)_0 \end{aligned} . \tag{43}$$

4.3 Discretization of Poisson equations

The discretization of Poisson's equation is a boundary value problem, whereas the previously described discretization is an initial value problem. Respectively, its discretization by LRBFCM is described separately. It is used for the solution of electric scalar potential in EMBR (Eq. 23), vector potential in EMS (Eq. 28), and pressure in FSM method, which is used for coupling of the mass conservation with the momentum conservation equations. A detailed description of FSM pressure-velocity coupling solution procedure and its implementation in the LRBFCM can be found in [2-5]. The discretization procedure for Poisson's equation is presented for a general variable. The global $\mathbf{p}_n$ and the local influence domain ${}_l\mathbf{p}_i$ nodes coincide and are connected by the following relation $\mathbf{p}_{n(l,i)} = {}_l\mathbf{p}_i$. The general variable is represented on each of the influence domains as a linear combination of basis functions and expansion coefficients as

$${}_l\Phi(\mathbf{p}) = \sum_{i=1}^{N_{domain}} \psi_{n(l,i)}(\mathbf{p}) \, {}_l\alpha . \tag{44}$$

The expansion coefficients are determined from collocation, as presented in Eq. 35. Φ is therefore determined in each of the subdomains as

$$ {}_l\Phi(\mathbf{p}) = \sum_{i=1}^{N_{Basis}} \sum_{j=1}^{N_{Basis}} {}_l\psi_{n(l,i)}(\mathbf{p})\, {}_l\Psi_{ij}^{-1} \Phi_{n,(l,j)} . \tag{45} $$

The discretized general Poisson equation is a result of collocation and application of Laplace operator in global node $\mathbf{p}_k$

$$ \sum_{k=1}^{N_{domain}} \Psi_{kl} \Phi_l = S_k , \tag{46} $$

where Ψ_{kl} is the sparse matrix element and S_k is a function of Φ. For scalar potential $S_k = \nabla \cdot (\mathbf{v} \times \mathbf{B})$, for vector potential $S_k = -\mu_0 (\mathbf{j}_{ind} + \mathbf{j}_{ext})$ and for pressure $S_k = \frac{\rho}{\Delta t} \nabla \cdot \mathbf{v}^*$, where $\mathbf{v}^*$ is the intermediate velocity.

4.4 Discretization of boundary conditions

In general, there are three types of boundary conditions: the Neumann, Dirichlet and Robin ones, all of which are used in the solution of governing equations for the CC of steel under the influence of magnetic field. The implementation of Dirichlet boundary conditions is straightforward

$$ {}_l\Phi(\mathbf{p}) = {}_l\,\Phi_{BC} . \tag{47} $$

The implementation of Neumann boundary conditions on the other hand requires the application of collocation

$$ \frac{\partial}{\partial \mathbf{n}} {}_l\Phi(\mathbf{p}) = \sum_{j=1}^{N_{domain}} \frac{\partial}{\partial \mathbf{n}} {}_l\psi_j(\mathbf{p})\, {}_l\alpha_j . \tag{48} $$

The same is true for Robin boundary conditions

$$ a \frac{\partial}{\partial \mathbf{n}} {}_l\Phi(\mathbf{p}) + b\, {}_l\Phi(\mathbf{p}) = \sum_{j=1}^{N_{domain}} \left(a \frac{\partial}{\partial \mathbf{n}} {}_l\psi_j(\mathbf{p}) + b\, {}_l\psi_j(\mathbf{p})\right) {}_l\alpha_j , \tag{49} $$

where $\mathbf{n}$ is normal to the boundary, and a and b are the weights.

5 CONCLUSIONS

In this paper, a local meshless procedure for discretization of governing equations and their boundary conditions for a coupled multiphysics problem, resulting from EMBR and EMS in CC is presented. The general local form of discretized equations is applied to the transport equations of mass, momentum, species, energy, turbulent kinetic energy and dissipation rate. The solution of Poisson's equation, which is representative for solving the electric vector potential and pressure Poisson's equation, is discretized. The discretisation results in solving of a global sparse matrix for all the nodes in calculation domain. In both cases the collocation with RBFs is used for local space discretization on influence domains and explicit Euler scheme is

used for time discretization. Several numerical results, stemming from the presented meshless computational scheme will be presented at the conference.

6 ACKNOWLEDGEMENTS

The research in this paper is sponsored by Slovenian Grant Agency under program group P2-0369 Modelling and Simulation of Materials and Processes and by the project L2-6775: Simulation of Industrial Solidification Processes under Influence of Electromagnetic Fields, co-sponsored by Štore Steel and IMPOL companies. The EMBR part of the research was sponsored by Centre of Excellence for Biosensors, Instrumentation and Process Control.

REFERENCES

[1] World Steel Association, *http: //www.worldsteel.org*, (2013).
[2] Vertnik, R. and Šarler, B. Solution of a continuous casting of steel benchmark test by meshless, *Eng. Anal. Bound. Elem.* (2014) **45**:45–1282.
[3] Mramor, K., Vertnik, R. and Šarler, B. Simulation of natural convection influenced by magnetic field with explicit local radial basis function collocation method, *Comp. Model. Eng. Sci.* (2013) **92**: 327–352.
[4] Mramor, K., Vertnik, R. and Šarler, B. Simulation of laminar backward facing step flow under magnetic field with explicit local radial basis function collocation method, *Eng. Anal. Bound. Elem.* (2006) **51:** 1269–1282.
[5] Šarler, B. and Vertnik, R. Meshfree explicit local radial basis function collocation method for diffusion problems, *Comp. Math. Appl.* (2006) **51**: 1269–1282.
[6] Birat, J.P. and Chone, J. Electromagnetic stirring on billet, bloom, and slab continuous casters: state of the art in 1982, *Iron. Steel.* (1982) **10**: 269–281.
[7] Spitzer, K. H., Dubke, M. and Schwerdtfeger, K. Rotational electromagnetic stirring in continuous casting of round strands, *Metal. Trans. B* (1986) **17**: 119–131.
[8] Yao, M., Ichimiya, M., Kiyohara, S., Suzuki. K.I., Mesaki, R., and Sugiyama, K. Three-dimensional analysis of molten metal flow in continuous casting mould. *Iron. Steel. Soc. Inc.* (1992) 123–129.
[9] Yang, H., Zhang, X., Deng, K., Li, W., Gan, Y., and Zhao, L. Mathematical simulation on coupled flow, heat, and solute transport in slab continuous casting process, *Metal. Mater. Trans. B* (1998) **29**: 1345–1356.
[10] Cukierski, K. and Thomas, B. G. Flow control with local electromagnetic braking in continuous casting of steel slabs, *Metal. Mater. Trans. B* (2008) **39**: 94-107.
[11] Chaudhary, R., Ji, C., Thomas, B.G., and Vanka, S. P. Transient turbulent flow in a liquid-metal model of continuous casting, including comparison of six different methods, *Metal. Mater. Trans. B* (2011) **42**: 987–1007.
[12] Chaudhary, R., Thomas, B.G., and Vanka, S. P. Effect of electromagnetic ruler braking (EMBr) on transient turbulent flow in continuous slab casting using large eddy simulations, *Metal. Mater. Trans. B* (2012) **43**: 532–553.
[13] Ji, C.-B., Li, J.-S., Yang, S.-F., and Sun, L.-Y. Large eddy simulation of turbulent fluid flow in liquid metal of continuous casting, *J. Iron. Steel. Res. Int.* (2013) **20**: 34–46.

[14] Miao, X., Timmel, K., Lucas, D., Ren, Z., Eckert, S., and Gerbeth, G. Effect of an electromagnetic brake on turbulent melt flow in a continuous casting mold, *Metal. Mater. Trans. B* (2012) **43**: 954–972.
[15] Özişik, M. N. *Finite Difference Methods in Heat Transfer.* CRC Press (1994).
[16] Wang, L., Shen, H., and Liu, B. Effect of electromagnetic field on fluid flow in continuous casting mold. *MCSP6-2004* (2004) 13-20.
[17] Zienkiewicz, O. C. and Taylor, R. L. *The Finite Element method*. McGraw Hill, Vol. I., (1989), Vol. II. (1991).
[18] Trinidade, L. B., Vilela, A. C. F., Vilhena, M. T. M. B., Soares, R. B. et al. Numerical model of electromagnetic stirring for continuous casting billets, *IEEE Trans. Mag.* (2002) **38**:3658-3660.
[19] Versteeg, H. K. and Malalasekera W. *An Introduction to Computational Fluid Dynamics: the Finite Volume Method*. Prentice Hall, (2007).
[20] Haiqi, Y., Baofeng, W., Huiqin, L., and Jianchao, L. Influence of electromagnetic brake on flow field of liquid steel in the slab continuous casting mold, *J. Mater. Process. Tech.* (2008) **202**:179-187.
[21] Atluri, S.N. and Zhu, T. A new meshless local Petrov-Galerkin (MLPG) approach in computational mechanics, *Comp. Mech.* (1998) **22**:117-127.
[22] Liu, G. R. and Gu, Y. T. A point interpolation method for two-dimensional solids, *Int. J. Numer. Method. Eng.* (2001) **50**:937-951.
[23] Golberg, M. A. and Chen, C. S. *Boundary Integral Methods-Numerical and Mathematical Aspects*. WIT Press, (1998).
[24] Vertnik, R. and Šarler, B. Local collocation approach for solving turbulent combined forced natural convection problems, *Adv. Appl. Math. Mech.* (2011) **3**:259-279.
[25] Kosec, G. and Šarler, B. Simulation of macrosegregation with mesosegregation in binary metallic casts by meshless method, *Eng. Anal. Bound. Elem.* (2014) **23**:189-204.
[26] Abe, K., Kondoh, T., and Nagano, Y. A new turbulence model for predicting fluid flow and heat transfer in separating and reattaching flows- I. Flow field calculations, *Int. J. Heat. Mass. Transf.* (1994) **37**:139-151.
[27] Li, T. T.-K. Tri-axial square Helmholtz coil for neutron EDM experiment, *Chinese Uni. Of Hong-Kong* - Report (2004) **1**: 1-23.
[28] Misakian, M. Equations for the magnetic field produced by one or more rectangular loops of wire in the same plane. *J. Res. Nat. Inst. Stand. Tech.* (2000) **105:**557-564.
[29] Maldovan, M., Príncipe, J., Sánchez, G., Pignotti, A. and Goldschmit, M. Numerical modeling of continuous casting of rounds with electromagnetic stirring, *ECOMAS* (2010).
[30] Bondeson, A., Rylander, T., and Ingelström, P. *Computational Electromagnetics*. Springer, (2010).
[31] Hon, Y. C., and Schaback, R. On unsymmetric collocation by radial basis functions, *Appl. Math. Comp.* (2001) **119**:177-186.

SIMULATION OF MULTISCALE INDUSTRIAL SOLIDIFICATION PROBLEM UNDER INFLUENCE OF ELECTROMAGNETIC FIELD BY MESHLESS METHOD

B. ŠARLER[1,2], A. Z. GUŠTIN[1], V. HATIĆ[1], N. KOŠNIK[3,4], B. MAVRIČ[1] AND R. VERTNIK[1,5]

[1] Institute of Metals and Technology, Lepi pot 11, SI-1000 Ljubljana, Slovenia, e-mail: bozidar.sarler@imt.si, www.imt.si

[2] University of Nova Gorica, Vipavska 13, SI-5000 Nova Gorica, Slovenia, e-mail: bozidar.sarler@ung.si, www.ung.si

[3] Jožef Stefan Institute, Jamova 39, SI-1000 Ljubljana, Slovenia

[4] Department of Physics, University of Ljubljana, Jadranska 19, SI-1000 Ljubljana, Slovenia

[5] Štore Steel, d.o.o., Železarska cesta 3, SI-3220 Štore, Slovenia

Key words: Solidification, Low Frequency Electromagnetic DC Casting, Multi-Scale Model, Multi-Physics Model, Coupled Problem, Meshless Methods.

Abstract. Simulation and control of macrosegregation, deformation and grain size under electromagnetic (EM) processing conditions is important in industrial solidification systems, since it influences the quality of the casts and consequently the whole downstream processing path. Respectively, a multiphysics and multiscale model is developed for solution of Lorentz force, temperature, velocity, concentration, deformation and grain structure of the casts. The mixture equations with lever rule, linearized phase diagram, and stationary thermoelastic solid phase are assumed, together with EM induction equation for the field imposed by the low frequency EM field or Ohm's law and charge conservation equation for stationary EM field. Turbulent effects are incorporated through the solution of a low-Re turbulence model. The solidification system is treated by the mixture-continuum model, where the mushy zone is modeled as a Darcy porous media with Kozeny-Karman permeability relation and columnar solid phase moving with the system velocity. Explicit diffuse approximate meshless solution procedure [1] is used for solving the EM field, and the explicit local radial basis function collocation method [2] is used for solving the coupled transport phenomena and thermomechanics fields. Pressure-velocity coupling is performed by the fractional step method [3]. The point automata method with modified KGT model is used to estimate the grain structure [4] in a post-processing mode. Thermal, mechanical, EM and grain structure outcomes of the model are demonstrated for low frequency EM casting of round aluminium billets. A systematic study of the complicated influences of the process parameters on the microstructure can be investigated by the model, including intensity and frequency of the electromagnetic field.

1 INTRODUCTION

Direct-Chill (DC) casting [5] turns out to be an efficient technology for production of aluminium semi-products. The process is robust and relatively simple, however it can induce a spectra of defects in the ingots. The control of the process, depending on the product geometry and alloy composition, basically relies on adjusting the mould level, the casting speed, the melt temperature, and the intensity of the water jets that chill the ingot. These control parameters indirectly influence the melt flow and thus the solidification characteristics. An additional handle that can directly affect the flow is offered by alternating EM fields [6], applied to the DC process by means of coils placed around the casting system. They induce the eddy currents in the melt and the ingot that dissipate heat and, more importantly, assert Lorentz force on them, thus providing a stirring force on the melt. A pioneering study of applying low frequency electromagnetic field to the casting process was presented in [7], while a recent magnetohydrodynamic study of this problem was done by our group in [8]. The main aim of the present paper is the extension of this model for including macrosegregation, mechanical effects and grain structure and thus forming a reasonably complete basic multiphysics and multiscale model of the process for the axisymmetric billets.

2 MACROSCOPIC MODELS

2.1 Thermofluid equations

We express all the governing macroscopic transport equations in cylindrical coordinates (r,z), with base vectors $\hat{\mathbf{e}}_r$ and $\hat{\mathbf{e}}_z$ with z coordinate opposite to the casting direction. We limit ourselves to the cases where velocity vector points in the $r-z$ plane. The mixture continuum concept [9], adjusted to DC casting with columnar only solid phase [10] is adopted. The momentum conservation accounts for the Darcy drag term, and the volume force, decomposed into thermal and solutal Boussinesq buoyancy term and the Lorentz force:

$$\rho_m \frac{\partial}{\partial t}\left(\mathbf{v}_m\right)+\rho_m \nabla\cdot\left(\mathbf{v}_m\mathbf{v}_m\right)=-\nabla P_l+\nabla\cdot\left(\mu_l\frac{\rho_m}{\rho_l}\nabla\mathbf{v}_m\right)-\frac{\mu_l}{K}\frac{\rho_m}{\rho_l}\left(\mathbf{v}_m-\mathbf{v}_{\text{cast}}\right)+\mathbf{b} \tag{1}$$

$$\mathbf{b}=-\rho_l\left[\beta_{T,l}\left(T-T_0\right)+\sum_{i=1}^{n}\beta_{C,l}\left(C_l^i-C_0^i\right)\right]\mathbf{g}+\left\langle\mathbf{b}_{\text{EM}}\right\rangle \tag{2}$$

The latter is denoted by $\left\langle\mathbf{b}_{\text{EM}}\right\rangle$ and will be detailed in the following section. The solid phase velocity $\mathbf{v}_S$ is assumed to be constant and equal to the casting velocity $\mathbf{v}_s=\mathbf{v}_{\text{cast}}$, that is parallel to the z coordinate. The mixture density and velocity is defined as $\rho_m=f_l\rho_l+f_s\rho_l$, $\mathbf{v}_m=\left(f_l\rho_l\mathbf{v}_l+f_s\rho_s\mathbf{v}_s\right)/\rho_m$, with f_l and f_s standing for the volume fraction of the liquid and solid phase, respectively. Mass conservation thus reads $\nabla\cdot\mathbf{v}_m=0$. The permeability constant of the Darcy term, relevant in the mushy zone, is modelled as $K=K_0 f_L^3\left(1-f_L\right)^{-2}$ and is put to 0 by hand below the consolidation temperature corresponding to $f_l^{\lim}$, where the liquid phase becomes trapped within the dendrites. The heat and species transfer is formulated with

the mixture convection-diffusion equations:

$$\frac{\partial}{\partial t}\left(\rho_m h_m\right)+\nabla\cdot\left(\rho_m h_m \mathbf{v}_m\right)=\nabla\cdot\left(\lambda_m\nabla T\right)-\nabla\cdot\left[\rho_m\left(h_l-h_m\right)\left(\mathbf{v}_m-\mathbf{v}_{\text{cast}}\right)\right] \tag{3}$$

$$\begin{aligned}&\frac{\partial}{\partial t}\left(\rho_m C_m^i\right)+\nabla\cdot\left(\rho_m C_m^i \mathbf{v}_m\right)=\\&=\nabla\cdot\left(\rho_m f_l D_l^i\nabla C_m^i\right)+\nabla\cdot\left[\rho_m f_l D_l^i\nabla\left(C_l^i-C_m^i\right)\right]-\nabla\cdot\left[\rho_m\left(C_l^i-C_m^i\right)\left(\mathbf{v}_m-\mathbf{v}_{\text{cast}}\right)\right]\end{aligned} \tag{4}$$

The enthalpy of the mixture is defined as $h_m=f_l h_l+f_s h_s$, where the constitutive relations are given as $h_S(\mathrm{T})=\rho_s c_{ps} T$ and $h_l(T)=\rho_l\left[c_{ps}T+c_{pl}(T-T_{sol})\right]+h_M$, with h_M denoting melting enthalpy. We assume that f_l rises consistent with the Lever rule in the mushy zone, $T>T_{liq}\left(C_m^i\right): f_l=1,\quad C_l^i=C_m^i;\ T_{liq}\left(C_m^i\right)\geq T\geq T_{sol}\left(C_m^i\right): C_m^i=\left(1-f_l\right)k^i C_l^i,\ T=T_{liq}\left(C_l^i\right); T<T_{sol}\left(C_m^i\right): f_l=0$, $C_l^i=C_m^i/k^i$ and that the heat conduction is a weighted average $\lambda=f_l\lambda_l+f_s\lambda_s$. On the inlet boundary (north) of the axisymmetric domain we impose the fully developed Poisseuille flow profile with known casting temperature and nominal concentration of the alloying elements C_0^i:

$$\mathbf{v}_S=\mathbf{v}_L=-\hat{\mathbf{e}}_z v_{cast}\frac{2R^2}{R_{inlet}{}^2}\left(1-\frac{r^2}{R_{inlet}{}^2}\right),\quad T=T_{cast} \tag{5}$$

where R_{inlet} and R are the inlet and mould radii, respectively. On the top free surface we set the temperature to the casting temperature and concentration to the nominal alloy concentration, and for the velocity, we set $\partial v_r/\partial z=0, v_z=0$. On the left (symmetry) boundary, we prescribe $v_r=0,\ \partial v_z/\partial r=0,\ \partial T/\partial r=0, C_l^i=C_0^i$. On the right boundary (cooling side), the sticking boundary condition in the moving system is $\mathbf{v}=-v_{\text{cast}}\hat{\mathbf{e}}_z$. The cooling side temperature boundary condition is modelled with the Robin boundary condition $-k\partial T/\partial r=h(z)\left(T-T_{env}\right)$, where $h(z)$ takes into account variation of the heat transfer coefficient with coordinate . It vanishes in the hot-top region. In the mould chill region we set it to a constant value h_{MC} whereas in the direct chill region it starts at value h_{DC} just below the mould chill and linearly rises to $3h_{DC}$ on the bottom of the direct chill region. Finally, on the bottom boundary we require constant field derivatives in the casting direction $\partial T/\partial z=0,\ \partial v_z/\partial z=0,\ v_r=0,\ \partial C_l^i/\partial z=0$.

2.2 Electromagnetic field equations

The dynamical pair of Maxwell is:

$$\nabla\times\mathbf{E}=-\frac{\partial\mathbf{B}}{\partial t},\quad \nabla\times\mathbf{B}=\mu_0\mathbf{J}+\frac{1}{c^2}\frac{\partial\mathbf{E}}{\partial t},\quad \left(\mu_0=4\pi\times10^{-7}\,\mathrm{Vs/Am}\right) \tag{6}$$

where we assume $c=\infty$ and work in the quasi-static approximation. Electric $\mathbf{E}$ and magnetic $\mathbf{B}$ fields are both divergence-free, i.e., solenoidal. In addition, the Ohm's law in the

moving frame relates fields with induced eddy currents, $\mathbf{J}=\sigma\left(\mathbf{E}+\mathbf{v}\times\mathbf{B}\right)+\mathbf{J}_{ext}$, where σ is conductivity that is nonzero only in the conductor, whereas $\mathbf{J}_{ext}$ is nonzero only in the coil and can be viewed as a boundary condition. We work with the electromagnetic potential $\mathbf{A}$ defined through $\mathbf{E}=-\nabla\phi-\partial\mathbf{A}/\partial t$, $\mathbf{B}=\nabla\times\mathbf{A}$ and impose the Coulomb gauge, $\nabla\cdot\mathbf{A}=0$, for the vector potential. Maxwell equations thus reduce to induction equation for $\mathbf{A}$:

$$\nabla^2\mathbf{A}=\mu_0\sigma\left(\frac{\partial\mathbf{A}}{\partial t}-\mathbf{v}\times\nabla\times\mathbf{A}\right)-\mu_0\mathbf{J}_{\text{ext}},\ \ \phi=0 \tag{7}$$

We set $\mathbf{v}=0$ on the right-hand side since the effect of velocity on the EM field is small in regime of small magnetic Reynolds number. Details of equations in axisymmetry are given in [8].

2.3 Solid mechanics equations

The solid mechanics model is solved by assuming the calculated temperature field from the thermal model. The assumed thermoelastic governing equation is:

$$G\nabla^2\mathbf{u}+(G+\lambda)\nabla\nabla\cdot\mathbf{u}+\nabla\lambda\nabla\cdot\mathbf{u}+\nabla G\left(\nabla\mathbf{u}+\left(\nabla\mathbf{u}\right)^T\right)=\nabla\left(\beta_T\left(T-T_{ref}\right)\right)-\mathbf{f} \tag{8}$$

where G stands for shear modulus, λ for Lamé parameter and $\mathbf{f}$ is the body force. The coupling with the temperature field is described by the coefficient β defined as $\beta=(3\lambda+2G)\alpha$ which is composition dependent. Another free parameter is the reference temperature T_{ref}, which denotes the temperature at which the thermal expansion is considered to be zero.

3 MICROSCOPIC MODEL

The microscopic model is solved in a post-processing mode. The model inputs temperature and concentration data from the macroscopic model. The representative input data are taken from the billet center, middle and surface streamlines in Lagrangean sense. The time for microscopic model is calculated from the velocity and length of the streamline. A typical 3x3mm representative traveling area is considered on each of three considered streamlines.

3.1 Nucleation kinetics

To describe the nucleation event the continuous Log-normal Nucleation model is adopted. The variation of the grain distribution in each time step depends on the undercooling temperature in the bulk of the casting. The density of grains $n\left(\Delta T_{nuc}\right)$ at the given degree of undercooling ΔT_{nuc} is given by the integral of nucleation density distribution:

$$\frac{dn}{d\Delta T_{nuc}}=\frac{n_{\max}}{\sqrt{2\pi}\Delta T_\sigma}\frac{1}{\Delta T_{nuc}}\exp\left[-\frac{1}{2}\left(\frac{\ln\Delta T_{nuc}-\ln\Delta T_0}{\Delta T_\sigma}\right)^2\right] \tag{9}$$

where ΔT_0, ΔT_σ, $n_{\max}$ represent the mean nucleation undercooling, the standard deviation and the maximum density of nuclei that can form in the melt, respectively. These three

standard nucleation parameters are determined experimentally for each alloy and grain refiner by measuring the grain size.

3.2 Grain growth kinetics

A modified KGT model is used to predict the grain growth velocities. The growth process is driven [11] by the local undercooling, approximated from the linearized phase diagram of a multicomponent alloy $\Delta T = T_{liq} - T(t) + \sum_{i=1}^{n} m_i \left(C_l^i - C_0^i \right)$, where T_{liq}, T, m_i, C_l^i, C_0^i represent the liquidus temperature, the temperature recalculated from the macro model, the liquidus slope of species i, the liquid concentration of species recalculated from the macro model and the initial concentration of species i. In order to calculate the grain velocity V, the transport equations are solved for each species i. The solution relates the nondimensional supersaturations for each alloying element i, to the corresponding growth Péclet number $\Omega_i = F(P_i)$, $P_i = rV / 2D_i^l$ where r is a tip radius. The supersaturation Ω_i is expressed by the Ivantsov function for the purely diffusive growth regime $F(P_i) = \mathrm{P}_i \exp(\mathrm{P}_i) E_1(\mathrm{P}_i)$ where E_1 is the integral exponential function. The total undercooling can be expressed as:

$$\Delta T = \sum_{i=1}^{n} m_i C_0^i \left(1 - \frac{1}{1 - (1 - k_i) F(P_i)} \right) \tag{10}$$

where k_i is the partition coefficient of species i. The marginally stable plane front solution is applied in order to select the solution $V(\Delta T)$, with $r = V^{-1/2} \sqrt{4\pi^2 \Gamma / \sum_{i=1}^{n} m_i (k_i - 1) C_l^i / D_i^l}$, where Γ stands for the Gibbs-Thomson coefficient.

4 SOLUTION PROCEDURES

The time-averaged Lorentz force is calculated by discretizing the calculation domain in local domains with 9 nodes with the approximation function expressed as a minimum least squares fit of 6 monomials $\left(1, x, y, xy, x^2, y^2\right)$. The weight function in the least squares problem is chosen to be Gaussian, $e^{-|\mathbf{r}|^2/\xi}$. ξ is an average squared distance from the center node to all nodes in the local domain. The sparse linear system is solved using the sparse system solver [12]. The EM field calculation domain extends $1.5H$ beyond the casting system in vertical direction and is terminated at $3R$ in the radial direction. Unstructured node arrangement is used for the EM field calculation that has embedded the nodes for the thermofluid equations in a way that in the mould the EM and thermofluid nodes coincide. The transport equations for mass, momentum, energy, and concentration are solved by collocation with multiquadrics radial basis functions with 5-noded subdomains in case of thermal model (as described in [13, 14]) and 6-noded subdomains in case of mechanical model as elaborated in the present proceedings [15].

The phase change kinetics equations are solved by the PA method. The representative microscopic calculated domain with a dimension 3x3 mm is discretized into 360 000 randomly located nodes. The post processing calculations of temperature and concentration

for this domain are taken from the macro level of calculations. The details of the PA method are elaborated in [16]. The potential nucleation sites are randomly located in this domain in each time step. The newly arisen site represents a new family and gets a randomly chosen color at the time when it appears. As grains nucleate, they start to grow with respect to the PA neighborhood configuration which is associated with the position of the neighboring PA nodes which fall into a circle with assumed radius $R_h = 0.5\,\mu\text{m}$. Due to random arrangement of the nodes each grain gets different neighboring configuration. For the given values of the physical and thermodynamic properties m_i, k_i, D_l^i of each alloy element i, the variable grain growth velocity V for a specified undercooling temperature ΔT can be calculated. Independent calculations of the growth velocity for a given set of these values are carried out in each time step by using an iterative method. This requires to provide the micro model with values of T and C_l^i, taken from the macro model at every 0.05 s.

5 NUMERICAL EXAMPLES

Six cases of aluminium (Al-Cu-5.25%) alloy were simulated, where the majority of the casting parameters remained constant. These parameters are shown in Table 1. Other parameters (casting speed, EM amplitude, EM frequency) were varied in order to study their effect on the shape and dimension of the liquid sump. Overview of calculated cases is given in Table 2.

Table 1: Casting parameters.

Casting temperature	707 °C
Cooling water flow	140 m^3/h
Liquidus temperature	643 °C
Solidus temperature	562 °C
Liquid thermal conductivity	88.1 W/mK
Solid thermal conductivity	180 W/mK
Liquid specific heat	1160 J/kgK
Solid specific heat	1270 J/kgK
Permeability constant	$6.7 \cdot 10^{-11}$ m^2
Latent heat	$3.0 \cdot 10^5$ J/kg
Billets on a table	20
Number of coil turns	48
Coil width	20 mm
Coil height	80.97 mm
Coil inner radius	187.5 mm
Coil lowest point	-101.87

Table 2: Cases descriptions.

Case number	Casting speed	Amplitude	Frequency
Case 1	60 mm/min	0	0
Case 2	80 mm/min	0	0
Case 3	100 mm/min	0	0
Case 4	80 mm/min	10 A	10 Hz
Case 5	80 mm/min	10 A	15 Hz
Case 6	80 mm/min	10 A	20 Hz

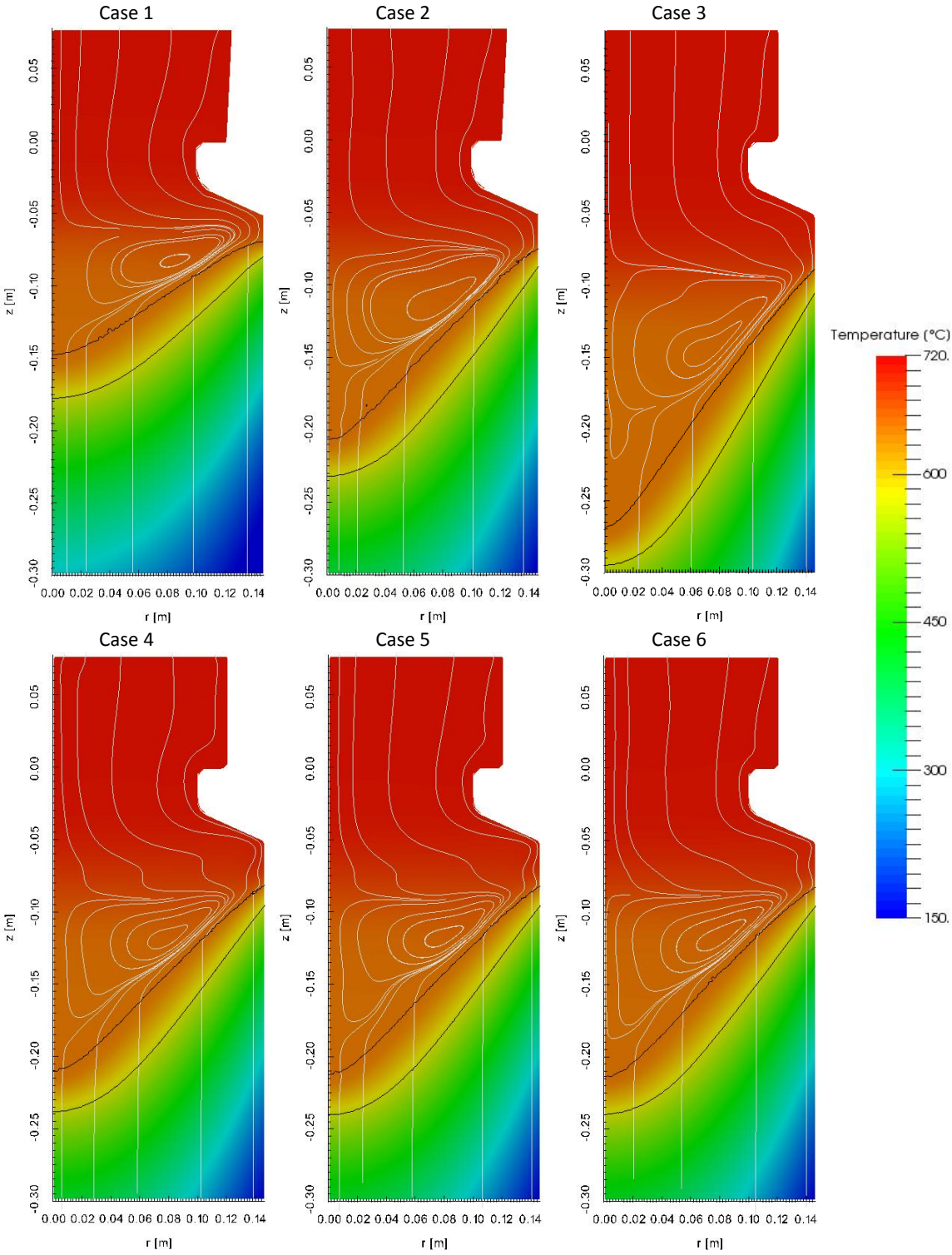


Figure 1: Temperature distribution and streamlines are shown for cases 1 - 6.

The effect of the casting speed is evident from the comparison of results of cases 1, 2 and 3, where the casting speed has values of 60, 80 and 100 mm/min respectively. It is obvious that an increase in velocity will result in deeper billet sump, which is in accordance with general observations in DC casting. The depth of the sump (solidus line at the center) increases as 18 cm, 23 cm, 29.5 cm in cases 1, 2 and 3 respectively. Furthermore, the thermo-solutal eddy in the central portion of the billet is also increasing with the casting speed.

Case 2 (no EM field) and cases 4, 5 and 6 (with EM field) can be compared to observe the effect of the EM field on the DC casting, as all cases were calculated for the same non EM field related casting parameters. Even though the sump depth is slightly smaller in the case of EM field, larger differences can be observed in the velocity field. The eddy formed in the sump due to natural convection is smaller in case 5. The streamlines are shallower, as the EM force amplifies the velocity in the horizontal direction.

As the frequency is relatively small (low frequency EM casting) in all three simulated cases accounting the EM field, the differences in the sump shape are barely visible. Nevertheless, the effect of electromagnetic frequency variations can be noticed from the streamlines, especially in the zone above the central eddy. If the EM frequency is increased, the streamlines of the inlet melt flow become more parallel with billet axes. Furthermore, higher frequency causes flatter central eddy.

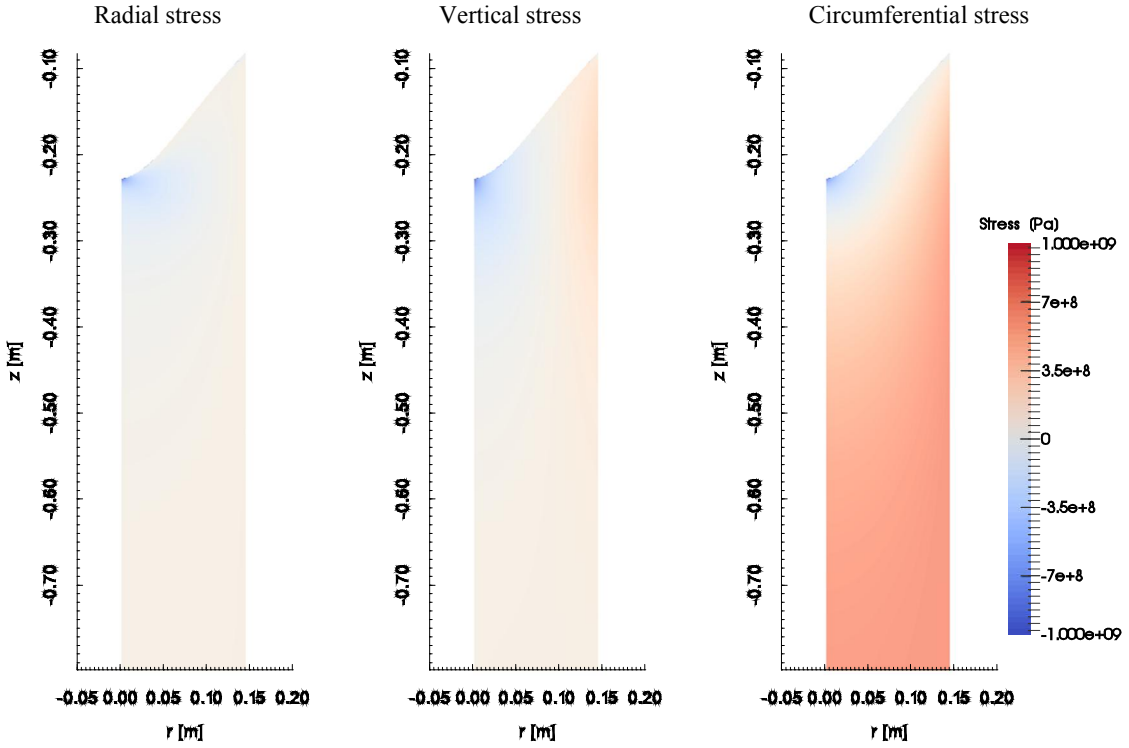


Figure 2: Stress calculations from the thermoelastic model for case 4.

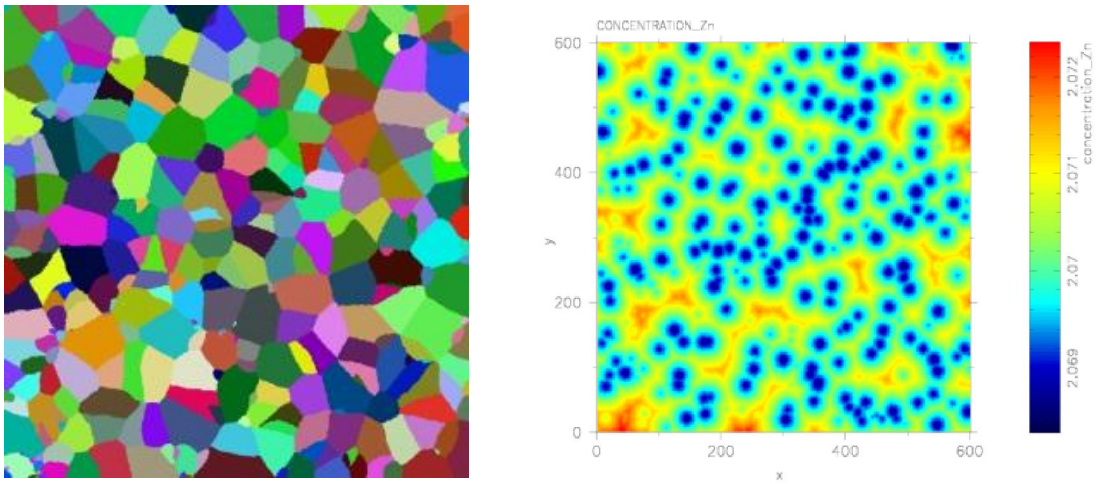


Figure 3: Results from the grain structure model: structure and Zn concentration for Al-Zn--5.35% Mg-2.35% Cr-1.5% Cu-0.5% alloy.

6 CONCLUSIONS

A reasonably complete coupled multiphysics and multiscale model of the low frequency electromagnetic casting has been developed. First the electromagnetic field is calculated, followed by the calculations of the momentum equation, pressure velocity-coupling, the energy equation, and species equations, respectively. The mechanical calculations are performed without feedback, based on the thermal calculations. The microstructure calculations are coupled in Lagrangean sense to the calculated velocity and concentration fields in a post-processing mode. The models are on the macro level solved by the diffuse approximate method and on the micro-level by the point automata method. This novel approach is completely meshless and no polygonisation is needed. The considered physical models incorporate only simple, basic elements that obviously need further improvements. In the field of macrosegregation, the movement of the solid phase and influence of the grain refiner represent the next necessary steps [17]. In the field of solid mechanics, the introduction of viscoelastic behavior and feedback to the thermal model (mould heat transfer) seems important. In the field of microstructure modelling, the inclusion of the secondary phases is necessary in the model. It should be pointed out that the extension of the present numerical approach to three dimensions seems straightforward, since the coding of the present axisymmetric meshless formulation is equivalent in higher dimensions. The capabilities of the model were demonstrated in figures 1, 2 and 3.

ACKNOWLEDGEMENTS

Support from the Slovenian Grant Agency in the framework of the project L2-6775 and young researchers grants, Slovenian Ministry for Economy in the framework of the Voucher projects and IMPOL Aluminium Industry is kindly acknowledged.

REFERENCES

[1] Vertnik, R., Založnik, M. and Šarler, B. *Eng.Anal.Bound.Elem.* (2006) **30** 847-55.
[2] Šarler, B. and Vertnik, R. *Comput. Math. Appl.* (2006) **51** 1269-82.
[3] Chorin, A. J. *J. Comput. Phys*. (1967) **2** 12-26.
[4] Lorbiecka, A. Z. and Šarler, B. *Materials Science Forum* (2014) **790-791** 115-20.
[5] Eskin, D. G. *Physical Metallurgy of Direct Chill Casting of Aluminium Alloys*. Boca Raton: CRC Press/Taylor & Francis (2008).
[6] Davidson, P. A. *An Introduction to Magnetohydrodynamics*. Cambridge: Cambridge University Press (2001).
[7] Dong, J. and Cui, J. Z. *Metall. Mater. Trans*. (2004) **35A** 2487-95.
[8] Košnik, N., Vertnik, R. and Šarler, B. *Materials Science Forum* (2004) **790-791** 390-395.
[9] Bennon, W. D. and Incropera, F. P. *Numer. Heat Transf.* (1988) **13A** 277-96.
[10] Du, Q., Eskin, D. G. and Katgerman, L. *Metall. Mater. Trans.* (2007) **38A** 180-89.
[11] Appolaire, B., Combeau, H. and Lesoult, G. *Mater. Sci. Eng. A* (2007) **487** 33-45.
[12] Schenk, O. and Gärtner, K. *Journal of Future Generation Computer Systems* (2004) **20** 475-487.
[13] Kosec, G., Založnik, M., Šarler, B. and Combeau, H. *Comput. Mater. Cont.* (2011) **22** 169-59.
[14] Kosec, G. and Šarler, B. *Eng. Anal. Bound. Elem.* (2014) **45** 36-44.
[15] Mavrič, B. and Šarler, B. *ICASP-4 Proceedings* (2014) (to appear)
[16] Lorbiecka, A. Z. and Šarler, B. *CMC* (2010) **18** 69-104.
[17] Nadella, R., Eskin, D. G., Du, Q. and Katgerman, L. *Progress in Materials Science* (2008) **53** 421-480.

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis(Eds)

CONTACT TYPES HIERARCHY AND ITS OBJECT-ORIENTED IMPLEMENTATION

IVAN I. KOSENKO*, KIRILL V. GERASIMOV† AND MIKHAIL E. STAVROVSKIY‡

*Moscow State Technical University of Radio Engineering, Electronics, and Automation
78 Vernadsky Avenue, 119454, Moscow, Russia
e-mail: kosenko@mirea.ru, web page: https://fcyb.mirea.ru/kafedri/internal/tm.html

†Faculty of Mechanics and Mathematics, Lomonosov Moscow State University
1 Leninskiye Gory, Main Building, GSP-1, 119991, Moscow, Russia
e-mail: kiriger@gmail.com

‡STAN Group
1 Vadkovsky lane, 127994, Moscow, Russia
e-mail: stavrov@list.ru, web page: http://www.stan-group.com/

Key words: Object-Oriented Modeling; Mechanical Constraint; Contact Model Template; Omni Wheel; Contact Tracking.

Abstract. Technology of the object-oriented implementation for the multibody dynamics models is the key feature when developing the corresponding computer structures. We are based on an approach originating from concepts explained earlier. Following the guidelines outlined there one can develop the family of the constraint abstractions being adapted to any type of the machinery applications and relatively easily implement corresponding family of Modelica models. One also can reorder these classes hierarchically using sequences of the behaviour inheritance. Solutions concerning contact problems and corresponding examples are under consideration.

1 INTRODUCTION

Development of a computer model for the multibody system (MBS) dynamics is connected usually with use of a unified technology of any type for constructing models in an efficient way. Object-oriented languages [1] are known to resolve such a problem successively step by step using their natural features. On the other hand one of the natural way for representing the MBS dynamics is the so-called multiport representation of the models initially based on the bond graph application [2]. This latter approach is based on the idea of energy interchange, and substantially on energy conservation for physically

interconnected subsystems of any engineering type. Consider in the sequel a technology for constructing a model of MBS dynamics with constraints of any specific type in a unified way. Note that the unilateral constraints, like ones of mechanical contact may also be included in the further consideration process.

A lot of methods to describe the structure of the MBS using different graph approaches is known, see for instance [3]. Consider the MBS consisting of $m+1$ bodies $B_0, \ldots, B_m$. Represent it as a set $\mathcal{B} = \{B_0, \ldots, B_m\}$. Here B_0 is assumed to be a base body. We suppose B_0 to be connected with an inertial frame of reference, or to have a known motion with respect to (w. r. t.) the inertial frame of reference. For example one can imagine the base body as a rotating platform, or as a vehicle performing its motion according to a given law. For definiteness and simplicity we suppose in the sequel all state variables describing the rigid bodies motion always refer to one fixed inertial coordinate system connected to the base body by default.

Some bodies are considered as connected by mechanical constraints. Suppose all constraints compose the set $\mathcal{C} = \{C_1, \ldots, C_n\}$. We include in our considerations constraints of the following types: holonomic/nonholonomic, scleronomic/rheonomic.

Thus one can uniquely represent a structure of the MBS via a undirected graph $G = (\mathcal{B}, \mathcal{C}, \mathcal{I})$. Here $\mathcal{I} \subset \mathcal{C} \times \mathcal{B}$ is an incidence relation setting in a correspondence the vertex incident to every edge $C_i \in \mathcal{C}$ of the graph. According to physical reasons it is easy to see that for any mechanical constraint C_i there exist exactly two bodies $B_k, B_l \in \mathcal{B}$ connected by this constraint.

2 BASE CLASSES FOR THE MULTIBODY DYNAMICS MODELS

It is clear that consideration of the graph G provides only a simplest structural information insufficient for the MBS dynamical description. Indeed, in addition to the force interaction represented usually by wrenches between bodies B_k, B_l through the constraint C_i there exist kinematic conditions specific for different kinds of constraints. Wrenches themselves can be represented in turn by constraint forces, reactions, and constraint torques couples. These forces and couples are connected by virtue of Newton's third law of dynamics.

Thus if the system of ODEs for translatory-rotary motion can be associated with the object of a model corresponding to rigid body, then the system of the algebraic equations in a natural way can be associated with the object of a model corresponding to constraint. Note that according to above consideration the set of algebraic equations comprises relations for constraint wrenches, and kinematic relations depending on the certain type of constraints.

Thus all the "population" of any MBS model is reduced to objects of two classes: *RigidBody* (objects $B_0, \ldots, B_m$), *Constraint* (objects $C_1, \ldots, C_n$). According to this approach simulation of the whole system behavior reduces to permanent information interaction between the objects of two considered types. Within the frame of Newton's laws of dynamics one can construct the MBS as a communicative network for this interaction. In this case

the objects of bodies "feel" the action of other ones through corresponding objects of constraints.

Physical interactions are conducted in models due to objects splitted also in two classes of ports: *WrenchPort*, *KinematicPort*. The first one is to be used to transfer wrench. In addition, *WrenchPort* has to be used for transferring the information about current location of the point reaction acts upon.

In our idealized model the force interaction between bodies supposed exactly at a geometric point. Its coordinates are fed outside constraint object through *WrenchPort* permanently in time.

Now it is possible to describe an architecture of information interactions within the particular constraint C_i corresponding to an individual edge of graph G, see Fig. 1. Thus computer model of the particular constraint is represented by the communication network.

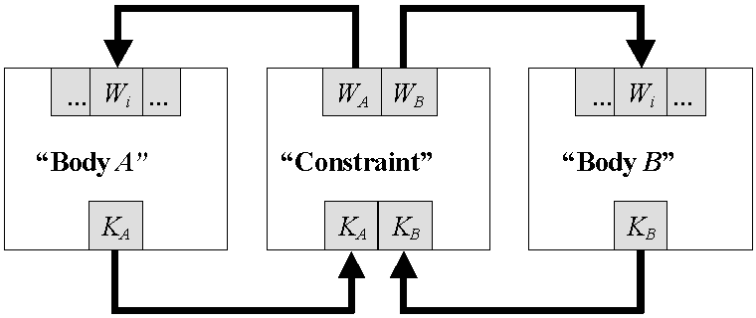


Figure 1: Architecture of constraint

KinematicPort is to be used to transfer the data of rigid body kinematics: configuration (position of center of mass, orientation), velocity (velocity of the center of mass, angular rate), and acceleration (acceleration of the center of mass, angular acceleration) containing in particular information about twist. When getting force information through ports $W_1, \ldots, W_s$ from the incident objects of class *Constraint* the object of class *RigidBody* simultaneously generates, due to an integrator, kinematic information being fed outside through the port K. On the other hand every object of class *Constraint* gets kinematic data from the objects corresponding to bodies connected by the constraint under consideration through its two "input" ports K_A, K_B. Simultaneously using the system of algebraic equations this object generates information concerning wrenches, and transmits the data to "output" ports W_A, W_B for the further transfer to objects of bodies under constraint.

For simplicity and clearness we apply now base classes from above for simulating the dynamics of MBSs with bilateral constraints [4]. Application of the components for the unilateral constraints [5, 6] we will see later.

In superclass *RigidBody* dynamics of rigid body is described here by means of Newton's differential equations for the body mass center, and by Euler's differential equations for the rotary motion. Note that to be able to have an invariant description of the rotary motion one can use an excellent tool: quaternion algebra $\mathbf{H}$. In this case we "lift" the configuration manifold from $SO(3)$ to $S_3 \subset \mathbf{H}$ and then implement dynamics of rotation in flat space $\mathbf{H} \cong \mathbf{R}^4$ taking into account that S_3 is an invariant manifold of the rotary dynamics redefined on $\mathbf{H}$. In this way we have only one flat chart $\mathbf{H}$ for the underlying due to double covering configuration manifold $SO(3)$, and need not in any special choices of the configuration angles or anything like that.

The double covering $S_3 \to SO(3)$, $\mathbf{q} \mapsto T$ mentioned is implemented inside the *RigidBody* class by the known formula. The rotation matrix T is fed outside the object through the *KinematicPort* permanently in time. The Euler equations are constructed using quaternion algebra in a way described in [4].

Remind that according to our technology of the constraint construction [4] two connected bodies are identified by convention with the letters A and B fixed for each body. All kinematic and dynamic variables and parameters concerned one of the bodies are equipped with the corresponding letter as a subscript.

All objects of the class *Constraint* must have classes-inheritors as subtypes of a corresponding superclass. According to Newton's third law this superclass must contain the equations of the form $\mathbf{F}_A + \mathbf{F}_B = \mathbf{0}$, $\mathbf{M}_A + \mathbf{M}_B = \mathbf{0}$ in its behavioral section. Here arrays $\mathbf{F}_A$, $\mathbf{M}_A$ and $\mathbf{F}_B$, $\mathbf{M}_B$ represent constraint forces and torques "acting in directions" of bodies A and B correspondingly. Kinematic equations for different types of constraints are to be added to the last equations in different classes-inheritors corresponding to these particular types of constraints.

3 JOINT MODEL AS ONE OF A BILATERAL CONSTRAINT

Class *Joint* plays a key role in the future model of a vehicle we will build. *Joint* is a model derived from the base class *Constraint*. Remind [5] that in order to make a complete definition of the constraint object behavior for the case of rigid bodies one has to compose a system of twelve algebraic equations w. r. t. twelve coordinates of vectors $\mathbf{F}_A$, $\mathbf{M}_A$, $\mathbf{F}_B$, $\mathbf{M}_B$ constituting the wrenches acting upon the connected bodies.

First six (**??**) always present in the base model *Constraint* due to Newton's third law. For definiteness suppose these six equations are used to express six components of $\mathbf{F}_B$, $\mathbf{M}_B$ depending on $\mathbf{F}_A$, $\mathbf{M}_A$. Thus six components of $\mathbf{F}_A$, $\mathbf{M}_A$ remain as unknowns. To determine them each constraint of rigid bodies need in six additional independent algebraic equations. These equations may include components of force and torque directly, or be derived from the kinematic relations corresponding to specific type of the constraint.

In the case of the joint constraint being under construction here let us represent the motion of the body B as a compound one including the body A convective motion w. r. t. an inertial frame of reference, and a relative motion w. r. t. the body A. An absolute motion is one of the body B w. r. t. the inertial system.

Define the joint constraint with help of the following parameters: (a) a unit vector $\mathbf{n}_A$ defining an axis of the joint in the body A; (b) a vector $\mathbf{r}_A$ fixed in the body A and defining a point which constantly stays on the axis of the joint; (c) a vector $\mathbf{r}_B$ fixed in the body B and defining a point which also constantly stays on the axis of the joint. The main task of the base joint class is always to keep geometric axes fixed in each of the bodies in coincidence.

First of all one has to compute the radii vectors of the points fixed in the bodies w. r. t. inertial system $\mathbf{R}_\alpha = \mathbf{r}_{O_\alpha} + T_\alpha \mathbf{r}_\alpha$ ($\alpha = A, B$), where [5] $\mathbf{r}_{O_\alpha}$ is the position of the α-th body center of mass, T_α is its current matrix of rotation. The joint axis has the following components $\mathbf{n}_{Ai} = T_A \mathbf{n}_A$ in the inertial frame of reference. According to the equation for relative velocity for the marked point of the body B defined by the position $\mathbf{R}_B$ we have $\mathbf{v}_{Ba} = \mathbf{v}_{Be} + \mathbf{v}_{Br}$, $\mathbf{v}_{Ba} = \mathbf{v}_{O_B} + [\boldsymbol{\omega}_B, T_B \mathbf{r}_B]$, $\mathbf{v}_{Be} = \mathbf{v}_{O_A} + [\boldsymbol{\omega}_A, \mathbf{R}_B - \mathbf{r}_{O_A}]$, where $\mathbf{v}_{Ba}$, $\mathbf{v}_{Be}$, $\mathbf{v}_{Br}$ are an absolute, convective, and relative velocities of the body B marked point, $\boldsymbol{\omega}_A$, $\boldsymbol{\omega}_B$ are the bodies angular velocities.

Furthermore, according to the computational experience of the dynamical problems simulation the precompiler work is more regular if the kinematic equations are expressed directly through accelerations. Indeed, otherwise the compiler tries to perform the formal differentiation of equations for the velocities when reducing an index of the total DAE system. Frequently this leads to the problems either in time of translation or when running the model.

Thus using the known Euler formulae for the rigid body kinematics and the Coriolis theorem we obtain an equations for the relative linear acceleration in the form

$$\begin{aligned} &\mathbf{a}_{Ba} = \mathbf{a}_{O_B} + [\boldsymbol{\varepsilon}_B, T_B \mathbf{r}_B] + [\boldsymbol{\omega}_B, [\boldsymbol{\omega}_B, T_B \mathbf{r}_B]], && \mathbf{a}_{Ba} = \mathbf{a}_{Be} + 2\,[\boldsymbol{\omega}_A, \mathbf{v}_{Br}] + \mathbf{a}_{Br}, \\ &\mathbf{a}_{Be} = \mathbf{a}_{O_A} + [\boldsymbol{\varepsilon}_A, \mathbf{R}_B - \mathbf{r}_{O_A}] + [\boldsymbol{\omega}_A, [\boldsymbol{\omega}_A, \mathbf{R}_B - \mathbf{r}_{O_A}]], && \mathbf{a}_{Br} = \mu \mathbf{n}_{Ai}, \end{aligned}$$

where $\mathbf{a}_{Ba}$, $\mathbf{a}_{Be}$, $\mathbf{a}_{Br}$ are an absolute, convective, and relative accelerations of the body B marked point, $\boldsymbol{\varepsilon}_A$, $\boldsymbol{\varepsilon}_B$ are the bodies angular accelerations, $\mathbf{v}_{Br}$ is a relative velocity of the body B marked point, $\boldsymbol{\omega}_A$, $\boldsymbol{\omega}_B$ are the bodies angular velocities.

We also need in an analytic representation of conditions that the only projections of the bodies angular velocities and accelerations having a differences are ones onto the joint axis. Corresponding equations have the form $\boldsymbol{\omega}_B = \boldsymbol{\omega}_A + \boldsymbol{\omega}_r$, $\boldsymbol{\varepsilon}_B = \boldsymbol{\varepsilon}_A + [\boldsymbol{\omega}_A, \boldsymbol{\omega}_r] + \boldsymbol{\varepsilon}_r$, $\boldsymbol{\varepsilon}_r = \lambda \mathbf{n}_{Ai}$, where $\boldsymbol{\omega}_r$, $\boldsymbol{\varepsilon}_r$ are the relative angular velocities and accelerations.

Besides the kinematic scalars μ, λ we will need in their reciprocal values $F = (\mathbf{F}_A, \mathbf{n}_{Ai})$, $M = (\mathbf{M}_A, \mathbf{n}_{Ai})$ correspondingly. Note that the class described above is a partial one (doesn't yet complete the constraint definition) and can be used to produce any imaginable model of the joint type constraint. To obtain a complete description of the joint model one has to add to the behavioral section exactly two equations. One of them is to define one of the values μ, F (translatory case). Other equation is intended to compute one of the values λ, M (rotary case).

4 CONTACT AS A PARTICULAR CASE OF MECHANICAL INTERCONNECTION

For simplicity we suppose here objects of unilateral constraint implement model of mechanical contact without impacts. Though in general according to the nature of unilateral constraint one can describe it using fundamental state variable which at any time instant can have one of three values: "Flight", "Sliding", "Rolling". Values enumerated have a sense transparent enough. State "Flight" means the constraint at a current time instant is in a disconnected condition i. e. bodies in fact are not connected and can perform free relative flying. As state variable has one of values "Sliding" or "Rolling" then bodies supposed to be in a contact. The difference is that the first state permits the relative slipping of the bodies but the second one does not.

The dynamics of a rigid body translatory–rotary motion was outlined above. However the mechanical constraint model representation undergoes an essential changes in compare with the bilateral case. We use the so called complementarity rules [7] as a base for the unified description of the unilateral constraint. Taking into account complementarity rules one can see easily that any constraint always is defined by the three scalar equations. To derive these equations first consider local geometry of the problem, see Figure 2.

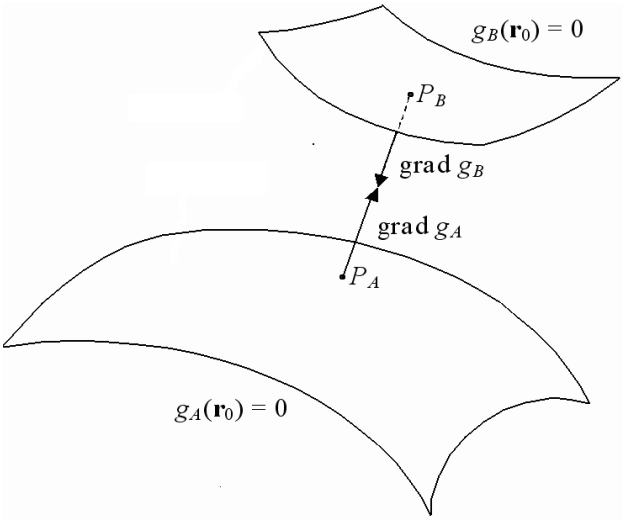


Figure 2: Area of Constraint

Outer surfaces supposed to be defined with respect to principal central axes of corresponding bodies by the equations $f_\alpha(\mathbf{r}_\alpha) = 0$ $(\alpha = A, B)$. Then in inertial frame of reference for the whole multibody system these equations will take the form $g_\alpha(\mathbf{r}_0) = f_\alpha[T^*_\alpha(\mathbf{r}_0 - \mathbf{r}_{O_\alpha})] = 0$ $(\alpha = A, B)$, with $\mathbf{r}_{O_A}$, $\mathbf{r}_{O_B}$ being a vectors of masscenters positions O_A, O_B for the bodies A and B, and T_A, T_B mean an orthogonal matrices for current bodies orientations. An asterisk denotes conjugating what equivalent to inverting of matrix for the case of orthogonality. Thus the functions $g_A(\mathbf{r}_0)$, $g_B(\mathbf{r}_0)$ depend upon the time

indirectly through the variables $\mathbf{r}_A$, $\mathbf{r}_B$, T_A, T_B.

Constraint object of our model is to compute at each current instant positions of the points P_A and P_B which are the nearest ones for interacting bodies A and B. By virtue of above assumptions such points are to be evaluated in a unique way. Denote the radii vectors of these points with respect to inertial frame of reference by $\mathbf{r}_{P_A}$, $\mathbf{r}_{P_B}$. Then using simple geometric considerations for the coordinates of the cited vectors one can derive the following system of algebraic equations

$$\begin{array}{rcl@{\qquad}rcl} \operatorname{grad} g_A(\mathbf{r}_{P_A}) & = & \lambda \cdot \operatorname{grad} g_B(\mathbf{r}_{P_B}), & g_A(\mathbf{r}_{P_A}) & = & 0, \\ \mathbf{r}_{P_A} - \mathbf{r}_{P_B} & = & \mu \cdot \operatorname{grad} g_B(\mathbf{r}_{P_B}), & g_B(\mathbf{r}_{P_B}) & = & 0. \end{array} \tag{1}$$

One can verify easily that the gradients can be computed by formulae $\operatorname{grad} g_\alpha(\mathbf{r}_{P_\alpha}) = T_\alpha \operatorname{grad} f_\alpha[T_\alpha^*(\mathbf{r}_{P_\alpha} - \mathbf{r}_{O_\alpha})]$, where $\alpha = A, B$. It easy to see the system (1) consists of eight scalar equations and has eight scalar unknowns: x_{P_A}, y_{P_A}, z_{P_A}, x_{P_B}, y_{P_B}, z_{P_B}, λ, μ. Variables λ, μ are an auxiliary ones. The equations (1) are in use either without or with a presence of the contact of bodies A, B. In a latter case the equation $\mu = 0$ instead of one of the surfaces equations is in use.

According to computational experience it is more reliable and convenient to use the equations of constraints in a differential form instead of those ones in the algebraic form (1). Such an approach is used frequently also when analyzing the properties of mechanical systems.

Normal vector $\mathbf{n}_A = \operatorname{grad} g_A / |\operatorname{grad} g_A|$ will play an important role in the further course. Normal for an outer surface of the body A is chosen here for definiteness. One can use the vector $\mathbf{n}_B$ as well.

Let us perform now a unified description of the unilateral constraint using kinematic and/or force equations. Denote by $\mathbf{F}_A$ the force acting on the body A from the body B. And by $\mathbf{F}_B$ denote the force acting on the body B from one of A vice versa. Each force mentioned acts at the point P_α, $\alpha = A, B$. In addition, let us introduce auxiliary notations $F_{An} = (\mathbf{F}_A, \mathbf{n}_A)$, $\mathbf{F}_{A\tau} = \mathbf{F}_A - F_{An}\mathbf{n}_A$, $\mathbf{v}_r = \mathbf{v}_{P_A} - \mathbf{v}_{P_B}$, $v_{rn} = (\mathbf{v}_r, \mathbf{n}_A)$, $\mathbf{v}_{r\tau} = \mathbf{v}_r - v_{rn}\mathbf{n}_A$.

If the bodies are disconnected and the constraint is in a state "Flight" then the force of reaction is equal to zero. Thus we have three scalar equations. For unifying the system of constraint equations further and for taking into account arbitrary directions of the normal $\mathbf{n}_A$ let us introduce auxiliary scalar variable κ in a way such that

$$F_{An} = 0, \quad \mathbf{F}_{A\tau} - \kappa \mathbf{n}_A = \mathbf{0}.$$

Actually one has obtained the system of four equation with four unknown variables F_{Ax}, F_{Ay}, F_{Az}, κ.

In case of bodies contact the condition $F_{An} = 0$ is substituted by the kinematic one $v_{An} = 0$. States "Sliding" and "Rolling" differ one from another using conditions in a tangent plane. Implementation of the Coulomb friction model supposed for the simplicity. Then one can obtain the vector force equation in the tangent plane

$$\mathbf{F}_{A\tau} - d \cdot F_{An}\mathbf{v}_{r\tau} / |\mathbf{v}_{r\tau}| - \kappa \mathbf{n}_A = \mathbf{0}, \tag{2}$$

where d is the coeffitient of friction.

For rolling the tangent velocity has to be zero

$$\mathbf{v}_{A\tau} - \kappa \mathbf{n}_A = \mathbf{0}.$$

In the case of sliding the model equation (2) "works" properly if the relative velocity isn't very small. However the problem of regularization for the equation of constraint (2) arises at the instance of transition from "Rolling" to "Sliding". It turns out that one can apply here the known approximation for Coulomb's friction using regularized expression for the tangent force

$$\mathbf{F}_{A\tau} - \kappa \mathbf{n}_A = d \begin{cases} F_{An}\mathbf{v}_{r\tau} / |\mathbf{v}_{r\tau}| & \text{as} \quad |\mathbf{v}_{r\tau}| > \delta, \\ F_{An}\mathbf{v}_{r\tau} / \delta & \text{as} \quad |\mathbf{v}_{r\tau}| \leq \delta, \end{cases}$$

where one supposes that $\delta \ll 1$.

It is known [8] that in this case the solution of the regularized problem remains close to the solution of the original one on asymptotically large time intervals. Implementation and further simulation show this closeness holds with the very high degree of accuracy. Such an approach resolves completely the problem of modeling for accurate transitions between states of "Sliding" and "Rolling".

Thus properties of the frictional contact as: (a) contact tracking algorithm; (b) contact velocities computation; (c) contact forces computation; (d) contact surfaces particular properties are combined in an object-oriented manner forming lines of inheritance in a natural way.

5 CLASS PARAMETRIZATION IN CASE OF COMPLIANT CONTACT

Implementation of the mechanical contact model with compliance assumes much more possibilities for the contact properties than in the rigid case. According to experience while developing the models for elastic contacting of rigid bodies interactions in the multibody dynamics a flexibility provided by an object-oriented approach can be used to utilize a wide variety of different properties concerning a contact of solids. The properties are mainly of the following categories:

(a) geometric properties for surfaces in vicinity of the contact patch (gradients of the functions defining surfaces, their Hesse matrices);

(b) a model to compute the contact area dimensions and normal elastic force;

(c) model for the normal viscous force of resistance;

(d) model for the tangent forces along the plane of the contact area.

A submodel of the geometry properties is to describe analytically algebraic surfaces of the structure complex enough. To implement the normal force computation one can choose from at least two approaches: the Hertz model and its volumetric modification. Force of viscous resistance also can be modeled in several different ways: linear, non-linear, etc. In the models for tangent forces one can adopt either "simplest" approaches based

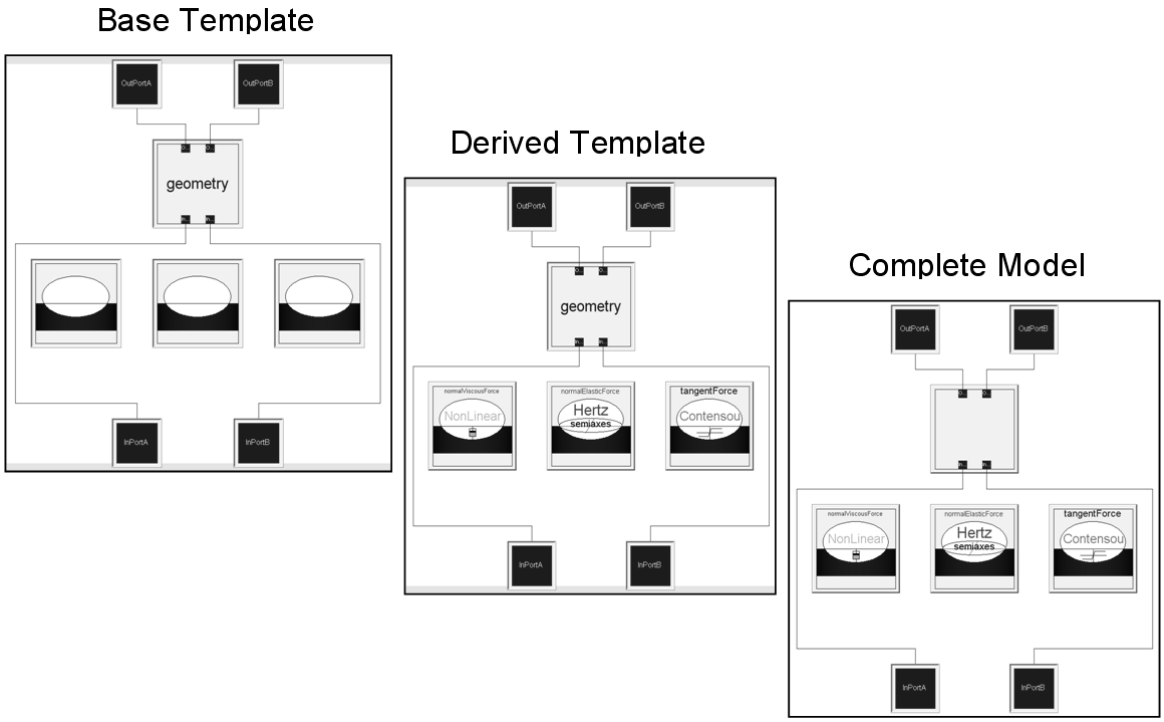


Figure 3: The model of mechanical contact by stages of inheritance.

on the Amontons–Coulomb friction or more complex ones represented by the Contensou–Erismann, and other models.

While developing a mechanical contact model architecture we used the base class *Constraint* described above as a starting point to construct its inheritor *ContactConstraintTemplate* being simultaneously a base class of new family of models to simulate mechanical contacts. Really this class is a base template represented as a container having four "sockets" to instantiate there the specific parameter classes of four types enumerated above, see its visual model in Figure 3 at a top left corner.

To develop complete model one can move along different ways. Class parametrization implemented in Modelica language for example is the facility in line to apply to the problem under description. In our case we have four class parameters corresponding to the submodel categories enumerated above. An example to construct specific contact interaction model see in Figure 3. The example includes two stages of inheritance:

1. to derive a template with the forces models, namely: the Hertz model for normal force, non-linear viscous force, the Contensou–Erismann model for the dry friction forces (to "fill in" three sockets in the middle of the base template visual model, see the derived template visual model at a central position of the Figure 3);

2. to complete the whole construct one should define a specific geometry submodel for the surfaces in contact (to "seal" the socket for geometry properties, see the complete visual model at a bottom right corner of the Figure 3).

On all the stages of inheritance the templates considered have an internal information interconnections between the submodels to be instantiated. These interconnections are implemented via the set of equations hidden behind the visual models and can vary for different models requiring different variables for the algorithms to compute normal and tangent forces of the complete model. So the whole picture remind us known construct of a card with the sockets and the interconnection wiring in its internal layers as a base template, and a chips to be instantiated in the sockets as a models of four types from above. With one exclusion: we have the derived template playing a role of additional card with its own additional wiring servicing already instantiated models "covering the card" of the base template.

One can remark finally an approach under presentation allows us to create and to change fast enough different types of an elastic contact models while developing the multibody dynamics systems simulators.

6 EXAMPLE OF THE OMNI VEHICLE DYNAMICAL MODEL

Investigation of omni vehicle dynamical properties is sufficiently popular topic in field of the multibody dynamics [9, 10, 11, 12]. The omni vehicle is one having omni wheels, wheels equipped by rollers along the rim. Simplified, idealized models having contacting rollers as an infinitely small discrete elements are known. Thus one has a resulting nonholonomic constraint being "uniformly distributed" over the wheel rim.

Our goal here is to develop a technique for building up a dynamical prototype for the "real" model of the omni vehicle explicitly involving dynamics of physical rollers. Here we rely upon the "simple" 3D multibody dynamics library classes from above which was utilized previously in several examples of the multibody systems dynamics [13]. Simultaneously this library enables us to create complex dynamical models including unilateral constraints of different nature.

We will pay here main attention to the process of the omni vehicle dynamical model development. Once again for simplicity we apply for the mechanical contact model the rigid point-contact one. This model has three structural levels of complexity: level of the roller; level of the omni-wheel; one of the whole vehicle, see Figure 4).

All the classes used in the vehicle model were mentioned above. In addition, one has to note that the contact tracking algoritm is possible to be arranged extremely simple and effective for the case of the wheel vertically aligned.

In the model of Figure 4 one has a vehicle with three omni-wheels. Each wheel has four rollers. Computational experiments were performed for different numbers of rollers. In addition, computational comparisons were performed for almost limit cases of simplified models with rollers inertial properties almost vanishing. The verification process was completed successfully.

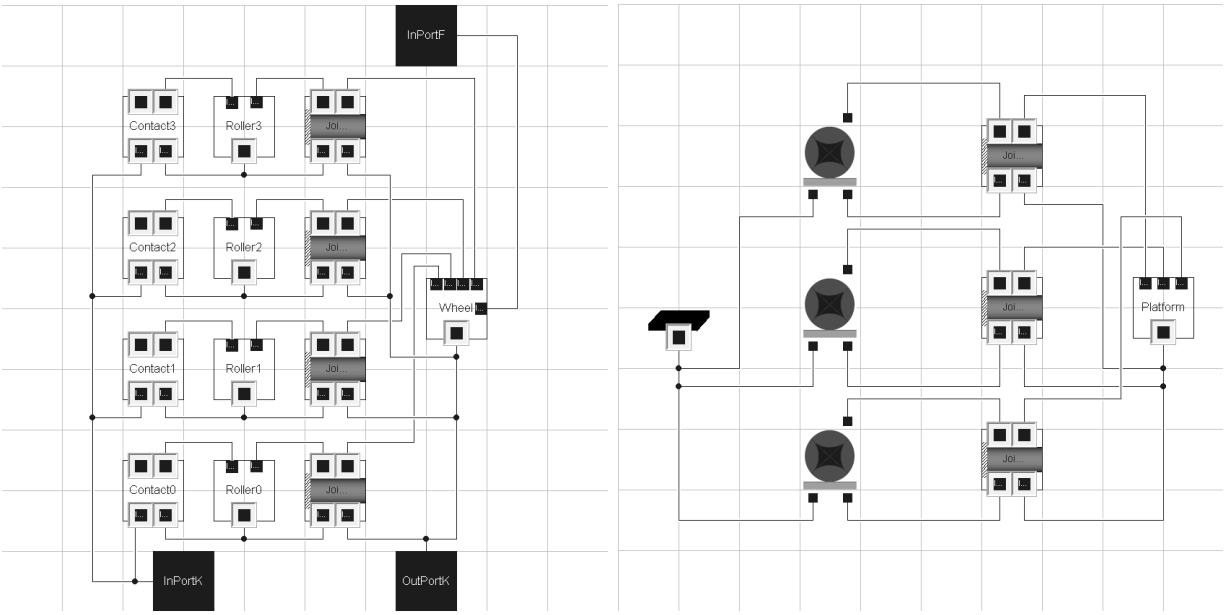


Figure 4: Left: The omni wheel visual model. Right: The omni vehicle visual model.

7 CONCLUSIONS

-The process of the models development and debugging becomes fairly easy and simple if one uses physically-oriented approach for the MBS dynamics simulation.

-An acausal modeling accelerates the model development releasing an engineer from the problem of causality assignment if s/he takes into account some requirements like complementarity rules.

-An object-oriented representation makes it possible to develop the constraints models adopted to the specific types of the bodies interconnections in a fast and effective manner.

-The bond graph theory guidelines are useful for the MBS model building process to create consistent resulting DAE system.

-Introducing the compliance into the model may be useful and effective preserving the principal properties of the MBS like anholonomity etc.

-The most natural and effective way to implement computer model of the mechanical contact is use of the templates with the object-oriented class parametrization.

-There exist a possibility for smooth impactless switching between rollers at contact upon rolling of omni wheel.

-Efficient and simplified contact tracking algorithm was implemented.

-Influence of friction model on dynamics of the omni vehicle was analyzed.

8 ACKNOWLEDGEMENT

The investigation was fulfilled under financial support provided by RSF, project 14-21-00068.

REFERENCES

[1] Fritzson, P. *Principles of object–oriented modeling and simulation with Modelica 2.1.* IEEE Press. (2004).

[2] Paynter, H. M. *Analysis and design of engineering systems.* The M. I. T. Press. (1961).

[3] Wittenburg, J. *Dynamics of systems of rigid bodies.* B. G. Teubner. (1977).

[4] Kossenko, I. I., Stavrovskaia, M. S. How one can simulate dynamics of rolling bodies via Dymola: approach to model multibody system dynamics using Modelica. *Proceedings of the 3rd International Modelica Conference*, Linköpings universitet, Linköping, Sweden, November 3–4. (2003) 299309.

[5] Kossenko, I. I. Implementation of unilateral multibody dynamics on Modelica. *Proceedings of the 4th International Modelica Conference*, Hamburg University of Technology, Hamburg–Harburg, Germany, March 7–8. (2005) 1323.

[6] Kosenko, I. Implementation of unilateral constraint model for multibody systems dynamics on Modelica language. *Proceedings of ACMD2006, The Third Asian Conference on Multibody Dynamics 2006*, Institute of Industrial Science, The University of Tokyo, Tokyo, Japan, August 1–4. (2006) 8pp.

[7] Pfeiffer, F. Unilateral multibody dynamics. *Meccanica*. (1999) **34**:437–451.

[8] Novozhilov, I. V. *Fractional Analysis : Methods of Motion Decomposition*, Birkhauser. (1997).

[9] Campion, G., Bastin, G., d'Andréa-Novel, B. Structural Properties and Classification of Kinematic and Dynamic Models of Wheeled Mobile Robots. *IEEE Transactions on Robotics and Automation*, (1996) **12**:47–62.

[10] Kálmán, V. Controlled Braking for Omnidirectional Wheels. *International Journal of Control Science and Engineering*, (2013) **3**:48–57.

[11] Tobolár, J., Herrmann, F., Bünte, T. Object-oriented modelling and control of vehicles with omni-directional wheels. *Computational Mechanics 2009*. Hrad Nectiny, Czech Republic, November 9–11. (2009).

[12] Zobova, A. A., Tatarinov, Ya. V. The Dynamics of an Omni-Mobile Vehicle. *Journal of Applied Mathematics and Mechanics*, (2009) **73**:8–15.

[13] Kosenko, I. I., Loginova, M. S., Obraztsov, Ya. P., Stavrovskaya, M. S. Multibody Systems Dynamics: Modelica Implementation and Bond Graph Representation. *Proceedings of the 5th International Modelica Conference*, arsenal research, Vienna, Austria, September 4–5. (2006) 213–223.

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis (Eds)

FURTHER DEVELOPMENT OF THE THEORY OF MULTICOMPONENT DRY FRICTION

ALEXEY A. KIREENKOV*

* Institute for Problems in Mechanics of the Russian Academy of Sciences
Prospekt Vernadskogo 101, korp.1, 119526 Moscow, Russia
e-mail: kireenk@ipmnet.ru, kireenk@mail.ru

Key words: Coupled Problems, Combined Dry Friction.

Abstract. *It is proposed further development of the theory of multicomponent dry friction which consists in presenting a more convenient form for the problems of the dynamics of the earlier developed models with a reduced number of coefficients*.

1 INTRODUCTION

In recent years there has been a surge of interest in the dynamics of systems of solids with friction under conditions of the combined kinematics. Fact is that researchers in the field of dry friction has long been known that in case of combined of the kinematics, when the rubbed solids are participated, simultaneously, in the sliding, spinning and rolling motion, the use of the classical Coulomb's law is not correct, and the friction law is undergoing significant changes.

One of the first attempts to describe the relationship of friction and spinning in the case of non-point contact of moving solids was undertaken by Contensou Contensou got numerical dependence of the dry friction force from the ratio of the slip velocity to the linear velocity of spinning.

The principle new development of the theory was given by Zhuravlev in With the aid of transition of the coordinaton system origin, to the instantaneous center of velocities, he obtained exact analytical expressions of the resultant vector and friction torque for circular contact spot, on the assumption that the distribution of contact pressure in the contact area subject to the law of Hertz. To use the obtained relationships in the dynamics problems, Zhuravlev built their fractional-linear but Pade. The convenience of the use of the Pade approximations making it possible to describe the effects of the combined dry friction for the entire range of angular and linear velocities has led in subsequent to the development of principally new models of friction on their basis [3-10].

Shortly after the publication of [2] in Russia and some European countries, several research groups focus their attention on studying the effects of dry friction in the combined kinematics. Main subjects of their research were the construction of mathematical models of dry friction, explaining the conditions of equilibrium of solids with dry friction and solution of various problems of classical dynamics. The narrow field of theoretical mechanics that began in the early 2000's with a few publications by Zhuravlev [2], Ivanov [11],

Kireenkov [12], Leine [13] and Karapetyan [14] are formed in integral scientific direction. Dozens publications on the subject were published.

Starting with the work [12] author of this publication began to develop a theory of the multi-component dry friction, one of the directions of it consist in the construction of the phenomenological dry friction models which are suitable for using in differential equations of motion.

The main distinguishing feature of this approach is that, at first, under the assumption of validity both the classical Coulomb's law in differential form for small element of the area inside the contact area and its generalized forms, there are constructed the exact coupled integral dry friction model, obtained by integrating the differentials of the principal vector and torque on the contact spot.

It is worth explain the used of the term "Exact integral model" because any model can not be exact, because it is only an approximation to a real phenomenon. This notion is used in the sense that, after the initial assumptions about the validity of Coulomb's law in classical and generalized differential form and general properties of the normal contact distributions inside of contact spot, all other computations, from a mathematical point of view, are being made exactly, without the use of approximate methods. Thus, after writing expressions for the differentials of the dry friction principal vector and torque, all subsequent transformations are exact results, reflecting the nature of the phenomenon.

The integral model gives a good description of the dry friction effects in the case of combined kinematics, but is inconvenient to be used in problems of dynamics, because it is required to calculate multiple integrals in the right-hand sides of the equations of motion. Previously, to escape this difficult procedure, the exact integral models are replaced by approximate models based on the Pade expansions of the first or second order. These replacements substantially simplify the combined dry friction modeling, making the calculation of double integrals over the contact area unnecessary. But in the case of arbitrary (in sign) velocities spinning ω and sliding v, the approximated models contain non-smooth functions (modules of velocities ω and v) in the denominators of the corresponded Pade expansions. A new type of approximated models which are the ratio of the linear form to square root of the quadratic form makes it possible to avoid this inconvenience.

2 EXACT INTEGRAL MODELS IN THE CASE OF COMBINED KINEMATICS

The dry friction exact integral models in the case of simultaneously sliding, spinning and rolling are constructed for circular contact sites under the assumption that the Coulomb law in generalized differential form holds for the small surface element dS in the interior of the contact spot, according to which the differentials of the resultant vector $d\mathbf{F}$ and the moment of friction dM_C with respect to the contact spot center are determined by the formulas [8]:

$$d\mathbf{F} = -f\sigma\frac{\mathbf{V}}{|\mathbf{V}|}\left(1+\mu_1|\mathbf{V}|^3-\mu_2|\mathbf{V}|\right)dS,\ dM_C = -f\sigma\frac{\mathbf{r}\times\mathbf{V}}{|\mathbf{V}|}\left(1+\mu_1|\mathbf{V}|^3-\mu_2|\mathbf{V}|\right)dS,$$
$$\mathbf{V} = (v-\omega y, \omega x),\ \mathbf{r} = (x, y) \tag{1}$$

where f is the coefficient of friction, $\mathbf{r} = (x, y)$ is the position vector of an elemental area in the interior of the contact spot with respect to its center Fig. 1 (left), ω is the angular velocity of rotation of the contact spot center, but μ_1 and μ_2 are the coefficients which can be defined in practice from experiments.. Necessity of using of the generalized differential form of the Coulomb law is caused by the numerous experimental investigations [15].

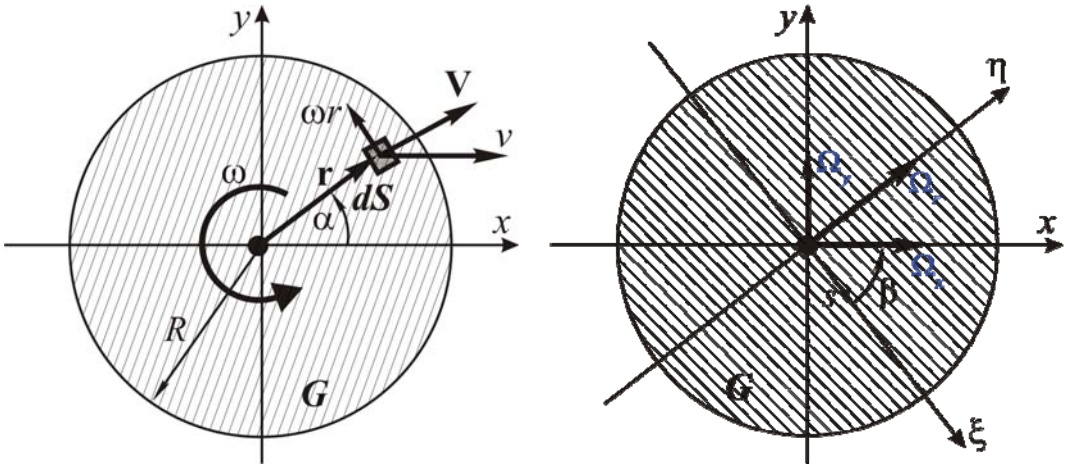


Figure 1. Kinematics inside the contact spot

In addition, in process of the exact integral models construction there are is used well known results from the theory of elasticity that tangent stresses lead to shift in the symmetric diagram of the normal contact stresses in the direction of the instantaneous sliding velocity v or in the rolling direction.

To use these results in the dynamics problems, it is proposed the simple asymptotic representations for the contact stresses distributions based on their general properties [6-7,16]:

$$\sigma(x, y) = \sigma_0 \left(1 + k_x x/R + k_y y/R\right) \tag{2}$$

where R - radius of contact spot and where $\sigma_0 = \sigma_0(r)$ - distribution of normal contact stresses at absence of motion having the properties of central symmetry

Presence of the simultaneously sliding and rolling in the different directions leads only to summarization of the corresponded coefficients. The influence of each of these effects can be investigated, due to linearity, independently. The only difference is that in case of rolling these coefficients depend on the direction and module of the rolling velocity Ω_r Fig 1 (right):

$$k_{xr} = \frac{k_r \Omega_y}{R\Omega_r}, k_{xr} = -\frac{k_r \Omega_x}{R\Omega_r}, \Omega_r = \sqrt{\Omega_x^2 + \Omega_y^2}, \ |k_r| \le 1, k_r \equiv 0 \text{ if } \Omega_r = 0 \tag{3}$$

To define the corresponded coefficients can be used or the results of theory of elasticity [6] or procedure developed in [6].

Thus, we have substantial approximation to the real situation in dependence on the general properties of the normal contact stresses distribution and real differential characteristics of the friction law.

Integration of the corresponded differentials over the contact spot yields the resultant vector $\mathbf{F}$ of the friction force and torque $\mathbf{M}_\mathbf{C}$:

$$\begin{aligned}
&\mathbf{F} = -f\iint\limits_G \sigma(x,y)\frac{\mathbf{V}}{|\mathbf{V}|}dxdy - f\iint\limits_G \sigma(x,y)\mathbf{V}\left(\mu_1\mathbf{V}^2 - \mu_2\right)dxdy \\
&\mathbf{M}_\mathbf{C} = -f\iint\limits_G \sigma(x,y)\frac{\mathbf{r}\times\mathbf{V}}{|\mathbf{V}|}dxdy - f\iint\limits_G \sigma(x,y)\mathbf{r}\times\mathbf{V}\left(\mu_1\mathbf{V}^2 - \mu_2\right)dxdy \\
&G = \{(x,y): x^2 + y^2 \le R^2\},\ \mathbf{F} = (F_{\|}, F_{\perp})
\end{aligned} \tag{4}$$

where $F_{\|}$ and $F_{\perp}$ denote the respective components of the resultant vector directed along the tangent and the normal to the trajectory of motion modulus of which has form:

$$\begin{aligned}
&F_{\|} = f\iint\limits_G \left(\frac{(v-\omega y)(1+k_y\, y/R)}{\sqrt{\omega^2(x^2+y^2)+v^2-2\omega v y}} + \mu_1 v^3 - \mu_2 v + 2\mu_1 v\omega^2(x^2+y^2)\right)\sigma_0 dxdy \\
&F_{\perp} = \frac{k_x f}{R}\iint\limits_G \frac{\omega x^2\sigma_0}{\sqrt{\omega^2(x^2+y^2)+v^2-2\omega v y}}dxdy,\ G = \{(x,y): x^2+y^2 \le R^2\} \\
&M_C = f\iint\limits_G \frac{(\omega(x^2+y^2) - v y)(1+k_y\, y/R)}{\sqrt{\omega^2(x^2+y^2)+v^2-2\omega v y}}\sigma_0 dxdy
\end{aligned} \tag{5}$$

Expressions (5) are calculated under supposition that influence of nonlinearity in the Coulomb law on the friction torque is neglegible and the coefficients k_x, k_y are the small parameters. These dependencies present the exact dry friction integral model in the case of combined kinematics.

It is convinient to present the double integrals in formulas (5) in the polar coordinates: $x = r\cos\alpha$, $y = r\sin\alpha$, $r \in [0,1]$, $\alpha \in [0, 2\pi]$ (Fig. 1) with origin in the contact spot center. In these coordinates, the polynomials terms in expressions for the friction force component $F_{\|}$ directed along the tangent to the trajectory of motion are the first moments of the normal contact stresses distribution of the first and third orders [7, 10].

3 APPROXIMATED DRY FRICTION MODELS

The exact integral model (5) gives a good description of the dry friction effects in the case of combined kinematics, but is inconvenient to be used in problems of dynamics, because it is required to calculate multiple integrals in the right-hand sides of the equations of motion. Previously, to escape this difficult procedure, the exact integral models are replaced by approximate models based on the Pade expansions of the first or second order. These replacements substantially simplify the combined dry friction modeling, making the calculation of double integrals over the contact area unnecessary. But in the case of arbitrary (in sign) velocities ω and v, the approximated models contain non-smooth functions (modules of velocities ω and v) in the denominators of the corresponded Pade expansions. A new type of approximated models which are the ratio of the linear form to

square root of the quadratic [16] form makes it possible to avoid this inconvenience. Possibility to use this kind of expansions was first mentioned in [9].

Procedures of the approximated model construction are based on the analytical properties of the double integrals (first terms in proposed integral model) as functions of the velocities $u = \omega R$ and v.

$$
\begin{aligned}
F_{\parallel}(u,v) &= fN\int_0^{2\pi}\int_0^1 \frac{(v-ur\sin\alpha)r(1+k_y r\sin\alpha)\sigma_0(r)}{\sqrt{u^2r^2+v^2-2uvr\sin\alpha}}drd\alpha \\
F_{\perp}(u,v) &= k_x fN\int_0^{2\pi}\int_0^1 \frac{ur^3\sigma_0(r)\cos^2\alpha}{\sqrt{u^2r^2+v^2-2uvr\sin\alpha}}drd\alpha \\
M_C(u,v) &= fRN\int_0^{2\pi}\int_0^1 \frac{(ur^2-vr\sin\alpha)r(1+k_y r\sin\alpha)\sigma_0(r)}{\sqrt{u^2r^2+v^2-2uvr\sin\alpha}}drd\alpha
\end{aligned}
\tag{6}
$$

These integrals can be considered independently due to additivity of the integral model (5) with the aid of formal zeroing of the parameters μ_1 and μ_2 - case correspondents to the classical Coulomb low in differential form. Using Coulomb's law in generalized differential form leads only to appearances of the additional polynomial terms.

One of the main analytical properties the exact integral expressions (6) is that the absolute values of the friction force components and toque are the homogeneous functions of the variables u and v of zero order of homogeneity: $F_{\parallel}(\lambda u,\lambda v) = \lambda^0 F_{\parallel}(\lambda u,\lambda v)$, $F_{\perp}(\lambda u,\lambda v) = \lambda^0 F_{\perp}(\lambda u,\lambda v)$, $M_C(\lambda u,\lambda v) = \lambda^0 M_C(\lambda u,\lambda v)$ Consequently, their approximations have to be the homogeneous functions of the variables u and v of zero order of homogeneity. This fact significantly reduces the possible type of approximations, coefficients of which are defined from the behavior $F_{\parallel}$, $F_{\perp}$ and M_C as functions variables $\{u,v\}$ as well as the behavior of their first derivatives at zero and at infinity . In result, the approximated analytical dry friction model in case of using of the Coulomb law classical differential form is

$$
M_C = \frac{M_0(u+k_y m_1 v)}{\sqrt{u^2+mv^2}},\quad F_{\parallel} = \frac{F_0(v+k_y a_1 u)}{\sqrt{v^2+au^2}},\quad F_{\perp} = \frac{k_x b_1 u}{\sqrt{(u^2+bv^2)}}
\tag{7}
$$

and in case of using of the Coulomb law in generalized differential form is

$$
\begin{aligned}
F_{\parallel} &= \frac{F_0(v+k_y a_1 u)}{\sqrt{v^2+au^2}} + 2\pi F_0\left((\mu_1 v^3-\mu_2 v)I_1 + 2\mu_1 v u^2 I_3\right),\ F_{\perp} = \frac{k_x b_1 u}{\sqrt{(u^2+bv^2)}} \\
M_C &= \frac{M_0(u+k_y m_1 v)}{\sqrt{u^2+mv^2}},\quad I_1 = \int_0^1 r\sigma_0(r)dr,\ I_3 = \int_0^1 r^3\sigma_0(r)dr
\end{aligned}
\tag{8}
$$

Coefficients of these models (7-8) can be calculated analytically [6-10] or numerically if the distribution of the normal contact stresses inside of contact spot is described by a priory known laws:

$$\frac{1}{m}=\left(\frac{v}{M_0}\frac{\partial M_C}{\partial u}\bigg|_{u=0}\right)^2,\ \frac{1}{a}=\left(\frac{u}{F_0}\frac{\partial F_{\|}}{\partial v}\bigg|_{v=0}\right),\ \frac{1}{b}=\left(\frac{v}{\mu F_0}\frac{\partial F_{\perp}}{\partial u}\bigg|_{u=0}\right)^2 \tag{9}$$

where functions $F_{\|}$, $F_{\perp}$ and M_C are defined by the formulas (6). In the other cases they can be estimated from the experiments [15].

The dynamics coupling of the dry friction models (5), (7-8) is defined by the coefficient k_x. If the external forces are absent then it can be calculated from the simple equation [10]: $F_{\|}h = Ns$, where h - distance from the center mass of the moving solids to the plane of sliding and s - the shifting of the center of gravity of the contact spot in the direction of sliding or rolling (Fig. 1). In result, the approximated analytical dry friction model has form:

$$\begin{gathered} F_{\|} = F_0\left(\frac{v+k_y a_1 u}{\sqrt{v^2+au^2}}+2\pi\left((\mu_1 v^3-\mu_2 v)I_1+2\mu_1 v u^2 I_3\right)\right) \\ F_{\perp}=\frac{\mu_x F_0 uv}{\sqrt{(u^2+bv^2)(v^2+au^2)}},\ M_C = M_0\frac{u+k_y m_1 v}{\sqrt{u^2+mv^2}} \end{gathered} \tag{10}$$

New kind of the approximated dry friction models permits to escape using not smooth functions in the cases when velocities are changed their signs. These models are are defined by the same coefficients amounts as models based on Pade expansions and completely satisfies to the all integral model analytical properties. Moreover, the accuracy of these models are corresponded to accuracy of the models based on the second order Pade expansions

4 CONCLUSIONS

The proposed dry friction models enables as well to describe the relationship between force and kinematical characteristics by smooth analytical functions over the entire range of angular and linear velocities as to take into account the more realistic representation about normal contact stresses distribution and differential characteristics of the friction law. The approximate models preserve all analytical properties of the models based on the exact integral expressions and correctly describe the behaviour of the net vector and torque of the friction forces and their first derivatives at zero and infinity.

Moreover, the models coefficients are numbers that can be identified from experiments. Consequently, these models may be considered as phenomenological models of the combined dry friction.

The procedure of approximate dry friction models construction is universal method of the dry friction model developing in more difficult cases.

REFERENCES

[1] Contensou, P. Couplage Entre Frottenment de Glissement et Frottenment de Pivotement Dans la Th ſeorie de la Toupie. *In Kreiselprobleme Gyrodynamics: IUTAM Symp. Celerina* (1962):201-216.

[2] Zhuravlev, V.Ph. The Model of Dry Friction in the Problem of the Rolling of Rigid Bodies. *Prikl. Mat. Mekh.* (1998) **62(5)**:762–767.
[3] Zhuravlev, V.Ph. and Kireenkov, A.A. Pade Expansions in the Two-Dimensional Model of Coulomb Friction. *Izv. Akad. Nauk. Mekh. Tverd. Tela*. (2005) **2**: 3–14.
[4] Kireenkov, A.A. About the motion of the symmetric rigid solid alone the plane. *8th CONFERENCE on Dynamical Systems: Theory and Applications. December 12-15, 2005. Lodz, Poland. Proceedings*. (2005) **1**: 95-102.
[5] Kireenkov A.A. Three-dimensional Model of Combined Dry Friction and Its Application in Non-holonomic Mechanics. *The Fifth EUROMECH Nonlinear Dynamics Conference, August 7-12, 2005, TUE, Eindhoven, The Netherlands, CD-Rom Proceedings*. (2005): 571-577
[6] Kireenkov, A.A.: Coupled Models of Sliding and Rolling Friction. *Doklady Akademii Nauk*, (2008). **419(6)**: 759–762.
[7] Kireenkov A.A. Combined model of sliding and rolling friction in dynamics of bodies on a rough plane. *Izv. Akad. Nauk. MTT*. (2008). **3**: 116-131.
[8] A.A.Kireenkov, A.A. Generalized two dimensional model of the sliding and spinning friction. *Doklady Physics*. (2010), **55(4)**:482—486.
[9] Andronov, V.V.; Zhuravlev, V.Ph.: *Dry Friction in Problems of Mechanics*. Moscow-Izhevsk: NITs Reg. Khaot. Dinam., (2010).
[10] Kireenkov, A.A.: Coupled Model of Sliding and Spinning Friction. *Doklady Akademii Nauk*. (2011). **441(6)**: 750–755.
[11] Ivanov A.P.: A dynamically consistent model of the contact stresses in the plane motion of a rigid body, *Prikl. Mat. Mekh.*, (2009). **73(2)**: 134-144.
[12] Kireenkov, A.A. On the Motion of a Homogeneous Rotating Disk along a Plane in the Case of Combined Friction. *Izv. Akad. Nauk. Mekh. Tverd. Tela*. (2002) **1**:60–67.
[13] Leine, R.I., Ch.Glocker, Ch.: A Set-Valued Force Law for Spatial Coulomb-Contensou Friction. *European J. of Mech.* (2003). **22(2)**: 193–216.
[14] Karapetyan, A.V.: A two-parameter friction model. *Prikl. Mat. Mekh.*, (2009). **73(4)**: 515-519.
[15] A.A. Kireenkov A.A. and Semendyaev S.V. and Filatov V.V. Experimental Study of Coupled Two-Dimensional Models of Sliding and Spinning Friction. *Izv. Akad. Nauk. Mekh. Tverd. Tela*. (2010) **6**:192-202
[16] Kireenkov A.A. Theoretical and experimental modelling of the Combined Dry Friction Effects. *IFAC – MATHMOD - 2012 Proceedings*, www.IFAC-PapersOnLine.net.

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis(Eds)

ROTATION OF THE BODY WITH MOVABLE INTERNAL MASSES AROUND THE CENTER OF MASS ON THE ROUGH PLANE

ALEKSANDR SAKHAROV

Moscow Institute of Physics and Technology (State University)
9 Institutskiy per., Dolgoprudny, Moscow Region, Russian Federation
e-mail: sah.aleksandr@gmail.com, web page: http://www.mipt.ru/

Key words: Dry Friction, Combined Friction, Normal Stress, Multibody System, Movable Masses, Slider

Abstract. We consider motion of a system consisting of a rigid body and internal movable masses on a rough surface. The possibility of rotation of the system around its center of mass because of the internal movable mass motion is investigated. To describe the friction between the body and the reference surface selected local Amonton–Coulomb law. To determine the distribution of the normal stress in the contact area between the body and the surface linear dynamically collaborative model is used. As examples we consider two configurations of internal masses: hard horizontal disk and two material points, moving parallel to the longitudinal axis of the body symmetry in the opposite way. Motions of the system analyzing for selected configurations.

1 INTRODUCTION

At the moment mobile robotics find wide application in various fields of human activity: research, medicine, space, etc. Among the large variety of robotic systems moving on a plane, class of the devices which move without any external moving parts, such as wheels, tracks, legs can be distinguished. Such systems have a number of advantages associated with their hermiticity and isolation from external adverse factors. For example a robot-ball could be called.

Motion of the robot without external drivers can be achieved by using movable internal masses and interaction with the supporting surface by friction forces. In the works [1, 2] investigated one-dimensional motion of a dual-mass system in the presence of dry friction forces. Optimal parameters which give the highest average speed of the body with imposed restrictions is determined. In the paper [3] investigated the body with one and two internal masses and three models of resistance to the robot motion: a piecewise-linear friction, quadratic friction and dry friction.

Two-dimensional movement of the three-mass robot in a liquid medium is investigated in the work [4, 5]. The motion of the system is achieved by moving two internal masses and the forces of viscous resistance. The possibility of the body rotation on a rough surface with the rotor inside the body, which rapidly rotating around the longitudinal axis of the body symmetry while the body move direct this axis is proved in [6, 7].

In this work, we study the rotation of the symmetrical hollow body around its center of mass with by moving material points inside the body. It is assumed that the points are located on a horizontal plane passing through the body center of mass and system center of mass does not move inside the body. In contrast to [6, 7], in the present work applies the local friction of rest.

2 MODEL OF THE SYSTEM

Consider a rigid body of mass m_0, representing a hollow parallelepiped composed of six homogeneous rectangles. Length of the body equal to a, width – b, height – $2h$. The body relies on a rough horizontal surface. We introduce a laboratory coordinates system $O'xyz$ with start on the surface and connected with the body system $O\xi\eta\zeta$, where O – body center of mass (Fig. 1). Axis $O'z$ and $O\zeta$ forward vertically, the axis $O\xi$ is parallel to the larger edge of the parallelepiped base, the axis $O\eta$ directed so that the system formed right-hand triple. We suppose that in the initial moment corresponding axis coincide. The main axis of the central ellipsoid of inertia of the body coincide with the coordinates system $O\xi\eta\zeta$. Moment of inertia with respect to axis $O\zeta$ is equal to C. Inside the body material points are located such way that center of mass of system coincides with the center of mass of the body O. The system consists of the body and internal masses will called a slider.

In the general case position of the body is determined by three coordinates. Coordinates x_O, y_O determine point O in the laboratory coordinates system. The angle φ between the axes $O'x$ and $O\xi$ determine rotation of the slider relative to its initial position. But in this work we suppose that at the initial time body stay at rest. Due to the special chose of the internal mass configuration and motion point O will be fixed during the whole time of motion. Thus position of the body is determined by the angle of rotation φ and coordinates of points O' and O are $x_O = x_{O'} = 0$, $y_O = y_{O'} = 0$, $z_O = h$, $z_{O'} = 0$.

3 FORCES AND MOMENTS

Let $\sum_{i=1}^{s} m_i$ – is the sum of the inside movable material points masses, s – the number of points. Write theorem on the slider center of mass motion and the momentum change of the slider relative to the point O:

$$m\ddot{\boldsymbol{r}}_O = \boldsymbol{F}_g + \boldsymbol{F}_n + \boldsymbol{F}_t, \quad m = m_0 + \sum_{i=1}^{s} m_i, \tag{1}$$

$$\dot{\boldsymbol{K}}_O = \boldsymbol{M}_O^g + \boldsymbol{M}_O^n + \boldsymbol{M}_O^t, \tag{2}$$

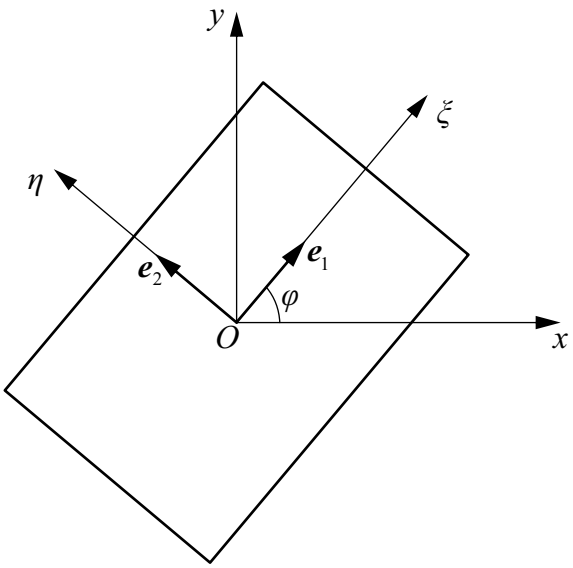


Figure 1: Laboratory and body systems of coordinates

where $\boldsymbol{r}_O = \overrightarrow{O'O}$ – the radius-vector directed from the point O' to the body center of mass O, $\boldsymbol{F}_g$ – the force of gravity acting on the slider, $\boldsymbol{F}_n$ and $\boldsymbol{F}_t$ – the main vectors of norm reactions and friction forces respectively. $\boldsymbol{K}_O$ – kinetic moment of the slider relative the point O, $\boldsymbol{M}_O^g$ – the main moment of gravity force relative to the point O, $\boldsymbol{M}_O^n$ and $\boldsymbol{M}_O^t$ – the main moment of normal reactions and friction forces relative the point O respectively.

The radius-vector $\boldsymbol{r}_O$ and unit vectors of the body coordinates system is equal to

$$\boldsymbol{r}_O = \begin{bmatrix} 0 \\ 0 \\ h \end{bmatrix}, \quad \boldsymbol{e}_1 = \begin{bmatrix} \cos\varphi \\ \sin\varphi \\ 0 \end{bmatrix}, \quad \boldsymbol{e}_2 = \begin{bmatrix} -\sin\varphi \\ \cos\varphi \\ 0 \end{bmatrix}, \quad \boldsymbol{k} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix},$$

Kinetic moment of the slider $\boldsymbol{K}_O$ can be represented as a sum of the kinetic moments of the body and internal movable masses:

$$\boldsymbol{K}_O = \boldsymbol{K}_{O0} + \sum_{i=1}^{s} \boldsymbol{K}_{Oi}, \tag{3}$$

where $\boldsymbol{K}_{O0} = C\dot{\varphi}\boldsymbol{k}$ – kinetic moment of the body, $\boldsymbol{K}_{Oi}$ – kinetic moment of i-th internal mass.

The force of gravity, main vectors of normal reactions and friction forces, as well as the corresponding moments are determined by the following expression:

$$\boldsymbol{F}_g = -mg\boldsymbol{k}, \quad \boldsymbol{F}_n = \boldsymbol{k}\iint n_A ds, \quad \boldsymbol{F}_t = \iint \boldsymbol{t}_A ds, \tag{4}$$

$$\boldsymbol{M}_O^g = -\sum_{i=1}^{s} \boldsymbol{r}_{Oi} \times m_i g\boldsymbol{k}, \quad \boldsymbol{M}_O^n = \iint n_A \boldsymbol{r}_{OA} \times \boldsymbol{k} ds, \quad \boldsymbol{M}_O^t = \iint \boldsymbol{r}_{OA} \times \boldsymbol{t}_A ds. \tag{5}$$

Here and elsewhere the integration extends over the region $D(\xi, \eta)$, consisting of body and surface contact points. n_A and $\boldsymbol{t}_A$ – normal and tangential stresses at point $A \in D$, $\boldsymbol{r}_{Oi} = \xi_i(t)\boldsymbol{e}_1 + \eta_i(t)\boldsymbol{e}_2$ and $\boldsymbol{r}_{OA} = \xi\boldsymbol{e}_1 + \eta\boldsymbol{e}_2 - h\boldsymbol{k}$ — radius-vectors from the point O to the i-th internal mass and point A, respectively. Friction is locally described by the locally Amontons–Coulomb law, which extend to the case of the body at rest:

$$\boldsymbol{t}_A = -\mu n_A \boldsymbol{f}_A, \quad \boldsymbol{f}_A = \begin{cases} \upsilon_A^{-1}\boldsymbol{\upsilon}_A, & \upsilon_A \neq 0, \\ \boldsymbol{\tau}_A, & \upsilon_A = 0, \end{cases}$$

where μ – coefficient of friction, $\boldsymbol{\upsilon}_A = \dot{\boldsymbol{r}}_A$, $\boldsymbol{r}_A = \overrightarrow{O'A}$. Since $\boldsymbol{\upsilon}_O = 0$, velocity of the arbitrary point A of the slider base express by the formula:

$$\boldsymbol{\upsilon}_A = \dot{\varphi}\boldsymbol{k} \times \boldsymbol{r}_{OA} = \dot{\varphi}\left(\xi\boldsymbol{e}_2 - \eta\boldsymbol{e}_1\right),$$

where $\dot{\varphi}\boldsymbol{k}$ – angular velocity of the body. Thus, unit vector $\upsilon_A^{-1}\boldsymbol{\upsilon}_A$ expressed:

$$\upsilon_A^{-1}\boldsymbol{\upsilon}_A = \frac{\dot{\varphi}}{|\dot{\varphi}|}\frac{-\eta\boldsymbol{e}_1 + \xi\boldsymbol{e}_2}{\sqrt{\xi^2 + \eta^2}}.$$

The module of the vector $\boldsymbol{\tau}_A$ is determined by attached to the slider forces and not greater than one. This approach was previously proposed by John H. Jellett [8, 9]. Because point O fixed, vectors $\boldsymbol{\upsilon}_A$ directed perpendicular to the radius-vector $\boldsymbol{r}_{OA}$. Suppose that vectors $\boldsymbol{\tau}_A$ have the same orientation. Combining vectors $\upsilon_A^{-1}\boldsymbol{\upsilon}_A$ and $\boldsymbol{\tau}_A$ in one formula we obtain the following expression (Fig. 2):

$$\boldsymbol{f}_A = f_\xi\boldsymbol{e}_1 + f_\eta\boldsymbol{e}_2, \quad f_\xi = -f_A\frac{\eta}{\sqrt{\xi^2 + \eta^2}}, \quad f_\eta = f_A\frac{\xi}{\sqrt{\xi^2 + \eta^2}}, \quad -1 \leq f_A \leq 1.$$

Continuity condition of the contact between the body and the surface imposes on the kinematic characteristics three independent constraints. Therefore, the model of normal stresses n_A should include three independent parameters λ_0, λ_1 and λ_2, which is determines from these limitations each time. This model of normal stress distribution is dynamically collaborative [8, 10].

$$n_A = \lambda_0 + \lambda_1\xi + \lambda_2\eta, \quad (\xi, \eta) \in D. \tag{6}$$

The physical interpretation of this formula can serve as a representation of small deformations of the surface in the contact area, leading to the normal stresses by Hooke's law.

Substituting equation expressing the Amontons–Coulomb law and formula (6) in expressions (4), (5), then substituting the resulting expression into a system (1), (2), we get the equation of motion of the slider in closed form. Each time are unknown coordinate φ (coordinates x_O, y_O assumed to be constant) and the coefficients λ_0, λ_1 and λ_2 of the linear model (6).

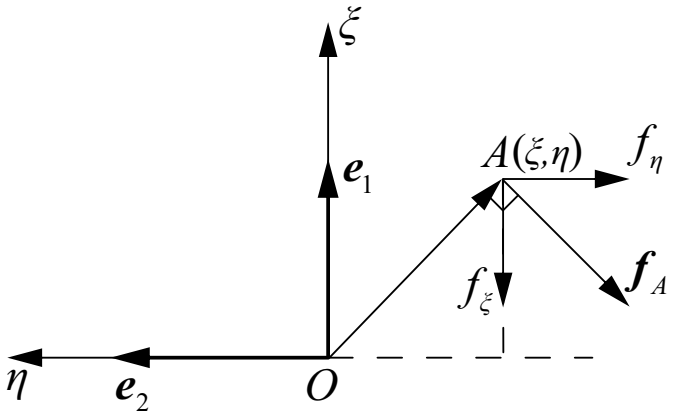


Figure 2: Decomposition of function $\boldsymbol{f}_A$ on the basis $\boldsymbol{e}_1$, $\boldsymbol{e}_2$

4 COEFFICIENTS OF THE NORMAL STRESS MODEL

To determine the coefficient λ_0 in the decomposition (6), simplify the expression for the main vector of normal reaction calculating the integral in the second expression of the (4). For this purpose it is necessary to take into account the symmetry of the slider and surface contact area $D(\xi, \eta)$:

$$-\frac{a}{2} \leq \xi \leq \frac{a}{2}, \quad -\frac{b}{2} \leq \eta \leq \frac{b}{2}.$$

Then the second expression of the (4) is easily integrated:

$$\boldsymbol{F}_n = \boldsymbol{k} \iint (\lambda_0 + \lambda_1 \xi + \lambda_2 \eta) d\xi d\eta = \lambda_0 ab\boldsymbol{k}. \tag{7}$$

The continuity condition of the contact between the slider and the surface is equal to zero projections of the velocities (and accelerations) of the body points on the vertical. This condition can be expressed in three independent equations. First equation is obtained from the required that projection of the main force on the vertical equality to zero, i.e. the projection of the equation (1) on unit vector $\boldsymbol{k}$:

$$(\boldsymbol{F}_g + \boldsymbol{F}_n) \cdot \boldsymbol{k} = 0. \tag{8}$$

Substituting first formula (4) and (7) in (8), we find the coefficient λ_0:

$$\lambda_0 = \frac{mg}{ab}. \tag{9}$$

Physical sense of the value λ_0 is the average pressure of the body weight mg on a surface area ab.

To obtain expressions for the coefficients λ_1 and λ_2 we project equation (2) on unit vectors $\boldsymbol{e}_j$, $(j = 1, 2)$. Main moment of the gravitational forces equals to zero because of

the configuration symmetry of the internal masses: $\boldsymbol{M}_O^g = 0$. For the projections of the main moments of the normal reactions and friction forces $\boldsymbol{M}_O^n$ and $\boldsymbol{M}_O^t$ we obtain:

$$\boldsymbol{M}_O^n \cdot \boldsymbol{e}_j = \iint n_A \left(\boldsymbol{r}_{OA} \times \boldsymbol{k}\right) \cdot \boldsymbol{e}_j ds = \iint n_A \boldsymbol{k} \cdot \left(\boldsymbol{e}_j \times \boldsymbol{r}_{OA}\right) ds,$$
$$\boldsymbol{M}_O^t \cdot \boldsymbol{e}_j = -\iint \mu n_A \left(\boldsymbol{r}_{OA} \times \boldsymbol{f}_A\right) \cdot \boldsymbol{e}_j ds = -\iint \mu n_A \boldsymbol{f}_A \cdot \left(\boldsymbol{e}_j \times \boldsymbol{r}_{OA}\right) ds.$$

We also use the decomposition (6), consider that $\dot{\boldsymbol{K}}_O \cdot \boldsymbol{e}_j = 0$ (all internal masses move at the horizontal plane containing point O) and then obtain a system of two equations:

$$\iint \psi_j \left(\lambda_0 + \lambda_1 \xi + \lambda_2 \eta\right) ds = 0, \quad \psi_j = \left(\boldsymbol{k} - \mu \boldsymbol{f}_A\right) \cdot \left(\boldsymbol{e}_j \times \boldsymbol{r}_{OA}\right). \tag{10}$$

Perform calculations in the expression of ψ_j to simplify it:

$$\psi_1 = \eta - \mu h f_\eta, \quad \psi_2 = -\xi + \mu h f_\xi.$$

We rewrite the expression (10) in the form of a two linear equations relative λ_1 and λ_2:

$$a_{11}\lambda_1 + a_{12}\lambda_2 = a_{10}\lambda_0,$$
$$a_{21}\lambda_1 + a_{22}\lambda_2 = a_{20}\lambda_0,$$

where coefficients:

$$a_{11} = \iint \psi_1 \xi ds, \quad a_{12} = \iint \psi_1 \eta ds, \quad a_{10} = -\iint \psi_1 ds,$$
$$a_{21} = \iint \psi_2 \xi ds, \quad a_{22} = \iint \psi_2 \eta ds, \quad a_{20} = -\iint \psi_2 ds.$$

The solution of the system:

$$\lambda_1 = \frac{a_{10}a_{22} - a_{12}a_{20}}{a_{11}a_{22} - a_{12}a_{21}}\lambda_0, \quad \lambda_2 = \frac{a_{11}a_{20} - a_{10}a_{21}}{a_{11}a_{22} - a_{12}a_{21}}\lambda_0. \tag{11}$$

It is easy to verify that in this case λ_1 and λ_2 are equal to zero. Indeed, coefficients $a_{10} = a_{20} = 0$ due to the odd of subintegral function ψ_j and symmetry of the region D. It remains to show that the denominator in the formula (11) is not equal to zero:

$$a_{11}a_{22} = -(\mu h)^2 \iint f_\eta \xi ds \iint f_\xi \eta ds =$$
$$= (\mu h f_A)^2 \iint \frac{\xi^2}{\sqrt{\xi^2 + \eta^2}} ds \iint \frac{\eta^2}{\sqrt{\xi^2 + \eta^2}} ds > 0,$$
$$a_{12}a_{21} = \left(\frac{ab^3}{12} - \mu h \iint f_\eta \eta ds\right)\left(-\frac{a^3 b}{12} + \mu h \iint f_\xi \xi ds\right) = -\frac{a^4 b^4}{144} < 0.$$

It follows that

$$a_{11}a_{22} - a_{12}a_{21} > 0.$$

Thus, in this case, we find that $\lambda_1 = \lambda_2 = 0$. Within the selected linear model, the distribution of the contact stresses is constant: $n_A = \lambda_0$.

5 EQUATION OF MOTION

To obtain the equation of the slider motion projecting equation (2) on the unit vector $\boldsymbol{k}$:

$$\dot{\boldsymbol{K}}_O \cdot \boldsymbol{k} = \boldsymbol{M}_O^t \cdot \boldsymbol{k}. \tag{12}$$

Discover the right part of this equation:

$$\boldsymbol{M}_O^t \cdot \boldsymbol{k} = -\mu \iint n_A \left[(\xi \boldsymbol{e}_1 + \eta \boldsymbol{e}_2 - h\boldsymbol{k}) \times \boldsymbol{f}_A\right] \cdot \boldsymbol{k} ds = -\mu\lambda_0 \iint \left(f_\eta \xi - f_\xi \eta\right) ds.$$

Substituting this expression, formula (3) and expressions for f_ξ, f_η in equation (12), we obtain following equation of motion

$$C\ddot{\varphi} + \sum_{i=1}^{s} \dot{\boldsymbol{K}}_{Oi} \cdot \boldsymbol{k} = -f_A T_\varphi, \tag{13}$$

where

$$T_\varphi = \mu\lambda_0 \iint \sqrt{\xi^2 + \eta^2} ds.$$

Left part of this equation depends on the specific configurations of the internal masses and the laws of their relative motion. Let us consider two examples.

5.1 Disk

Let us put inside the body a disk mass m_d and radius d with horizontal plane and the centre coincided with the point O (Figure 3a). Assume that the disk could rotate around its axis by engine. Then its kinetic moment equals to:

$$\sum_{i=1}^{s} \boldsymbol{K}_{Oi} = \boldsymbol{K}_{Od} = J_d \left(\dot{\varphi} + \omega_d\right) \boldsymbol{k},$$

where $J_d = \dfrac{m_d d^2}{2}$ – axial moment of inertia, $\omega_d \boldsymbol{k}$ – angular velocity of the disk relative to the body.

Substituting this formula into equation (13), we obtain the equation of the slider motion with the horizontal disk located inside:

$$\left(C + J_d\right) \ddot{\varphi} = R_d - f_A T_\varphi, \tag{14}$$

where $R_d = -J_d \dot{\omega}_d$. Considering the double inequality $-1 \leq f_A \leq 1$, the value f_A is defined by following expression:

$$f_A = \begin{cases} \dot{\varphi}|\dot{\varphi}|^{-1}, & \dot{\varphi} \neq 0, \\ R_d T_\varphi^{-1}, & \dot{\varphi} = 0,\ |R_d| \leq T_\varphi, \\ R_d |R_d|^{-1}, & \dot{\varphi} = 0,\ |R_d| > T_\varphi. \end{cases}$$

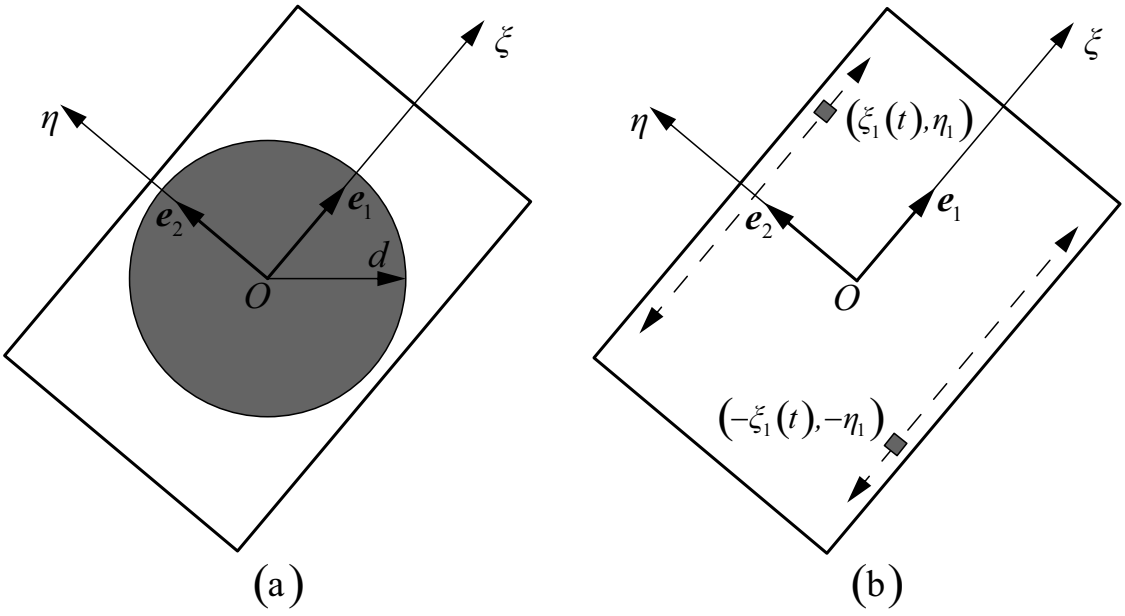


Figure 3: Slider with the disk (a) and two point masses (b)

From this formula is possible to find the minimum value of angular acceleration of the rotor $\dot{\omega}_d$ is required to start rotation of the body:

$$|\dot{\omega}_d| > \frac{T_\varphi}{J_d}. \tag{15}$$

As an example of a simple control law of the relative angular acceleration of the disk we choose the two-phase piecewise constant periodic function. Then the relative angular velocity of the disc must satisfy the periodicity conditions: $\omega_d(0) = \omega_d(t^*) = 0$, where t^* is the period of the disk rotation. Such conditions the following function satisfies:

$$\omega_d = \Omega \cdot \begin{cases} t\tau^{-1}, & 0 \le t < \tau, \\ (t-t^*)(\tau-t^*)^{-1}, & \tau \le t \le t^*; \end{cases} \qquad \dot{\omega}_d = \Omega \cdot \begin{cases} \tau^{-1}, & 0 \le t < \tau, \\ (\tau-t^*)^{-1}, & \tau \le t \le t^*; \end{cases}$$

where $0 < \tau < t^*$, Ω — maximum relative angular velocity of the disc. According with condition (15), the body begin to move if one of the following conditions satisfied:

$$\tau \le |\Omega|\frac{J_d}{T_\varphi}, \quad \text{or} \quad \tau \ge t^* - |\Omega|\frac{J_d}{T_\varphi}.$$

On Figure 4 showed spaces of parameters (τ, t^*) (Figure 4a) and $(\tau, |\Omega|)$ (Figure 4b). Shaded areas correspond to the values of the parameters at which the body begins to move.

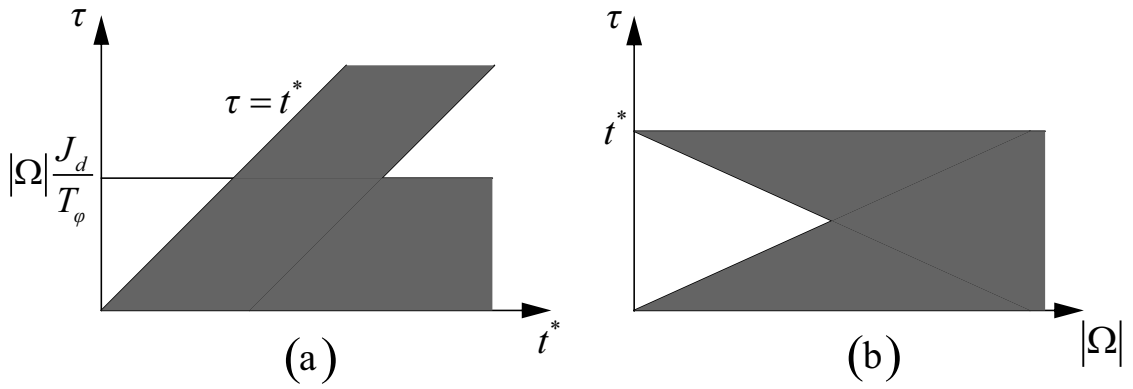


Figure 4: Shaded areas of the parameters (τ, t^*) (a) and $(\tau, |\Omega|)$ (b) corresponding beginning of the body movement

For the selected control law right part of the equation (14) piecewise constant. Depending on the functions f_A, equation takes the following three types:

$$\ddot{\varphi} = 0, \quad \ddot{\varphi} = \frac{R_d}{C + J_d}\left(1 \pm \frac{T_\varphi}{R_d}\right).$$

Thus the trajectory of the body $\varphi(t)$ for piecewise constant control of the relative acceleration of the disk consists of a set of lines and parabolas.

5.2 Two point masses

Let us place inside the body two material points with masses $m_1 = m_2$. Mass can move relative to the body on two rails parallel to its longitudinal axis of symmetry and lying in a horizontal plane passing through the body center of mass O (Figure 3b). The rails are positioned at equal distances $\eta_1 = -\eta_2 > 0$ of the point O. We assume that the masses are moved inside the body in antiphase: $\xi_1(t) = -\xi_2(t)$. Kinetic moment of i-th point mass $(i = 1, 2)$:

$$\boldsymbol{K}_{Oi} = \boldsymbol{r}_{Oi} \times m_1 \boldsymbol{v}_i = m_1 \boldsymbol{r}_{Oi} \times \left(\dot{\varphi}\boldsymbol{k} \times \boldsymbol{r}_{Oi} + \dot{\xi}_i(t)\,\boldsymbol{e}_1\right) = m_1\left(\dot{\varphi} r_{Oi}^2 - \dot{\xi}_i(t)\,\eta_i\right)\boldsymbol{k},$$

where $\boldsymbol{v}_i$ – absolute velocity of i-th internal mass. Because of selected restrictions on the function $\xi_i(t)$ kinetic moment will be the same for both masses. Summarizing, calculating the time derivative and projecting on the unit vector $\boldsymbol{k}$, we obtain:

$$\sum_{i=1}^{s} \dot{\boldsymbol{K}}_{Oi} \cdot \boldsymbol{k} = 2m_1\left(\left(\xi_1^2(t) + \eta_1^2\right)\ddot{\varphi} + 2\xi_1(t)\dot{\xi}_1(t)\dot{\varphi} - \ddot{\xi}_1(t)\eta_1\right).$$

Thus, substituting this formula in (13), we find the equation of the slider motion with two internal masses moves in antiphase:

$$\left(C + 2m_1\left(\xi_1^2(t) + \eta_1^2\right)\right)\ddot{\varphi} = R_p - f_A T_\varphi, \tag{16}$$

where $R_p = -2m_1\left(2\xi_1(t)\dot{\xi}_1(t)\dot{\varphi} - \ddot{\xi}_1(t)\eta_1\right)$. The value of f_A is defined similarly to the previous subsection:

$$f_A = \begin{cases} \dot{\varphi}|\dot{\varphi}|^{-1}, & \dot{\varphi} \neq 0, \\ R_p T_\varphi^{-1}, & \dot{\varphi} = 0,\ |R_p| \leq T_\varphi, \\ R_p|R_p|^{-1}, & \dot{\varphi} = 0,\ |R_p| > T_\varphi. \end{cases}$$

Absolute value of the relative acceleration $\ddot{\xi}_1(t)$ of the internal mass is needed to start rotation of the body is bounded below:

$$|\ddot{\xi}_1(t)| > \frac{T_\varphi}{2m_1\eta_1}. \tag{17}$$

Let's consider two-phase piecewise linear function $\xi_1(t)$ as a control law of the internal mass relative motion. The function must satisfy the boundary conditions: $\xi_1(0) = -a_1$, $\xi_1(\tau) = a_1$, $\xi_1(t^*) = -a_1$, where a_1 – amplitude of the movable mass displacement inside the body. Prescribed conditions satisfies the following function:

$$\xi_1(t) = a_1 \cdot \begin{cases} 2t\tau^{-1} - 1, & 0 \leq t < \tau, \\ 1 - 2(t-\tau)(t^* - \tau)^{-1}, & \tau \leq t \leq t^*. \end{cases}$$

Thus for the relative velocity $\dot{\xi}_1(t)$ we obtain:

$$\dot{\xi}_1(t) = 2a_1 \cdot \begin{cases} \tau^{-1}, & 0 \leq t < \tau, \\ (\tau - t^*)^{-1}, & \tau \leq t \leq t^*. \end{cases}$$

Relative acceleration of the internal mass is always equals to zero, except the moments 0, τ and t^*, when the relative velocity changes abruptly:

$$\ddot{\xi}_1(t) = \frac{2a_1}{\tau}\delta(t) - \left(\frac{2a_1}{\tau} + \frac{2a_1}{t^* - \tau}\right)\delta(t-\tau) + \frac{2a_1}{t^* - \tau}\delta(t - t^*), \quad 0 \leq t \leq t^*,$$

where $\delta(t)$ – delta–Dirac function. It is obvious that for the chosen unction $\xi_1(t)$ condition (17) will be executed at time 0, τ and t^* with any parameters of the control law. Stepwise change of the relative velocity of the movable mass means the presence of an impact in the system, which has a decisive influence on the dynamics of the body. The magnitude of the stepwise change of angular velocity is founded by integrating the equation of motion (16) for infinitely small period of time which includes an impact. Thus, at moments 0, τ and t^* the body has the instantaneous increment of angular velocity:

$$\dot{\varphi}(t+\varepsilon) - \dot{\varphi}(t-\varepsilon) = \frac{2m_1\eta_1}{C + 2m_1\left(\xi_1^2(t) + \eta_1^2\right)} \int\limits_{t-\varepsilon}^{t+\varepsilon} \ddot{\xi}_1(t)dt, \quad \varepsilon \to 0, \quad t = 0,\ \tau,\ t^*.$$

In the remaining time points relative angular acceleration of the internal mass is zero and the function $R_p = -4m_1\xi_1(t)\dot{\xi}_1(t)\dot{\varphi}$, $t \neq 0,\ \tau,\ t^*$.

6 CONCLUSIONS

In this work we considered a rigid body based on a rough surface. The body has a hollow parallelepiped shape and could contain a movable masses. Masses are arranged in such way that the slider center of mass remained constant inside the body and coincide with the body center of mass. It was assumed that the initial velocity of the body center of mass is zero. Thus the position of the body is specified by the rotation angle relative to the initial position. In the work the local Amonton–Coulomb law was used. As a model of normal stress distribution in the contact area between the body and the surface dynamically consistent linear model was selected. It was shown that within the framework of this model, the distribution of normal stresses are constant. The equation of body motion with an arbitrary configuration of movable masses was obtained.

As example considered horizontal disk and two point masses moving in antiphase. Considered a two-phase piecewise constant control law of the relative acceleration of the disk. Area of the control law parameters which correspond to overcome the force of static friction and start the motion was found. Motion was qualitatively described. For two point masses proposed two-phase piecewise linear control law of shift points. It is established that the motion of the system is determined by the presence of impact resulting by stepwise changes of the mass relative velocity.

Author express thanks to N. N. Bolotnik for helpful comments. The work was supported by the basic part of the state tasks in the field of scientific activity No. 2014/120 “Investigation of the regularities in the dynamics of systems with friction and the development of mobile robots without external drivers” (research No. 2583) and the Russian Foundation for Basic Research (No. 14-01-00432).

REFERENCES

[1] Chernousko, F.L. Analysis and optimization of the motion of a body controlled by means of a moveable internal mass. *J Appl Math Mech*, 2006. Vol. 70, N 6. Pp. 819–842.

[2] Chernousko, F.L. The optimal periodic motions of a two-mass system in a resistant medium. // *J Appl Math Mech*, 2008. Vol. 72, N 2. Pp. 116–125.

[3] Chernousko, F.L. and Bolotnik, N.N. Mobile robots controlled by internal bodies movement. *Tr. IMM UrO RAN*, 2010. Vol. 16, N 5. Pp. 213–222. (in Russian)

[4] Volkova, L.Iu. and Iatsun, S.F. Control of the three-mass robot motion moving in a liquid medium. *Nonlinear Dynamics*, 2011. Vol. 7, N 4. Pp. 845–857. (in Russian)

[5] Iatsun, S.F. and Volkova, L.Iu. Simulation of dynamic modes of vibration robot moving along the surface of the viscous resistance. *Specialized machinery and communication*, 2012. N 3. Pp. 25–29. (in Russian)

[6] Ivanov, A.P. and Sakharov, A.V. On the dynamics of a rigid body with moving masses and a rotor on a rough plane. *Nonlinear Dynamics & Mobile Robotics*, 2013. Vol. 1, N 1. Pp. 23–32.

[7] Sakharov, A.V. Rotation of a body without external drivers by a rotor. *Proceedings of MIPT*, 2014. Vol. 6, N 2. Pp. 80–91. (in Russian)

[8] Ivanov, A.P. *Fundamentals ot the theory of systems with friction.* Moscow–Izhevsk: Regular & Chaotic Dynamics, Institute of Computer Sciences, 2011. 304 p. (in Russian)

[9] Jellett, J.H. *A treatise on the theory of friction.* Dublin, London: MacMillan, 1872. 220 p.

[10] Ivanov, A.P. A Dynamically consistent model of the contact stresses in the plane motion of a rigid body. *J Appl Math Mech*, 2009. Vol. 73, N 2. Pp. 134–144.

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis (Eds)

CONSOLIDATED FIRE TESTING – A FRAMEWORK FOR THERMO-MECHANICAL MODELLING

P. SCHULTHESS[*], M. NEUENSCHWANDER[*], M. KNOBLOCH[†] AND M. FONTANA[*]

[*] Institute of Structural Engineering
Swiss Federal Institute of Technology (ETH Zurich)
Stefano-Franscini-Platz 5, 8093 Zurich, Switzerland
e-mail: schulthess@ibk.baug.ethz.ch, www.ibk.ethz.ch

[†] Institute of Steel, Lightweight and Composite Structures
Ruhr-Universität Bochum
Universitätsstraße 150, 44801 Bochum, Germany
e-mail: markus.knobloch@rub.de, www.ruhr-uni-bochum.de/stahlbau

Key words: Consolidated Fire Testing, Physical Sub-modelling, Fire Engineering

Abstract Consolidated testing facilitates the investigation of the global behavior of structures subjected to fire and therefore may become increasingly important in structural fire engineering. In order to develop a consolidated testing procedure that meets the requirements arising from structural fire engineering and considers thermal strains, thermal creep effects as well as strength and stiffness degradation, a consolidated testing benchmark problem is elaborated. The benchmark problem allows to perform coupled experimental and numerical tests that can be verified by pure physical testing. Furthermore, a framework for a consolidated test setup is developed, including a tangent stiffness update algorithm. Two preliminary tests at ambient temperature show the eligibility of the consolidated testing framework and are presented in this paper.

1. INTRODUCTION

In structural fire engineering the mechanical response of a structure under a fire exposure must be realistically predicted. Elevated temperatures lead to stiffness and strength degradation of construction materials and influence strongly the mechanical response of a structure. Additionally, thermal expansion, stress-inducing constraints to thermal expansion within a global structure, high-temperature creep phenomena and strain rate effects have to be considered. In todays structural design practise, the fire resistance is usually assessed based on the structural fire behaviour of isolated load-bearing members and connections under natural fire or standard fire exposure. However, global fire tests [1] and experience from real building fire incidents have proven that in case of fire entire structures perform better, than as predicted on the basis of their single components performance. This is due to global structural effects, like load redistributions or change of the structural system. Therefore, experimental and

numerical analysis of the global structural behaviour is crucial for a realistic performance assessment and economic design of a structure in fire [2]. On the other hand, global testing is very expensive and rare [1] and purely numerical simulations remain afflicted with uncertainty. The proposed consolidated testing method combines physical element testing and global structural numerical simulation, to overcome these restrictions. Yet, considerable complexity is added to the part of the physical element testing by the aforementioned temperature-dependent aspects of material behaviour. Therefore, a suitable benchmark problem for consolidated thermo-mechanical modelling was first developed and will be presented in this paper. This proposed basic framework for advanced thermo-mechanical modelling can be extended, to assess the global structural behaviour in fire and to develop a tool for fire resistance time verification of a global structure, e.g. 30, 60 or 90 minutes for standard fire as well as for a specific time or until full burnout in case of natural fires.

2. FRAMEWORK FOR CONSOLIDATED TESTING IN FIRE ENGINEERING

Effective structural fire design, considering the global structural behaviour is very costly and time consuming. However, by using a consolidated (coupled experimental and numerical) assessment method for thermo-mechanical problems, many difficulties due to experimental cost and uncertainty of testing can be overcome efficiently. Therefore, first some basic methodological problems have to be studied by means of a benchmark problem, which can be verified solely by physical testing. This methodology is described in this section.

2.1 Thermo-mechanical benchmark problem

Figure 1 overviews the developed concept of a thermo-mechanical benchmark test for the proposed consolidated assessment method. This benchmark problem, when solved consolidated, can be verified by physical testing. Figure 1a shows the examined structure of a simply supported beam with an additional restraint at midspan by a truss element. The connection between the beam and the truss element is hinged and the midspan deflection of the beam, w, is identical with the axial displacement of the truss element, u_{tr}. The entire structure is loaded with a concentrated force, $P(t)$, whereas only the truss element is additionally exposed to increasing temperature, $\theta(t)$. Figure 1b shows the mechanical and thermal action on the structure in function of time. From the beginning of the test at time, t_0, until the start of the temperature exposure at time, t_1, the load is linearly increased to a magnitude of, P_0. After the onset of the linear temperature increase, the magnitude of the load is held constant. The beam is assumed to remain in the elastic range and at ambient temperature throughout the entire test.

Figure 1d shows the implementation of the examined structure as a complete physical model in an universal testing machine, combined with an electric furnace. The truss element extending between the points B and C represents the gauge length on a test specimen inside the furnace. The change in length, u, of this gauge length is continuously measured with a high-temperature resisting extensometer. Additionally, the specimen can be exposed to a controlled temperature increase, $\theta(t)$, and the actual specimen temperature is constantly recorded. The truss element extending from point A to B stands for the connecting rods and attachments between the specimen inside the furnace and the machine cross-head and the loading frame. These two truss

elements represent the truss element restraining the beam at midspan in the examined structure. The beam itself is represented in the physical model with two connecting rods with a stiffness equal to the beam stiffness, $k_b=48EI/L^3$. With load-cells the forces corresponding to the supporting forces of the beam, F_1, and, F_3, and the force in the restraining truss element, F_2, are registered.

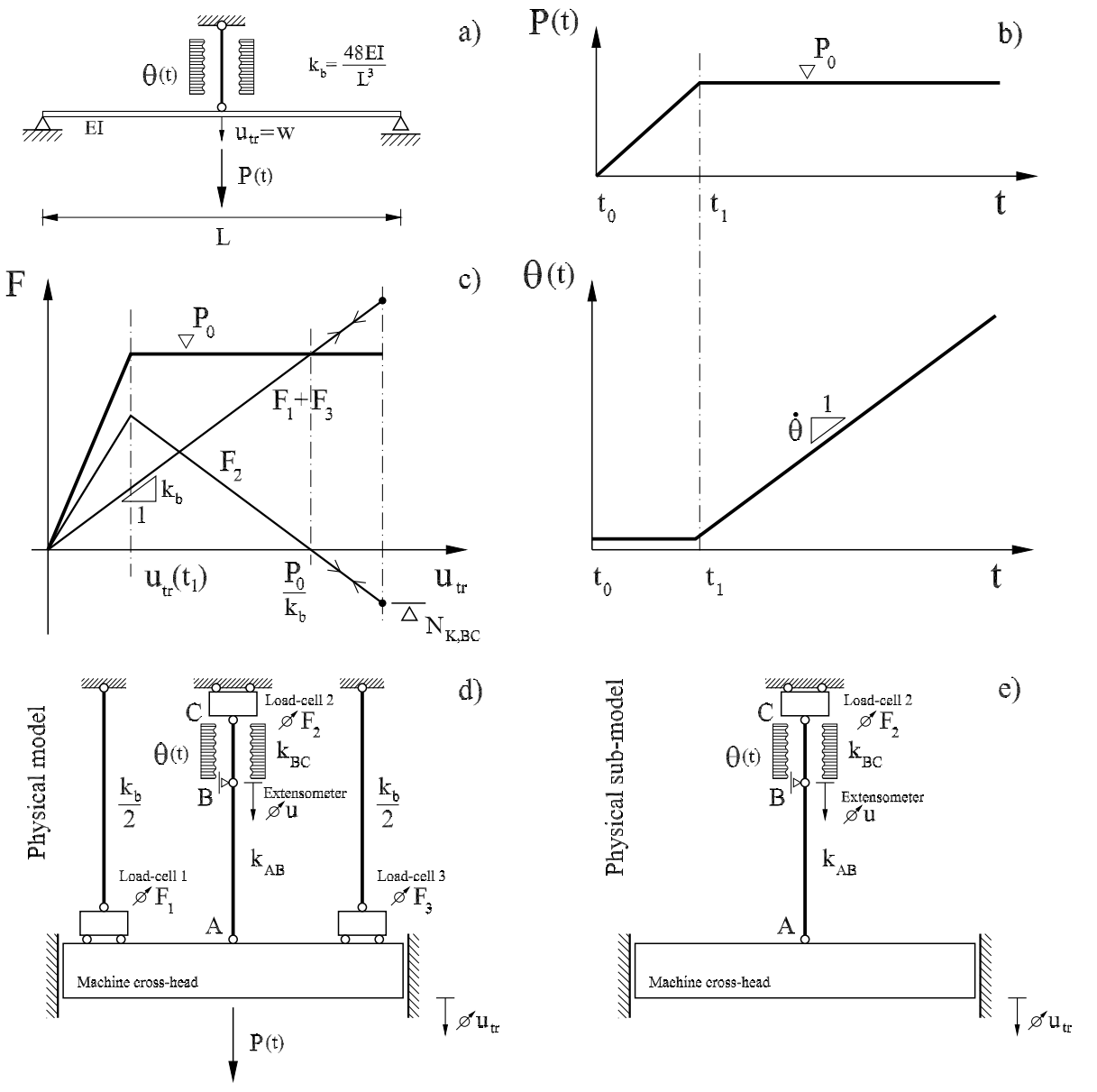


Figure 1: Benchmark test for thermo-mechanical consolidated testing. (a) Benchmark problem, (b) mechanical and thermal loading in function of time, (c) response of physical model, (d) test setup of physical model and (e) experimental setup of the physical sub-model of the consolidated model.

Figure 1c illustrates the response of the physical model under the mechanical and thermal action as shown in Figure 1b. With increasing cross-head displacement until, $u_{tr}(t_1)$, the actual sum of the supporting forces of the beam, F_1+F_3, follows the straight line of the beam characteristic with a slope of, k_b, whereas the force in the truss elements, F_2, follows the truss elements in series characteristic with a slope of, $k_{tr}=k_{AB}\cdot k_{BC}/(k_{AB}+k_{BC})$. Depending on the stiffness ratio between the beam and the truss elements the shares of the preload, P_0, that are finally taken by the beam and the truss elements can vary, but they always balance $F_1+F_2+F_3=P_0$, when the entire preload has been applied at time, t_1. With the onset of the temperature increase mainly the truss element *BC*, but as well the truss element *AB* start to expand. However, with increasing cross-head displacement the supporting forces of the beam increase too. Consequently, the force in the truss elements must decrease in order to balance together with the supporting forces of the beam the sustaining load P_0. The unloading of the truss elements continues until the entire sustaining load P_0 is balanced solely by the supporting forces of the beam. This takes place at a cross-head displacement defined by the ratio of the magnitude of the preload and the stiffness of the beam. Beyond this point, ongoing temperature increase will lead to a compressive force in the truss elements and accordingly to supporting forces of the beam exceeding in the sum the preload magnitude P_0. When the truss element BC starts buckling, the cross-head displacement decreases, until the compressive force in the truss elements has unloaded completely and the beam alone balances again the entire sustaining load P_0.

The benchmark problem can be used to verify a consolidated model, whose numerical part consists of the beam, while the heated truss elements represent the physical sub-model. Figure 1e shows the experimental test setup of the physical sub-model, which is identical with the setup of the physical model, with the exception that the connecting rods are omitted, since the action of the beam is modelled numerically. Performing a consolidated test with the same mechanical and thermal loading as in the physical test, allows to validate the consolidated model and therefore verify the method for applications in structural fire engineering. For further studies investigating the influence of (1) the heating rate, (2) the stiffness ratio of the truss element and the beam or (3) the level of the preload, the physical sub-model is reduced to the truss element BC and the displacement, u, measured directly on the specimen inside the furnace is set equal to the midspan deflection, w, of the beam in the numerical model.

2.2 Setup for consolidated thermo-mechanical testing

Figure 2a illustrates the thermo-mechanical consolidated test setup. Its two main components are a numerical and a physical sub-model. The physical sub-model consisted of a steel tensile coupon specimen that could be heated with a split-tube electric furnace (manufacturer: Könn) in a universal testing machine (manufacturer: ZWICK). The numerical model was developed in the FEM-software ABAQUS Standard which features user defined elements. A user defined element is an interface that allows to script a subroutine for the elemental calculations (nodal forces, element stiffness). This subroutine is called in every Newton-Raphson iteration. During such a subroutine call, control is handed over to the user and the FEM-program is put on hold upon completion of the subroutine. This offers the possibility to update the nodal forces and the

element stiffness of the user defined element with actual data of the physical sub-model. The necessary communication between the subroutine and the testing machine is facilitated via a server developed by the authors. Most thermo-mechanical problems in structural fire engineering can be treated as sequentially coupled problems, where only the thermal solution influences via temperature-dependent material properties the mechanical solution. Therefore, an increment of a sequentially coupled thermo-mechanical problem can be split into a sequence of a thermal step followed by a mechanical step, which is illustrated in Figure 2a with the two outermost bounding boxes.

At the beginning of a new increment, first the thermal step is performed before any mechanical equilibrium iterations are executed. Therefore, the target temperature is updated and sent via the server to the furnace. During the heating phase the testing machine stands by in a force hold mode and the server monitors whether the target temperature has already been reached. As soon as the target is reached the server requests the current specimen displacement and force and forwards the data to the subroutine. This terminates the thermal step and initiates the mechanical equilibrium iterations.

At the start of the mechanical step, the current estimate for the displacement is sent to the testing machine as new target, u_{target}. The testing machine starts ramping from the previous position, $u_{target,old}$, to the current target. This can be seen in Figure 2b which shows the evolution in time of the specimen displacement, u, together with the current and previous target values, indicated with stepped solid and dashed lines respectively. While approaching the target position, a limit monitoring procedure is activated, to provide the subroutine with an intermediate specimen displacement and force data set. This is triggered as soon as the displacement exceeds, $u_{intermediate}$, (dotted lines). This mechanism is illustrated in Figure 2c with the displacement command, $u^{j=1}$, of the first iteration, leading from point A to the target position in point C. When the specimen displacement reaches the intermediate position in point B, the current displacement, u_{actual}, and force, F_{actual}, are forwarded to the subroutine. While the testing machine is still approaching the target, the tangent stiffness, $K_{tan}^{j=1}$, is updated using the intermediate data set, the nodal force in point C' is extrapolated, and a solver pass is executed if the equilibrium check is not satisfied. Therefore, the new estimate for the second Newton-Raphson iteration, $u^{j=2}$, is already available, by the time when the testing machine reaches the target in point C. This procedure is repeated for the second (C to E) and third iteration (E to G). At the beginning of iteration 4 the displacement estimate, $u^{j=4}$, is reduced with respect to the precedent iteration. This would lead to spurious unloading of the specimen in case if it was used as a displacement command. To prevent this from happening, no displacement command leading to a reduction of the actual specimen displacement is sent to the testing machine and the nodal force in point H' is extrapolated, using the previous stiffness

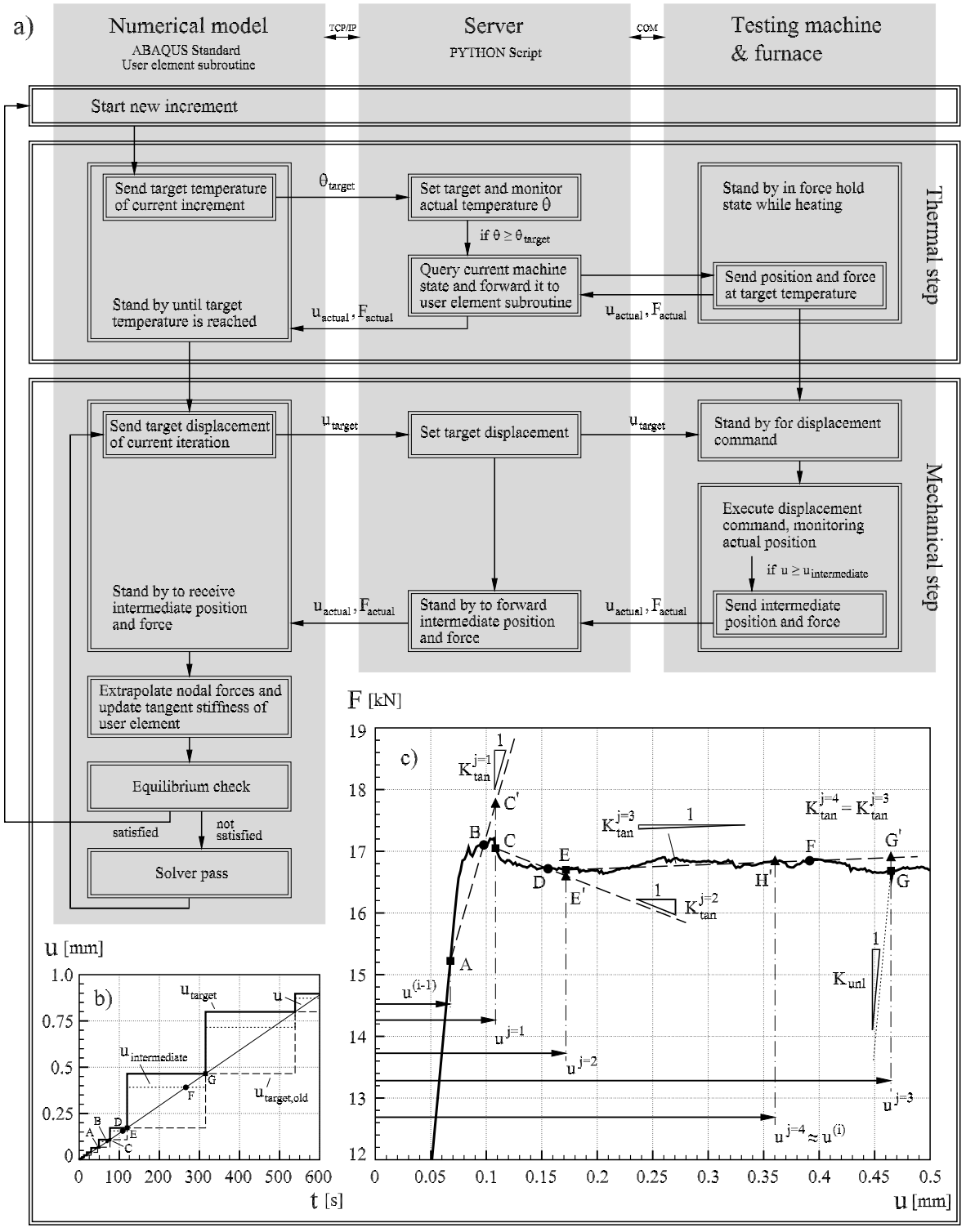


Figure 2: Consolidated thermo-mechanical test setup. (a) Solution procedure of thermo-mechanical increment in consolidated model, (b) displacement commands evolution with time, (c) illustration of tangent stiffness update algorithm with experimental test data.

3. PRELIMINARY EXPERIMENTAL RESULTS

Two preliminary coupled experimental and numerical tests were performed at ambient temperature to proof the eligibility of the consolidated testing framework. The preliminary tests showed different initial load distributions between the truss element and the beam. In the first case 90% of the total applied load would be carried initially (elastic range) by the truss element, $\rho=0.9$, whereas in the second case the truss element would only take 60% of the total load, $\rho=0.6$. The truss element was represented with tensile coupon specimens from an SHS 160·160·5 section of steel grade S355 (mild steel) [3]. Given the elastic stiffness of the truss element (230 kN/mm) the values of the beam stiffness were determined accordingly. The consolidated model results are presented in Figure 3a to Figure 3b, which show in sequence as a function of displacement, the total load, $F_{total,ABAQUS}$, the load per component, $F_{truss,ABAQUS}$, and, $F_{beam,ABAQUS,}$ and the truss element force recorded during the experiment, $F_{truss,exp}$. Furthermore, the load distribution, ρ, referring to the second vertical axis is plotted against the

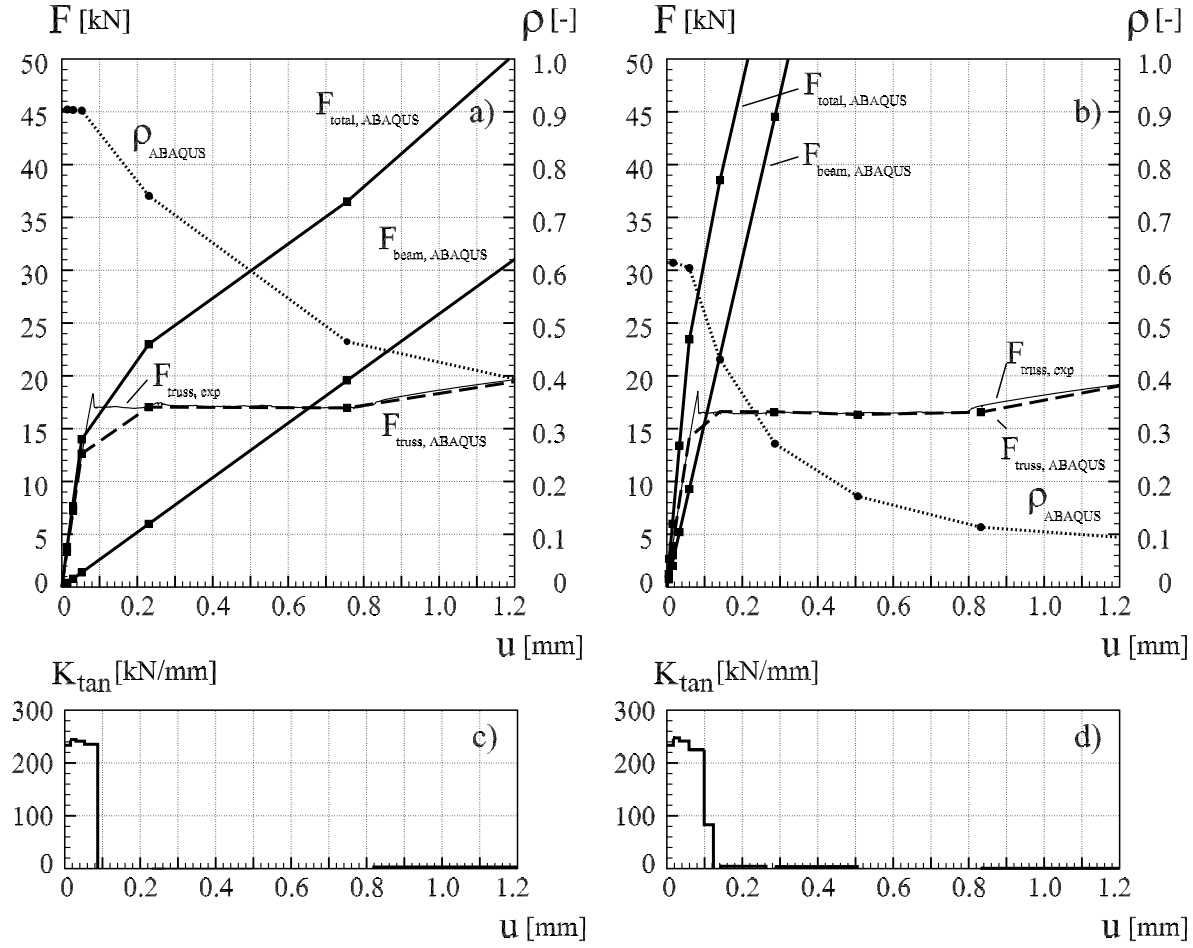


Figure3: Consolidated tests at ambient temperature. Force–deformation and load-share–displacement curves for ρ=0.9 (a) and ρ=0.6 (b), Tangent stiffness–deformation curves for ρ=0.9 (c) and ρ=0.6 (d).

displacement. The data points of converged increments of the consolidated model are denoted with markers. The forces in the two load-carrying components, $F_{truss,ABAQUS}$, and , $F_{beam,ABAQUS}$, always balance the externally applied load, $F_{total,ABAQUS}$. Additionally, the measured force in the truss element equals the force of the truss element in the consolidated model. This demonstrates, that the consolidated test was successful and represents accurately the structural behavior. The difference between the two lines at the beginning of the yield plateau is a consequence of the auto-incrementation feature, which was used in both consolidated tests. Reducing and fixing the increment size could mitigate the discrepancy of the two curves, at the cost of more iterations and calculation time. The change of load-share taken by the truss element throughout the consolidated test is illustrated with the ρ–u curves. The initial load distribution in the elastic range is in agreement with the chosen stiffness ratio between the truss element and the beam. As soon as the truss element reaches a plastic state, an additional external force increment can only be carried by the beam, while the displacement in the truss element increases at constant force. Consequently the ρ–u curve decreases once the truss element starts yielding. The plots of the tangent stiffness, given in Figure 3c and Figure 3d indicate that the stiffness update algorithm worked reliably. In both consolidated tests these plots exhibit the expected sequence of a constant stiffness in the elastic range, followed by a complete loss of stiffness during the yield plateau, and an increase in stiffness with the onset of hardening.

4. CONCLUSIONS

- A framework for consolidated thermo-mechanical modelling has been presented. This framework considers tangent stiffness updates in every Newton-Raphson iteration by using measured data of the physical sub-model.
- A benchmark problem was developed to analyse methodological aspects of thermo-mechanical modelling, and to verify the consolidated testing method with sole physical testing.
- The eligibility of the developed consolidated testing framework has been shown with two preliminary consolidated tests at ambient temperature.
- The framework might be extended to be used for realistic analysis of global structural fire behaviour and for the development of more efficient fire design solutions.

REFERENCES

[1] Bailey C, Both K, Burgess I, Dowling J, Kirby B, Kruppa J et al. *The behaviour of multi-storey steel framed buildings in fire*. Rotherham (1999).

[2] Fontana M, Knobloch M and Frangi A. Global structural behaviour in fire and consolidated testing of steel. *In: Zingoni E, editor. The Fifth International Conference on Structural Engineering, Mechanics and Computation*. Cape Town, South Africa: Taylor & Francis Group; (2013). pp. 1915-1918

[3] Knobloch M, Pauli J and Fontana M. Influence of the strain-rate on the mechanical properties of mild carbon steel at elevated temperatures. *Materials and Design* (2013) **49**:553-565.

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis (Eds)

DESIGN AND DEVELOPMENT OF TEN-ELEMENT HYBRID SIMULATOR AND GENERALIZED SUBSTRUCTURE ELEMENT FOR COUPLED PROBLEMS

SAEID MOJIRI, XU HUANG, OH-SUNG KWON, CONSTANTIN CHRISTOPOULOS

Department of Civil Engineering, University of Toronto
35 St. George St., Toronto, ON, Canada
E-mail: saeid.mojiri@mail.utoronto.ca, xu.huang@mail.uotoronto.ca,
os.kwon@utoronto.ca and c.christopoulos@utoronto.ca

Key words: Multi-platform Analysis, Coupled systems, Substructuring, Hybrid simulation, Pseudo-dynamic testing, Dynamic testing.

Abstract. The University of Toronto's ongoing work towards the improvement and further development of multi-platform simulation methods is presented in this paper. These developments include a ten-element hybrid simulator and a generalized OpenSees substructure element. Hybrid simulation which encompasses development of new integration algorithms, simulation frameworks, and applications has been an active research area in the past decade. Yet, unless the physically tested element significantly contributes to the overall lateral response of the structure in terms of stiffness, strength and energy dissipation, the improvement in the accuracy of results from the use of hybrid simulation may only be marginal. In most cases, the number of physically tested elements in the hybrid simulation is limited by the availability of experimental resources such as actuators, controllers, and laboratory space. As a step towards overcoming this limitation, a novel experimental apparatus, the UT10 Hybrid Simulator, is being developed at the University of Toronto. The UT10 is being developed to allow up to ten elements, such as braces and hysteretic dampers, to be concurrently tested and integrated into a hybrid simulation. The system can test up to ten physical specimens with peak force capacity ranging between 800 kN or 1,600 kN per specimen depending on the total number of tested specimens. The main design requirements, current development status, and potential applications of the UT10 Hybrid Simulator are presented in this paper. To integrate the potential numerical or physical substructures into distributed multi-platform simulations such as hybrid simulations, a generalized substructure element is being developed for OpenSees. The main research focus in this development is to standardize the data exchange format and communication protocol such that any other potential experimental and numerical substructure modules can be readily integrated into the simulation. The data exchange format is defined such that the number of degrees of freedom, data type, and error checks can be communicated in a seamless manner between modules. Designing a versatile data exchange format and a communication protocol is expected to facilitate simulation of coupled systems including diverse substructure modules and other loading scenarios such as thermal loading. The data exchange format and example implementations will be made available to the research community in the near future. To illustrate the current developments, an example of multi-platform simulation with a numerical substructure is presented in this paper.

1. INTRODUCTION

The inelastic dynamic response of structures under earthquake loads is a complex problem, which involves complicated mechanical response of structural elements and their materials. In addition, the recent development of performance based seismic design methods require performance assessment of a structure against multiple seismic hazard levels. Thus, the research community recognizes the need for more accurate and reliable seismic response prediction tools for complex structures subjected to seismic loading.

Although there have been considerable advances in the capabilities of numerical simulation platforms, most of the time a single platform is unable to handle all types of material behavior, loading conditions, and boundary configurations and therefore specialized software is required if a further detailed response is required for particular structural elements. Furthermore, since solving for the response of a complex and large structure with a similar level of model complexity for all elements is computationally expensive and even sometimes impossible, it may be preferred to use sophisticated models only for critical parts of the structure. Despite the advances in computational power and numerical techniques, physical testing, arguably the most realistic source of data, is still essential for validation and calibration of numerical elements. Therefore, structural analysis methods that can combine different numerical platforms and experimental components and benefit from advantages of each during the analysis are invaluable for more realistic response prediction of structures under earthquake loads. One approach to achieve this goal is by dividing a large structure into smaller substructures and model each substructure by the most suitable numerical or experimental platform with the desired level of model complexity. In these multi-platform simulation methods, the responses of different substructures are coupled during the analysis [1,2].

Numerical-experimental hybrid simulation, also known as pseudo-dynamic (PsD) or online testing, is a laboratory dynamic test method which incorporates realistic experimental behaviour of the critical structural components into the numerical model of a structural system, thereby enhancing the accuracy in the response predictions of the system under earthquake loads. In this method some structural components, mainly the most critical ones, can be tested as physical substructures, while the rest of the structure is modeled in a computer as a numerical substructure. The numerical substructure contains dynamic properties of the system like damping and mass. The analysis of the model involves stepwise time integration loops. In each time step, the displacements of the physical substructures due to earthquake loads are numerically predicted in the computer model and are applied quasi-statically on the physical substructures using hydraulic actuators in the laboratory. The forces and displacements from the physical substructures are then measured and sent back to the numerical model and are used to correct the predicted displacements. This process is repeated in each time step. Because the displacements on the physical substructures are applied with an expanded time scale, the PsD test is mainly suitable for physical substructures with rate independent mechanical behaviour [3].

Since its early development in 1969 [4] and 1980s [5,6], numerical-experimental hybrid simulation has evolved considerably in different aspects such as integration algorithms, simulation frameworks, actuator delay compensation methods, and applications. The hybrid simulation method has been successfully implemented in many applied research projects for

testing a variety of structural components [7,8,9,10]. However, in cases where the global response of the structure is of interest, this improvement is not necessarily considerable, unless the behaviour of the physical substructures dominates the global response of the structure. Furthermore, if there are many structural components considerably affecting the response of the structure, a sufficient number of these structural components should be tested as physical substructures. In most cases however, the number of physically tested components in hybrid simulations is limited by the availability of experimental resources in the laboratory such as actuators, controllers, and laboratory space. To overcome this limitation, a new experimental platform, *UT10 Hybrid Simulator,* is being developed for hybrid simulation in the Structural Testing Facilities at the University of Toronto. This simulator is capable of performing hybrid simulations with up to 10 large-scale uniaxial physical substructures such as braces and hysteretic dampers. Development of this facility can significantly affect the accuracy of the hybrid simulation by making it possible to include significantly more physical substructures into the simulations and therefore incorporate more realistic experimental data into the model. Based on the authors's knowledge, this facility will be the first of its kind when considering the number of physical substructures.

In order to facilitate multi-platform simulations using substructuring techniques, a new element named *SubStructure* is being developed for the finite element analysis software, OpenSees [11]. This element is developed for general purpose such that it can be used to represent either experimental or numerical substructures regardless of the number of nodes and degrees of freedoms. The *SubStructure* element is based on the *GenericClient* element which has been implemented in OpenSees to enable the data exchange between the numerical module and the physical substructure modules through the interface program OpenFresco [12]. The original *GenericClient* element is compatible with only four types of experimental elements in OpenFresco namely the Truss, Beam-Column, Two-Node Link, and Inverted-V Brace elements. In addition, it is difficult to use the *GenericClient* element for other substructure modules and different loading scenarios. For instance, in order to include thermal loads in a hybrid simulation, the data exchange format and the communication protocol need to be modified accordingly. To make substructure hybrid simulation methods more approachable, a standardized data exchange format and a communication protocol are being developed for the generic *SubStrucutre* element. The implementation of this versatile data exchange format can not only facilitate maintenance and extension of the hybrid simulation framework for developers but also help users with similar configuration inputs for diverse substructure modules.

Numerical-experimental hybrid simulation requires the implementation of two main modules: 1) The experimental module which is comprised of the physical substructures and the experimental platform including actuators to apply command displacements calculated by the numerical module on the physical substructures, control and communication platform to communicate between the numerical module and the actuators, and required frames to support and hold the specimens in place and 2) the numerical module which is the numerical substructure modeled in a software platform. The UT10 Hybrid Simulator is an experimental platform enabling hybrid simulations on uniaxial physical substructures. Different components of the UT10 Hybrid Simulator and the challenges involved for design and development of each component are described in the next section. The details of the development of the *SubStructure* element used in the numerical module are discussed in

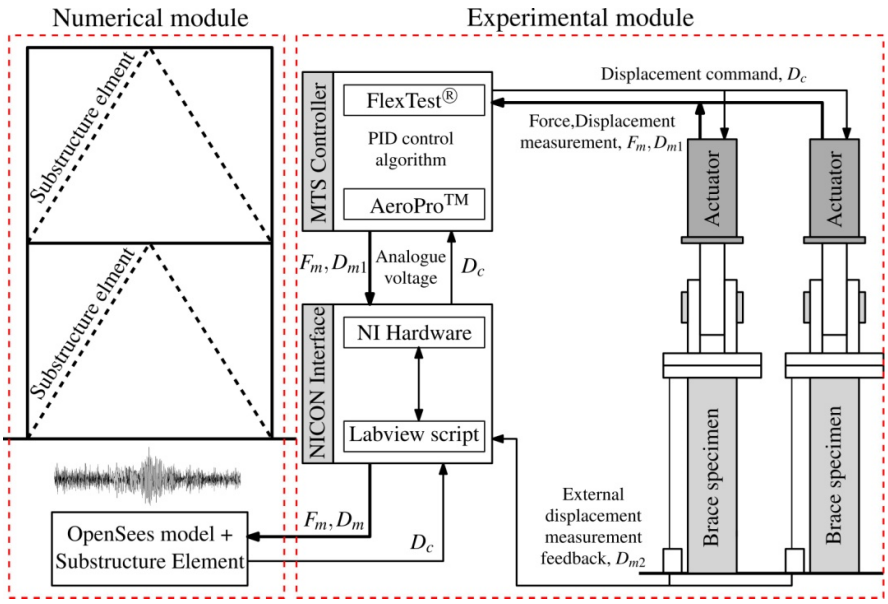


Figure 1: Flow of communication between components of hybrid simulation.

section three. A schematic showing both modules and the communication flow between their components is given in Figure 1.

2. DEVELOPMENT OF UT10 HYBRID SIMULATOR

2.1 Specimens

The UT10 Hybrid Simulator is designed for performing hybrid simulations with up to 10 large-scale uniaxial physical substructures with rate independent hysteretic properties hereafter referred to as specimens. These specimens can be simple steel braces, Buckling Restrained Braces (BRBs), friction dampers, Self Centering (SC) braces, etc. The specimens will be pinned at both ends and will be loaded in their axial direction. The hybrid simulator provides space and axial loading capacity for testing up to 10, 800 kN and 5, 1600 kN large scale specimens with a maximum length of 1,660 mm, simultaneously.

2.2 Actuators and specimen support frame

UT10 Hybrid Simulator uses the existing Shell Element Tester (SET) located in Structures Testing Facility at the University of Toronto. The SET was originally developed to study the behaviour of large scale reinforced concrete (RC) shell elements under various loading configurations. The SET is equipped with 40, 1000 kN in-plane and 20, 500 kN out-of-plane actuators. Figure 2.a shows the SET with a RC shell specimen.

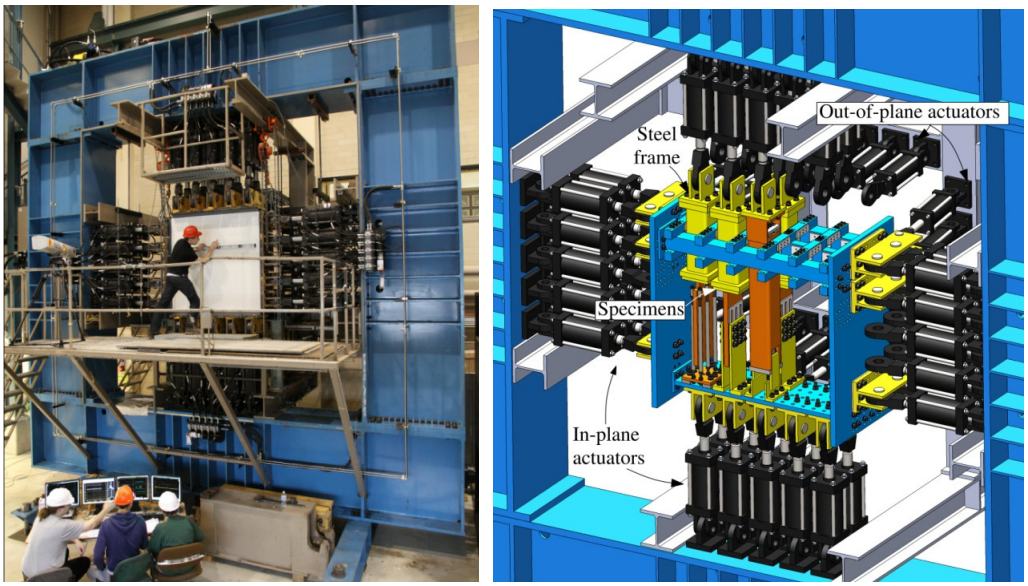


Figure 2: a) The Shell Element Tester (SET) with RC shell specimen in the lab, b) 3D illustration of the UT10 Hybrid Simulator steel frame with 4 specimens inside SET

For the purpose of hybrid simulations, uncoupled axial displacements are applied on 10 separate specimens, and hence, only ten in-plane vertical actuators at the top of the SET are used to control the specimens as shown in Figure 2.b. The other ten actuators at the bottom are used as fixed supports. Because both the top and bottom actuators are pinned at their both ends, the whole system develops a mechanism with four hinges and becomes unstable. To prevent this, the specimens are supported laterally in both the in-plane and out-of-plane directions while allowed to move freely in the vertical (axial) direction. A steel frame was designed and fabricated for the purpose of enabling the connection between the specimens and the in-plane vertical actuators for loading in the axial direction. The frame also provides lateral support to the specimens and connects to the horizontal in-plane and out-of-plane actuators to stabilize the system. Figure 2.b shows the 3D illustration of the steel frame with 4 specimens, installed in the SET. Figure 3 shows a 3D illustration of different components of the steel frame. As can be seen from Figure 3 each specimen is laterally supported by adjustable lateral support beams through a loading shaft which can be also part of the specimen. In order to reduce the amount of friction between the lateral support beams in the steel frame and the loading shaft when the specimen is moving in the axial direction, low friction PTFE sheets are used at the interface of the lateral support beams and the exterior of the loading shafts.

The axial deformations in UT10 Hybrid Simulator are applied by top vertical in-plane actuators that are pinned to the loading yokes which are connected to the loading shafts. The specimens are axially supported at their base by a base plate. The whole system is supported in all directions by in-plane and out-of-plane actuators that are pinned to the yokes which are in turn connected to the side plates and the base plate (see Figure 3). The loading shafts can be rectangular or square hollow structural sections (HSS) and the design of the steel frame is flexible to accommodate different sizes of loading shafts. The maximum size of the shaft that can be accommodated is 508 mmx254 mm.

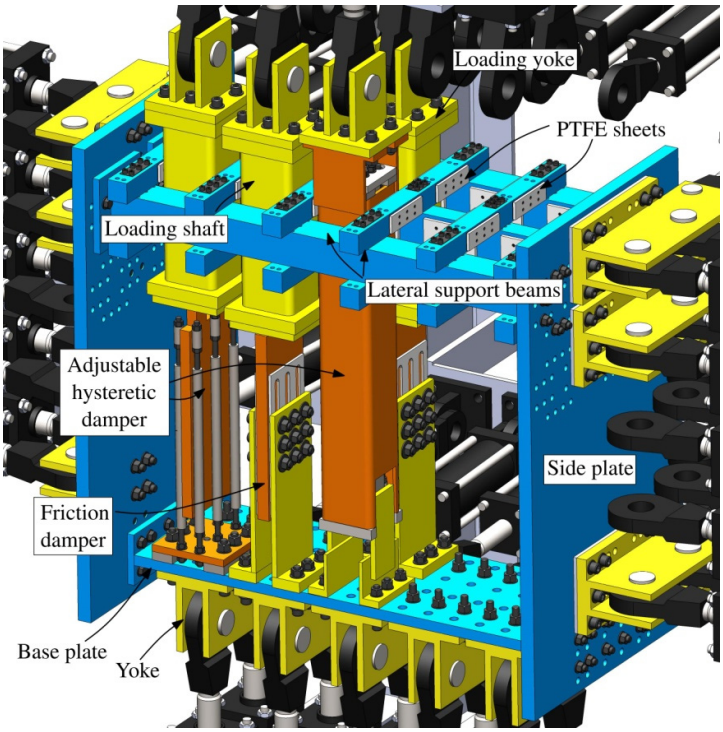


Figure 3: 3D illustration of different component of the steel frame with friction damper and adjustable hysteretic specimens

2.3 Control and communication platform

All 60 actuators in the SET are controlled simultaneously with an MTS FlexTest® controller and AeroPro^{TM} control software with both force and displacement control capabilities. Additional analogue to digital input and digital to analogue output cards are installed on the existing MTS controller in order to provide external communication with the controller through analogue voltages.

An interface program called Network Interface for Controllers (NICON) has been developed at the University of Toronto to facilitate communication between the numerical module and the MTS controller for hybrid simulations [13]. NICON uses National Instruments (NI) hardware and a LabVIEW script to receive displacement commands from the numerical module, communicate them to the MTS controller, and receive and send the feedback force and displacements from the controller to the numerical module. NICON has some added features like displacement and force limit checks, noise filtering, ramp generation, and coordinate transformation. The initial version of NICON can only handle problems with one independent degree of freedom. An updated version of NICON which is able to handle problems with several independent single degrees of freedom (SDOF) (which is needed for the UT10 Hybrid Simulator) and also for problems with multiple coupled degrees of freedom is being developed and is planned to be used in the UT10 Hybrid Simulator [14]. The communication between the controller and NICON is achieved with analogue voltages. Data flow between different components of the UT10 Hybrid Simulator is

shown in Figure 1.

It is predicted that the amount of axial movement of the actuators can be different from the real axial displacements transmitted to the specimens and this is one of the major challenges for the control of the actuators in the UT10 Hybrid Simulator. These differences are mainly due to the elastic deformations of the actuator reaction frames and their connections and also the slackness present in the pin connections between the actuators and loading yokes. One solution to correct for these differences and to ensure that the command displacements are imposed on the specimens with acceptable accuracy is that the actual movement of the specimens be measured externally (D_{m2} in figure 1) and fed back to NICON where they are compared to the actuator displacement feedbacks from the MTS controller (D_{m1} in figure 1) based on which the command displacements to the actuators (D_c) can be corrected. The implementation of this error correction scheme is still in progress.

3. DEVELOPMENT OF SUBSTRUCTURE ELEMENT

3.1 Implementation and characteristics

A generic *SubStructure* element is being built upon an object-oriented, open source software framework, OpenSees. A key feature of OpenSees is the ability to allow user-defined elements to be integrated into the application without the need to change the existing code. Therefore, the development and maintenance of the *SubStructure* element is independent from the main OpenSees software. The *SubStructure* element has the following characteristics:

- It does not have geometry and material descriptions and is only defined by the connected nodes and the number of degrees of freedom.
- It allows for the integration of any number of nodes and degrees of freedom for the substructure.
- All degrees of freedom are defined in the global coordinate systems and hence no coordinate transformation is required in OpenSees. Coordinate transformation for the substructure modules connected through the element is automatically handled,
- The represented substructure can have several structural elements.

The main contribution in development of the *SubStructure* element is the implementation of a standardized data exchange format and a communication protocol. The data exchange format defines a structure of data to be transmitted through network between the integration module which performs the analysis and the numerical and/or physical substructure. It should cover all information needed for various simulation purposes and have flexibility for further extensions. Figure 4 shows the proposed data exchange format which is under development. The data format includes a message header as shaded in the figure and actual data to be sent or received within this communication. Specifically, the *Version* parameter indicates the version of the data exchange format for the purpose of maintenance. The *Command* parameter indicates the communication action which can be sending target displacement to the substructures or asking restoring forces from them. *Test type* parameter indicates whether the simulation is Pseudo-dynamic or real-time. *Substructure type* describes the type of the substructure used in the simulation which can be either numerical or experimental. *Precision* parameter defines the precision of data appended to the message header. *Data type* indicates the type of the appended data which can be displacement, force, velocity, acceleration, and

temperature, or any combination of them. The size of the header is fixed to be 16 bytes while the size of the attached data depends on the parameters defined in the message header. For example, if an experimental truss element is represented by a *SubStructure* element and the parameters of the message header for the communication have been initialized with the *Number of DOFs* parameter of 4 (the total number of DOFs of the substructure), the *Command* parameter of 3 (sending target displacements to the substructure), and the *Precision* parameter of 2 (double precision), then the size of the data appended to the message header will be set to be 32 bytes (4 DOFs×8 bytes) for target displacements. An example using the proposed data exchange format and the communication protocol in a multi-platform simulation with numerical substructure is shown in the following section. More details on the data exchange format will be released to the research community in upcoming publications.

The *SubStructure* element is planned to be used with UT10 Hybrid Simulator for running numerical-experimental hybrid simulations. For this purpose, this element will be used to represent uniaxial physical substructures in the numerical module providing the connection and communication with the experimental modules.

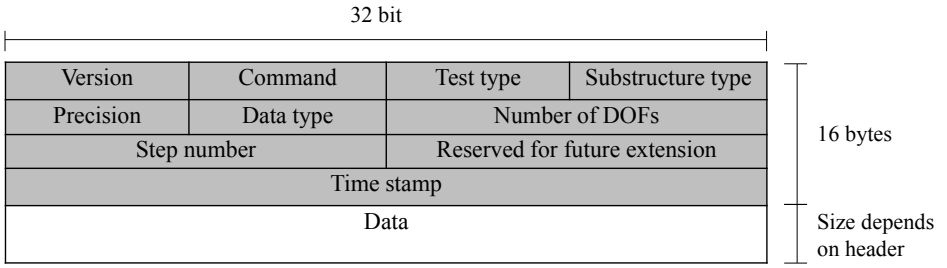


Figure 4: Standardized data exchange format

3.2 Test example

In order to verify the functionality of the new OpenSees *SubStructure* element, a 6 story simple steel frame equipped with Buckling Restrained Braces (BRBs) in concentric chevron configuration was modeled in OpenSees and its response under a ground motion was evaluated using nonlinear time history analysis. The ground motion was the scaled (SF = 1.89) 1994 Northridge earthquake recorded at Canoga Park-Topanga Can station with scaled peak ground acceleration of 0.68g. The record was selected from PEER NGA strong motion database [15]. The frame constitutes the main lateral load resisting system in one direction for a 6 story building located in Los Angeles where the earthquake loads are expected to govern the design of the structure for lateral loads. This building is designed and studied by Choi et al. [16].

For the analysis of the structure, it is divided into two parts as indicated in Figure 5. The 10 BRBs in the first 5 stories are modeled as a single numerical substructure in a separate OpenSees platform while the rest of the structural elements are modeled in the main OpenSees platform. The numerical substructure is represented by the *SubStructure* element that was developed in the main OpenSees platform. The frame is analyzed once using one OpenSees platform without substructuring (whole structure model) and once with two OpenSees platforms with one numerical substructure as explained above. Figure 6 shows the first 30 seconds of history of the first floor interstory drift ratio for both analyses. As can be

seen from this figure, the results completely match for both analyses showing that the *Substructrue* element is working properly.

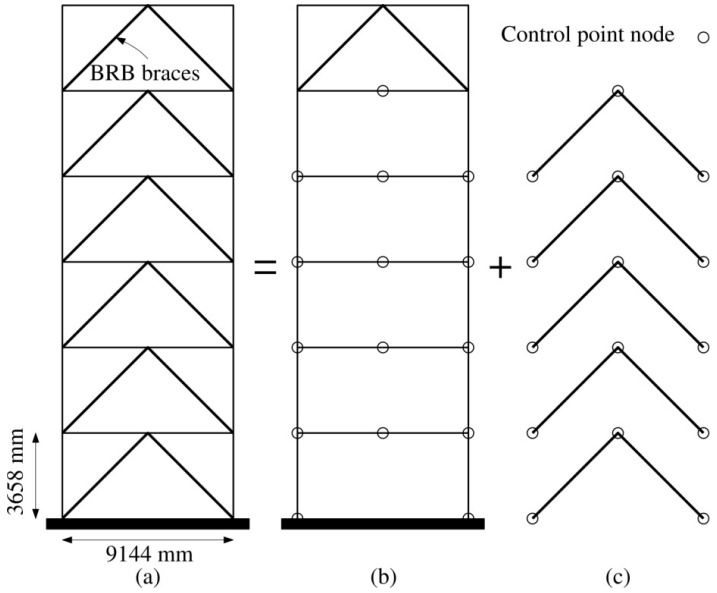


Figure 5: Six story frame model. a) Full frame with BRB braces, b) Main OpenSees model, c) Numerical substructure model

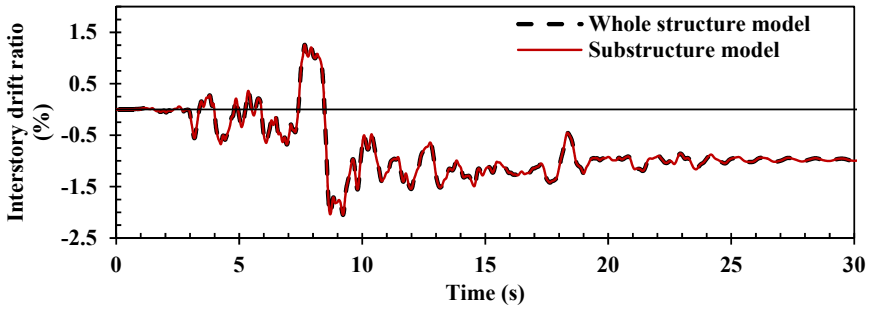


Figure 6: History of first floor interstory drift ratio for the 6 story frame

4. FIRST PLANNED RESEARCH PROJECT APPLICATIONS

4.1 Investigation of response of structures with hysteretic damper braces

Hysteretic energy dissipating braces are one of a number of effective devices for enhancing the seismic performance of structures [17]. The application of these devices is steadily increasing in North America. Several research projects have recently been completed at the University of Toronto on the development and deployment of novel hysteretic energy dissipation devices for use in high seismic performance steel structures [18,19]. Hybrid simulations are planned using the UT10 Hybrid Simulator to investigate the effect of the real properties of various hysteretic damper systems on the accuracy of commonly used numerical

models to predict the seismic performance of these structures. For this purpose and as a first step, the seismic performance of a 6-story steel structure equipped with BRBs will be evaluated using the UT10 Hybrid Simulator. In this study large scale BRB specimens will be used as specimens. In the second step, reusable specimens with adjustable hysteretic behaviour will be designed, fabricated, and used as the physical substructures in hybrid simulations. These specimens will be reusable and their important hysteretic parameters, like the post yield stiffness and self-centering capabilities, will be adjustable making it possible to perform experimental parametric studies. These specimens will be used to represent different types of hysteretic damper braces in the hybrid simulations. This experimental parametric study will not only help in verification of performance predictions inherent in existing design procedures, but will also provide a more realistic understanding from the behaviour of structures equipped with hysteretic damper braces.

4.2 Selection strategy for physical substructures

There are many cases in hybrid simulations where the number of critical structural elements that can be considered as physical substructure is more than what can be experimentally tested in the laboratory. In these cases, a strategy is required for the most efficient selection of physical substructures (number and location) which has the maximum effect on the response prediction of the structure. Preliminary studies of this kind have been completed at University of Toronto [13] and it was observed through experimental investigations on braced frames that the system-level response prediction of the structure can be significantly changed by changing the location of the brace (first story brace, second story brace, etc) that is represented by the physical substructures. A preliminary selection strategy for shear type buildings is also presented. More comprehensive studies of this kind using UT10 Hybrid Simulator with several specimens are planned.

5. CONCLUSIONS

The University of Toronto's ongoing projects for development and improvement of multi-platform simulation methods was presented in this paper. These developments include a novel experimental platform called the UT10 Hybrid Simulator which is capable of performing hybrid simulations on up to 10 uniaxial structural elements simultaneously. This facility uses the existing Shell Element Tester (SET) facility in the structural laboratory for applying the required loads on the specimens. An improved version of the NICON interface is developed to facilitate communication of the numerical models with the existing MTS FlexTest actuator controller. A new OpenSees element called *SubStructure* was also developed which represents the numerical or experimental substructures in the multi-platform simulations. This generic element provides a newly developed standardized data exchange format and a communication protocol for communication between the integrating numerical module and other numerical and/or physical substructures. The functionality of this element was verified by a multi-platform simulation with a numerical substructure. The developments are still ongoing and future experimental-numerical hybrid simulations are planned in the near future using the new developed systems.

REFERENCES

[1] Kwon, O., and Elnashai, A. Seismic analysis of Meloland road overcrossing using multiplatform simulation software including SSI. *Journal of Structural Engineering* (2008) 134, No. 4: 651-660.

[2] Kwon, O., Elnashai, A., and Spencer, B. F. A framework for distributed analytical and hybrid simulations. *Structural Engineering and Mechanics* (2008) 30, No. 3: 331-350.

[3] Saouma, V., and Sivaselvan, M. *Hybrid simulation: Theory, implementation and applications*. Taylor and Francis Group, London, UK, (2008).

[4] Hakuno M., Shidawara M., and Hara T. Dynamic destructive test of a cantilever beam controlled by an analog-computer. *Proceedings of the Japan Society of Civil Engineers* (1969) 171: 1–9.

[5] Takanashi, K., and Nakashima, M. Japanese activities on on-line testing. *Journal of Engineering Mechanics* (1987) 113, No. 7: 1014-1032.

[6] Mahin, S. A., and Shing, P. B. Pseudodynamic method for seismic testing. *Journal of Structural Engineering* (1985) 111, No. 7: 1482-1503.

[7] Negro, P., Mola, E., Molina, F. J., and Magonette, G. E. Full-scale PSD testing of a torsionally unbalanced three-storey non-seismic RC frame. In *Proceeding of 13th World Conference on Earthquake Engineering*, Vancouver, BC, Canada, (2004), No. 968.

[8] Yang, T. Y., Stojadinovic, B., and Moehle J. Hybrid simulation of a zipper-braced steel frame under earthquake excitation. *Earthquake Engineering & Structural Dynamics* (2009) 38, No. 1: 95-113.

[9] Elnashai, A., Spencer, B., Kim, S. J., Holub, C., and Kwon, O. Hybrid distributed simulation of a bridge-foundation-soil interacting system. In *4th International Conference on Bridge Maintenance, Safety, and Management,* Seoul, Korea. (2008).

[10] Kammula, V., Erochko, J., Kwon, O., and Christopoulos, C. Application of hybrid-simulation to fragility assessment of the telescoping self-centering energy dissipative bracing system. *Earthquake Engineering & Structural Dynamics* (2014) 43, No. 6: 811-830.

[11] Mazzoni, S., McKenna, F., Scott, M. H., and Fenves, G. L. OpenSees command language manual. *Pacific Earthquake Engineering Research (PEER) Center* (2006).

[12] Takahashi, Y., and Fenves, G. L. Software framework for distributed experimental–computational simulation of structural systems. *Earthquake engineering & structural dynamics* (2006) 35, No. 3: 267-291.

[13] Kammula, V. Application of hybrid simulation to fragility assessment of self-centering energy dissipative (SCED) bracing system. MASc dissertation, Civil Engineering Department, University of Toronto, Toronto, Canada, (2013).

[14] Zhan, H. Development of a generic controller interface program NICON for multi-degrees of freedom PSD hybrid simulation. MASc dissertation, Civil Engineering Department, University of Toronto, Toronto, Canada, (2014).

[15] Pacific Earthquake Engineering Research Center: NGA Database. (Accessed December, 2013). http://peer.berkeley.edu/peer_ground_motion_database.

[16] Choi, H., Erochko, J., Christopoulos, C., and Tremblay, R. Comparison of the seismic response of steel buildings incorporating self-centering energy-dissipative dampers, buckling restrained braces and moment resisting frames. Rep. No. 05-2008, Dept. of Civil Engineering, University of Toronto, Toronto, Ontario, (2008).

[17] Christopoulos, C., Filiatrault, A. *Principles of passive supplemental damping and seismic isolation*. IUSS Press, Milan, Italy, (2006).

[18] Gray, M. G., Christopoulos, C., and Packer J. A. Cast steel yielding brace system for concentrically braced frames: concept development and experimental validations. *Journal of Structural Engineering* (2013) 140, No. 4.

[19] Erochko, J., Christopoulos, C., and Tremblay, R. Design, testing, and detailed component modeling of a high-capacity self-centering energy-dissipative brace. *Journal of Structural Engineering* 10.1061/(ASCE)ST.1943-541X.0001166 (2014).

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis(Eds)

HETEROGENEOUS ASYNCHRONOUS TIME INTEGRATORS BUILT FROM THE ENERGY METHOD FOR COUPLING NEWMARK AND α-SCHEMES

M. BRUN*, A. GRAVOUIL† AND A. COMBESCURE†

*Université de Lyon, INSA-Lyon, LGCIE, 34 rue des Arts, F-69621 Villeurbanne, France

†Université de Lyon, INSA-Lyon, LaMCoS, CNRS UMR5259, 18-20 rue des sciences, F-69621 Villeurbanne, France.
e-mail: michael.brun@insa-lyon.fr - Web page: ttp://lamcos.insa-lyon.fr/

Key words: Structural dynamics, Heterogeneous Asynchronous time integrators, Hybrid multi-time methods, Energy conserving methods

Abstract. The time integration procedure selected in computational structural dynamics must possess at least the stability and accuracy properties required for the convergence to the exact solution. Other desired properties are the unconditional stability for linear dynamics, second-order of accuracy, high frequency dissipation capabilities, self-starting, no overshoot, one step method and no more than one set of implicit equations to be solved for each time step (single-step-single-solve format). In linear dynamics, the stability is classically assessed by a spectral study of the amplification matrix, whereas physical energy bounds are preferred in nonlinear dynamics. Popular a-schemes (HHT-α, WBZ-α, CH-α) are second-order accurate and provides numerical dissipation for spurious high frequencies due to the finite element discretization. To go beyond the standard approach based on the same time integration scheme (homogeneous time integration scheme) and the same time step for all the finite elements of the mesh (synchronous time integration), the purpose of this paper is to describe a general methodology for building Heterogeneous (different time integration schemes such as Newmark or a-schemes) Asynchronous (different time steps) Time Integrators (HATI) for computational dynamics. The key point for building the HATI methods is to cancel the interface pseudo-energy as introduced by Hughes in the so-called energy method employed for proving the stability of implicit-explicit algorithms in its pioneer works on heterogeneous time integrators. By canceling the pseudo-energy at the interface between subdomains and assuming a linear time variation of the Lagrange multipliers at the coarse time scale, the HATI method, called BCG-macro method, is derived. It can handle any dissipative α-schemes (HHT-α, WBZ-α, CH-α), while preserving the second-order of accuracy when adopting different time steps. In addition to the energy argument (cancelation of the interface pseudo-energy), the stability and order of accuracy is proved by the spectral study of the amplification matrix.

1 INTRODUCTION

Partitioning approaches are very appealing with a view of coupling dynamic parts of a complex structure in a hybrid (experimental/numerical) real time testing. These hybrid tests require different explicit or implicit time integration schemes in different subdomains (called substructure in the context of hybrid testing) according to their complexity and characteristics. The purpose of such hybrid experimental/numerical testing such as the continuous pseudo-dynamic testing (PsD) is to test parts of the structure expected to experience some damage while the remainder of the whole structure is dealt numerically. As a consequence, the hybrid numerical-experimental testing techniques need to select high performance coupling methods from the field of computational mechanics in order to be able to make interact physical and numerical parts in real time (*Pinto et al.* [2004], *Brun et al.* [2012]). Typically, a reduced number of degrees of freedom is considered for the tested part but with a large number of small time steps: explicit time schemes are then adopted or other non-iterative time integration schemes (*Combescure and Pegon* [1997], *Bonnet et al.* [2008]).
In the numerical/experimental hybrid context, heterogeneous (different time schemes) asynchronous (different time steps) time integrators (HATI) have been investigated. Gravouil and Combescure proposed an HATI (*Gravouil and Combescure* [2001] and *Combescure and Gravouil* [2002]), called GC method, enabling explicit/implicit multi-time step computations on a partitioned domain: the non overlapping subdomains are coupled following on a dual approach, that is by prescribing the continuity of the velocity at the interface with the help of Lagrange multipliers. The stability of the GC method when considering any Newmark time integration schemes has been proven using the energy method, initially proposed by Hughes for deriving stability conditions for hybrid (implicit/explicit) time integrators. Beyond the widely used Newmark schemes, it is of great interest in structural dynamics to choose more advanced time integration schemes endowed with the following desirable properties: unconditional stability, second-order accuracy, no more than one set of implicit equations involved in the time step, self-starting, and controllable algorithmic dissipation in the higher modes. An excellent state of the art on the generalized single-step-single-solve (GSSSS) framework algorithms that encompasses the class of liner multi-time step methods (LMS) can be found in the works of Tamma and co-authors (*Masuri et al.* [2009]).

In this paper, a general methodology based on energy considerations is described to obtain asynchronous coupling methods for different time integration schemes (Newmark, HHT-α, WBZ-α and CH-α)(*Brun et al.* [2015],*Gravouil et al.* [2015]). The gluing of subdomains with their own time integrator and time step is ensured with Lagrange multipliers and velocity continuity. The key point for building stable HATI is to ensure the zero interface pseudo-energy, involved in the energy method for the DAE (Differential Algebraic Equation) stability analysis. On the other hand, the energy method argument is confirmed by carrying out a stability analysis performed from the amplification

matrix linking state vectors at the beginning and at the end of the time step. The derived BGC-macro method, assumes a linear variation of the Lagrange multipliers over the macro-time step so as to lead to a kinematic condition at the interface, which, used together with the Lagrange multiplier variation assumption, guarantees the zero interface pseudo-energy. The method conserves the second order accuracy as highlighted by the convergence analysis.

2 GENERAL METHODOLOGY FOR COUPLING α schemes

2.1 Unified format for the α-schemes

The time discretization is introduced as the following partition of the time domain: $t_0 < t_1 < \cdots < t_n < t_{n+1} < \cdots < t_f$ and h is the time step size assumed to be constant. Following the concept of weak equilibrium, the discretized equilibrium equation is only ensured at some times inside the time step. For the α-schemes, it corresponds to different weights applied to inertial, stiffness and external terms as given below (viscous terms are omitted in the following for the sake of simplicity):

$$\mathbf{M}\mathbf{a}_{n+\xi_m} + \mathbf{K}\mathbf{u}_{n+\xi_f} = \mathbf{f}_{n+\xi_f} - \mathbf{L}^T\boldsymbol{\lambda}_{n+\xi_f} \tag{1}$$

where $\mathbf{M}$, and $\mathbf{K}$, are the consistent mass matrix and the stiffness matrix. $\mathbf{L}$ is a linear kinematic constraint matrix at the boundary conditions. The discrete displacements, velocities and accelerations are designated by $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{a}$. As external forces, the Lagrange multipliers required by the kinematic constraint is expressed at the time $t_{n+\xi_f}$. Using the two different averaged parameters ξ_m and ξ_f, the averaged displacements, velocities, accelerations and Lagrange multipliers read:

$$\begin{cases} \mathbf{a}_{n+\xi_m} = (1-\xi_m)\mathbf{a}_n + \xi_m\mathbf{a}_{n+1} \\ \mathbf{u}_{n+\xi_f} = (1-\xi_f)\mathbf{u}_n + \xi_f\mathbf{u}_{n+1} \\ \mathbf{f}_{ext,n+\xi_f} = (1-\xi_f)\mathbf{f}_{ext,n} + \xi_f\mathbf{f}_{ext,n+1} \\ \boldsymbol{\lambda}_{n+\xi_f} = (1-\xi_f)\boldsymbol{\lambda}_n + \xi_f\boldsymbol{\lambda}_{n+1} \end{cases} \tag{2}$$

The classical form for the Generalized-α scheme proposed by Chung and Hulbert (*Chung and Hulbert* [1993]) is characterized by two weighting parameters α_m and α_f. It writes:

$$\begin{aligned} (1-\alpha_m)\mathbf{M}\mathbf{a}_{n+1} + \alpha_m\mathbf{M}\mathbf{a}_n + (1-\alpha_f)\mathbf{K}\mathbf{u}_{n+1} + \alpha_f\mathbf{K}\mathbf{u}_{n+1} = \\ (1-\alpha_f)\mathbf{f}_{n+1} + (1-\alpha_f)\mathbf{f}_n - \mathbf{L}^T\boldsymbol{\lambda}_{n+1-\alpha_f} \end{aligned} \tag{3}$$

From the above averaged equilibrium equations in Eq. 1 and in Eq. 3, it can be easily seen that the following relationships between the algorithmic parameters hold: $\xi_m = 1 - \alpha_m$ and $\xi_f = 1 - \alpha_f$.

Optimized algorithmic parameters α_m, α_f, γ and β are given by the authors in order to achieve the unconditional stability, the second-order accuracy as well as the high-frequency dissipation while minimizing low-frequency dissipation. Three α-schemes are

Scheme	α_m	α_f	γ	β
$HHT-\alpha$	0	$-\alpha_{HHT}=\frac{1-\rho_\infty}{1+\rho_\infty}$	$\frac{1}{2}-\alpha_{HHT}$	$\frac{1}{4}(1-\alpha_{HHT})^2$
$WBZ-\alpha$	$\alpha_{WBZ}=\frac{\rho_\infty-1}{1+\rho_\infty}$	0	$\frac{1}{2}-\alpha_{WBZ}$	$\frac{1}{4}(1-\alpha_{WBZ})^2$
$CH-\alpha$	$\frac{2\rho_\infty-1}{1+\rho_\infty}$	$\frac{\rho_\infty}{1+\rho_\infty}$	$\frac{3}{2}-2\alpha_f$	$(1-\alpha_f)^2$

Table 1: Relations between the spectral radius at infinity and the algorithmic parameters for the α-schemes enabling unconditional stability, second-order accuracy, optimized dissipative features.

investigated : HHT-α (*Hilber et al.* [1977]), WBZ-α (*Wood et al.* [1981]) and CH-α (*Chung and Hulbert* [1993]). For all the three previous α-schemes, it is convenient to define the α_m and α_f parameters as a function of the the spectral radius at the high frequency limit (at infinity), noted as ρ_∞, characterizing the amount of dissipative energy in the high frequency range. The two other algorithmic parameters γ and β are involved in the classical Newmark formulae written below in terms of velocity increments:

$$\left\{\begin{array}{l}\Delta\mathbf{a}=\frac{1}{\gamma h}\Delta\mathbf{v}-\frac{1}{\gamma}\mathbf{a}_n\\ \Delta\mathbf{u}=\frac{\beta h}{\gamma}\Delta\mathbf{v}+h\mathbf{v}_n+\frac{\gamma-2\beta}{2\gamma}h^2\mathbf{a}_n\end{array}\right. \tag{4}$$

The algorithmic parameters γ and β are expressed in terms of α_m and α_f as it has been resumed in Table 1.

From Eq. 2, the averaged accelerations and displacements can be written as: $\mathbf{a}_{n+\xi_m}=\mathbf{a}_n+\xi_m\Delta\mathbf{a}$ and $\mathbf{u}_{n+\xi_f}=\mathbf{u}_n+\xi_f\Delta\mathbf{u}$. Incorporating theses expressions and the Newmark formulae in Eq. 4 into the equilibrium equation Eq. 1, we obtain a new format for the Generalized-α schemes as:

$$\mathbf{K}^*\Delta\mathbf{v}_{n+1}=\mathbf{g}_{n+1}-\mathbf{L}^T\boldsymbol{\lambda}_{n+\xi_f} \tag{5}$$

where the dynamic operator matrix is defined by $\mathbf{K}^*=\xi_m\frac{1}{\gamma h}\mathbf{M}+\xi_f\frac{\beta h}{\gamma}\mathbf{K}$. The right hand side vector $\mathbf{g}_{n+1}$ is expressed as:

$$\mathbf{g}_{n+1}=\mathbf{f}_{ext,n+\xi_f}-\mathbf{K}\mathbf{u}_n-\xi_f h\mathbf{K}\mathbf{v}_n+\xi_m\frac{1}{\gamma}M\mathbf{a}_n-M\mathbf{a}_n-\xi_f(\frac{\gamma-2\beta}{2\gamma})h^2\mathbf{K}\mathbf{a}_n \tag{6}$$

With a view to formulating the subdomain coupling algorithm, it is interesting to adopt a compact form for the complete-three stage time-stepping of the α-schemes as:

$$\mathbb{K}^*\Delta\mathbb{U}_{n+1}+\mathbb{L}^T\boldsymbol{\lambda}_{n+\xi_f}=\mathbb{G}_{n+1} \tag{7}$$

The vector $\Delta\mathbb{U}_{n+1}$ gathers the increments of velocities, displacements and accelerations: $\Delta\mathbb{U}_{n+1}^T=\left[\begin{array}{ccc}\Delta\mathbf{v} & \Delta\mathbf{u} & \Delta\mathbf{a}\end{array}\right]$. $\mathbb{L}$ is the constraint matrix: $\mathbb{L}=\left[\begin{array}{ccc}\mathbf{L} & \mathbf{0} & \mathbf{0}\end{array}\right]$. The right hand side vector is given by:

$$\mathbb{G}_{n+1}=\mathbb{F}_{n+\xi_f}-\mathbb{N}\mathbb{U}_n \tag{8}$$

where $\mathbb{F}^T_{n+\xi_f} = \left[\begin{array}{ccc} \mathbf{f}_{ext,n+\xi_f} & \mathbf{0} & \mathbf{0} \end{array} \right]$. The expressions of the matrix $\mathbb{K}^*$ and the matrix $\mathbb{N}$ are:

$$\mathbb{K}^* = \begin{bmatrix} \mathbf{K}^* & \mathbf{0} & \mathbf{0} \\ -\frac{\beta h}{\gamma}\mathbf{I} & \mathbf{I} & \mathbf{0} \\ -\frac{1}{\gamma h}\mathbf{I} & \mathbf{0} & \mathbf{I} \end{bmatrix}, \mathbb{N} = \begin{bmatrix} \xi_f h\mathbf{K} & \mathbf{K} & \xi_f \left(\frac{\gamma-2\beta}{2\gamma}\right) h^2\mathbf{K} + \left(1-\xi_m\frac{1}{\gamma}\right)\mathbf{M} \\ h\mathbf{I} & \mathbf{0} & \left(\frac{\gamma-2\beta}{2\gamma}\right) h^2\mathbf{I} \\ \mathbf{0} & \mathbf{0} & -\frac{1}{\gamma}\mathbf{I} \end{bmatrix} \tag{9}$$

2.2 Coupling with heterogeneous time steps

In order to formulate the coupling between two subdomains, we employ the previous unified form of the equilibrium equations at t_{γ_A} for the subdomain Ω_A (macro time scale), and at a series of times $t_{j-1+\gamma_B}$ for the subdomain Ω_B (micro time scale); first, for the macro subdomain Ω_A:

$$\left\{ \begin{array}{l} \mathbf{K}^*_A \Delta\mathbf{v}^A_m + \mathbf{L}^T_A \boldsymbol{\lambda}_{\xi_{A,f}} = \mathbf{g}^A_m \\ \Delta\mathbf{u}^A_m = \frac{\beta_A h_A}{\gamma_A}\Delta\mathbf{v}^A_m + h_A\mathbf{v}^A_0 + \frac{\gamma_A-2\beta_A}{2\gamma_A}h^2_A\mathbf{a}_0 \\ \Delta\mathbf{a}^A_m = \frac{1}{\gamma_A h_A}\Delta\mathbf{v}^A_m - \frac{1}{\gamma_A}\mathbf{a}^A_0 \end{array} \right. \tag{10}$$

Secondly, for the micro subdomain Ω_B:

$$\left\{ \begin{array}{l} \mathbf{K}^*_B \Delta\mathbf{v}^B_j + \mathbf{L}^T_B \boldsymbol{\lambda}_{j-1+\xi_{B,f}} = \mathbf{g}^B_j \\ \Delta\mathbf{u}^B_j = \frac{\beta_B h_B}{\gamma_B}\Delta\mathbf{v}^B_j + h_B\mathbf{v}^B_{j-1} + \frac{\gamma_B-2\beta_B}{2\gamma_B}h^2_B\mathbf{a}_{j-1} \\ \Delta\mathbf{a}^B_j = \frac{1}{\gamma_B h_B}\Delta\mathbf{v}^B_j - \frac{1}{\gamma_B}\mathbf{a}^B_{j-1} \\ \forall j \in \{1, m\} \end{array} \right. \tag{11}$$

The kinematic equation at the interface will complete the system of equations for coupling the two subdomains with heterogeneous time steps. It will be shown in the following that ensuring the zero energy at the interface in terms of the pseudo-energy norm as introduced by Hughes in the energy method (*Hughes* [1987]), enables a kinematic condition for velocity increments to be derived.

3 BGC-MACRO METHOD FROM THE ENERGY METHOD

The energy method proposed by Hughes has been widely used for obtaining the stability conditions for coupling schemes, mixing implicit and explicit schemes (*Hughes* [1987]). For example, the stability of the GC method and the PH method has been proven by this way (*Combescure and Gravouil* [2002], *Prakash and Hjelmstad* [2004]). It consists in proving that the interface pseudo-energy is equal or less than zero. Here, the interface pseudo-energy is employed as the starting point of the new coupling methods by ensuring the zero interface pseudo-energy. Thus, the following coupling schemes are built from the

discrete balance equation given in the energy method in terms of pseudo-energy. For a given subdomain (by omitting the scripts for belonging subdomains), the pseudo-energy balance equation is given by:

$$\left[\frac{1}{2}\mathbf{a}^T\mathbf{A}\mathbf{a}+\frac{1}{2}\mathbf{v}^T\mathbf{K}\mathbf{v}\right]_n^{n+1}=\frac{1}{h}\Delta\mathbf{v}^T\left\{(\mathbf{f}_{ext,n+1}-\mathbf{f}_{ext,n})\right\}-(\gamma-\frac{1}{2})\left\{\Delta\mathbf{a}^T\mathbf{A}\Delta\mathbf{a}\right\} \tag{12}$$

in which the matrix $\mathbf{A}$ is defined by $\mathbf{A}=\mathbf{M}+\left(\beta-\frac{1}{2}\gamma\right)h^2\mathbf{K}$. The previous balance equation can also be noted as:

$$\Delta E_{kin}+\Delta E_{int}=\Delta E_{ext}+\Delta E_{diss} \tag{13}$$

where ΔE_{kin}, ΔE_{int}, ΔE_{ext} and ΔE_{diss} are the increments over the time step of pseudo-energies which can be related to the classical kinetic, internal, external and dissipated energies.

Considering two subdomains Ω_A and Ω_B, the global discrete pseudo-energy balance equation takes into account the above contributions from both subdomains on the macro time step $h_A=[t_0;t_m]$ and on the series of the micro time steps $h_B=[t_{j-1};t_j]$ for j varying from 1 to m, plus an additional term corresponding to the interface energy. The discrete pseudo-energy balance equation over the macro time step $h_A=[t_0;t_m]$ for the whole domain is given below:

$$\Delta E^A_{kin,m}+\Delta E^A_{int,m}+\sum_{j=1}^{m}\left\{\Delta E^B_{kin,j}+\Delta E^B_{int,j}\right\}=\ldots$$
$$\Delta E^A_{ext,m}+\sum_{j=1}^{m}\Delta E^B_{ext,j}+\Delta E^A_{diss,m}+\sum_{j=1}^{m}\Delta E^B_{diss,j}+\Delta E_{interface} \tag{14}$$

The interface pseudo-energy (*Prakash and Hjelmstad* [2004]) can be written as:

$$\Delta E_{interface}=-\frac{1}{h_A}\Delta\mathbf{v}_{\mathbf{m}}^{\mathbf{A}\,T}\left\{\mathbf{L}_A^T(\boldsymbol{\lambda}_m-\boldsymbol{\lambda}_0)\right\}-\sum_{j=1}^{m}\left\{\frac{1}{h_B}\Delta\mathbf{v}_{\mathbf{j}}^{\mathbf{B}\,T}\left\{\mathbf{L}_B^T(\boldsymbol{\lambda}_j-\boldsymbol{\lambda}_{j-1})\right\}\right\} \tag{15}$$

The pseudo-energy method has to be distinguished from the discretized energy balance equation (*Hughes* [1987] and *Brun et al.* [2014]). The pseudo-energy is dedicated to stability analysis whereas the classical energy balance equation enables to build energy-preserving time integrators. It has to be noted that the above equation can be seen as a generalization of the pseudo-energy method related to the study of ODE (Ordinary Differential Equation) stability to the study of DAE (Differential Algebraic Equation) stability, that is with Lagrange multipliers. This stability method based on the energetic view point will be confirmed in the last section by a classical amplification matrix spectral analysis, with state vectors gathering kinematic quantities and Lagrange multipliers. The

coupling method is considered to be stable if the total pseudo-energy variation (kinetic plus internal pseudo-energy variations) over the macro time step under zero external loads is less than or equal to zero. Consequently, it is sufficient to prove that the pseudo-energy at the interface $E_{interface}$ is less than or equal to zero. It has to be reminded that the energy conservative features of the PH method (*Prakash and Hjelmstad* [2004]) has been demonstrated *a posteriori*. In contrast to the PH method , the coupling method here enables to satisfy the energy conservative feature *a priori* by ensuring the zero interface pseudo-energy over the macro time step. For this purpose, an assumption related to the time variation of the Lagrange multipliers over the macro time step is required: we assume that the Lagrange multipliers at the micro scale can be linearly interpolated from the Lagrange multipliers given at the beginning and the end of macro time step. So we have:

$$\boldsymbol{\lambda}_j - \boldsymbol{\lambda}_{j-1} = \frac{\boldsymbol{\lambda}_m - \boldsymbol{\lambda}_0}{m} \tag{16}$$

Consequently, the interface pseudo-energy can be simplified as:

$$\Delta E_{interface} = -\left[\frac{1}{h_A}\Delta \mathbf{v}_m^{A\,T}\mathbf{L}_A^T + \sum_{j=1}^{m}\left\{\frac{1}{mh_B}\Delta \mathbf{v}_j^{B\,T}\mathbf{L}_B^T\right\}\right](\boldsymbol{\lambda}_m - \boldsymbol{\lambda}_0) \tag{17}$$

Finally, the zero interface pseudo-energy requirement leads to the following kinematic equation:

$$\mathbf{L}_A\Delta \mathbf{v}_m^A + \sum_{j=1}^{m}\mathbf{L}_B\Delta \mathbf{v}_j^B = 0 \tag{18}$$

Considering two subdomains Ω_A and Ω_B with different time scales, m being equal to the time step ratio, the interface forces associated with the subdomain Ω_A are computed at the time $t_{\xi_{A,f}} = \xi_{A,f}t_0 + (1-\xi_{A,f})t_m$ belonging to the macro-time step $h_A = [t_0; t_m]$, whereas the interface forces associated with the subdomain Ω_B are computed at the time $t_{j-1+\xi_{B,f}} = \xi_{B,f}t_{j-1} + (1-\xi_{B,f})t_j$ belonging to the micro-time step $h_B = [t_{j-1}; t_j]$ for j varying from 1 to m. By using the linear relationship for Lagrange multipliers, we can write the interface forces for both subdomains as:

$$\begin{cases} \mathbf{L}_A^T\boldsymbol{\lambda}_{\xi_{A,f}} = \mathbf{E}_{A,m}^T\boldsymbol{\lambda}_0 + \mathbf{C}_{A,m}^T\boldsymbol{\lambda}_m \\ \mathbf{L}_B^T\boldsymbol{\lambda}_{j-1+\xi_{B,f}} = \mathbf{E}_{B,j}^T\boldsymbol{\lambda}_{j-1} + \mathbf{C}_{B,j}^T\boldsymbol{\lambda}_j \end{cases} \tag{19}$$

with:

$$\begin{cases} \mathbf{E}_{A,m}^T = (1-\xi_{A,f})\mathbf{L}_A^T \\ \mathbf{C}_{A,m}^T = \xi_{A,f}\mathbf{L}_A^T \\ \mathbf{E}_{B,j}^T = \left[(1-\xi_{B,f})(1-\frac{j-1}{m}) + \xi_{B,f}(1-\frac{j}{m})\right]\mathbf{L}_B^T \\ \mathbf{C}_{B,j}^T = \left[(1-\xi_{B,f})(\frac{j-1}{m}) + \xi_{B,f}(\frac{j}{m})\right]\mathbf{L}_B^T \quad \forall j \in [1,m] \end{cases} \tag{20}$$

Then, the global system of equilibrium equations is obtained by gathering the compact forms of the three-stage time stepping schemes for both subdomains. It writes:

$$\left[\begin{array}{ccccc|c|c} \mathbb{K}_B^* & & & & & & \mathbb{C}_{B,1}^T \\ \mathbb{N}_B & \mathbb{K}_B^* & & & & & \mathbb{C}_{B,2}^T \\ \mathbb{N}_B & \mathbb{N}_B & \mathbb{K}_B^* & & & & \mathbb{C}_{B,3}^T \\ & \ddots & \ddots & \ddots & & & \vdots \\ \mathbb{N}_B & \mathbb{N}_B & \mathbb{N}_B & \mathbb{N}_B & \mathbb{K}_B^* & & \mathbb{C}_{B,m}^T \\ \hline & & & & & \mathbb{K}_A^* & \mathbb{C}_{A,m}^T \\ \hline \mathbb{L}_B & \mathbb{L}_B & \mathbb{L}_B & \cdots & \mathbb{L}_B & \mathbb{L}_A & \mathbf{0} \end{array}\right] \left[\begin{array}{c} \Delta\mathbb{U}_1^B \\ \Delta\mathbb{U}_2^B \\ \Delta\mathbb{U}_3^B \\ \vdots \\ \Delta\mathbb{U}_m^B \\ \hline \Delta\mathbb{U}_m^A \\ \hline \boldsymbol{\lambda}_m \end{array}\right] = \left[\begin{array}{c} \mathbb{F}_{\xi_{B,f}}^B - \mathbb{N}_B\mathbb{U}_0^B - \mathbb{E}_{B,1}^T\boldsymbol{\lambda}_0 \\ \mathbb{F}_{1+\xi_{B,f}}^B - \mathbb{N}_B\mathbb{U}_0^B - \mathbb{E}_{B,2}^T\boldsymbol{\lambda}_0 \\ \mathbb{F}_{2+\xi_{B,f}}^B - \mathbb{N}_B\mathbb{U}_0^B - \mathbb{E}_{B,3}^T\boldsymbol{\lambda}_0 \\ \vdots \\ \mathbb{F}_{m-1+\xi_{B,f}}^B - \mathbb{N}_B\mathbb{U}_0^B - \mathbb{E}_{B,m}^T\boldsymbol{\lambda}_0 \\ \hline \mathbb{F}_{\xi_{A,f}}^A - \mathbb{N}_A\mathbb{U}_0^A - \mathbb{E}_{A,m}^T\boldsymbol{\lambda}_0 \\ \hline \mathbf{0} \end{array}\right] \tag{21}$$

in which the forms of the global matrices $\mathbb{L}_A$, $\mathbb{L}_B$, $\mathbb{C}_{A,m}$, $\mathbb{E}_{A,m}$, $\mathbb{C}_{B,j}$ and $\mathbb{E}_{B,j}$ are the same as in section 2.1.

4 CONVERGENCE ANALYSIS

The BGC-macro method preserves exactly the pseudo-interface energy, proving the stability of the coupling algorithm. It remains to analyse the accuracy order of the coupling algorithm. Along the lines of *Bonelli et al.* [2008] and *Bursi et al.* [2010], an alternative convergence analysis is investigated by the spectral analysis of the amplification matrix $\mathbf{A}$, linking a state vector at the time $t_n + h_A = t_{n+1}$ to the previous state vector t_n. In addition to the stability, the amplification matrix enables to derive the accuracy order as well as the algorithmic damping ratio and the period elongation error related to the coupling algorithm. The convergence of the BGC-macro method is carried out for a single degree of freedom system, split into two subdomains Ω_A and Ω_B. The BGC-macro method is recast into the following recursive form:

$$\mathbf{X}_{n+1} = \mathbf{A}\mathbf{X}_n + \mathbf{L}_n \tag{22}$$

where $\mathbf{X}$ is an appropriate state vector depending on the formulation of the time integration algorithm, $\mathbf{A}$ is the amplification matrix and $\mathbf{L}$ is the load vector depending on the external forces. The BGC-macro method adopts the following state vector at time t_n:

$$\mathbf{X}_n = \left[\mathbf{X}_n^A \;\; \mathbf{X}_n^B\right]^T \tag{23}$$

where $\mathbf{X}_n^A$, $\mathbf{X}_n^B$ are the state vectors related the subdomains Ω_A and Ω_B, respectively:

$$\mathbf{X}_n^A = \left[\mathbf{u}_n^A \;\; \mathbf{v}_n^A \;\; \mathbf{a}_n^A \;\; \boldsymbol{\lambda}_n\right]^T \;\; \textit{and} \;\; \mathbf{X}_n^B = \left[\mathbf{u}_n^B \;\; \mathbf{v}_n^B \;\; \mathbf{a}_n^B\right]^T \tag{24}$$

For the split oscillator, the state vector $\mathbf{X}_n$ gathers 7 components. In the following, we do not consider the load vector $\mathbf{L}_n$ because we assume that the leading error term of its

approximation is greater than the order of accuracy of the method. It has been in *Brun et al.* [2014] that Eq. 22 can be developed as:

$$\begin{bmatrix} \mathbf{X}_{n+1}^A \\ \mathbf{X}_{n+1}^B \end{bmatrix} = \begin{bmatrix} \mathbf{A}^{AA} & \mathbf{A}^{AB} \\ \mathbf{A}^{BA} & \mathbf{A}^{BB} \end{bmatrix} \begin{bmatrix} \mathbf{X}_n^A \\ \mathbf{X}_n^B \end{bmatrix} \tag{25}$$

where the matrices $\mathbf{A}^{AA}$, $\mathbf{A}^{AB}$, $\mathbf{A}^{BA}$ and $\mathbf{A}^{BB}$ depend on the time step ratio, noted in the following as ss. The split oscillator is defined by the following mass and stiffness decomposition: $m = m^A + m^B$, $k = k^A + k^B$ with the ratio $b_1 = \frac{m^A}{m^B} = \frac{k^B}{k^A}$. The amplification matrix is too complex for a closed-form expression as soon as different time steps are used. As a consequence, the amplification matrix is numerically computed for different values of time integrator parameters (ρ_∞ for CH-α schemes), mass and stiffness ratio b_1 and time step ratio ss. To address the accuracy, an alternative form of the amplification matrix is considered (*Bonelli et al.* [2008], *Bursi et al.* [2010]):

$$\bar{\mathbf{A}} = \bar{\mathbf{M}}\mathbf{A}\bar{\mathbf{M}}^{-\mathbf{1}} \tag{26}$$

with the state vector modified as:

$$\bar{\mathbf{X}}_n = \left[\mathbf{u}_n^A \;\; \mathbf{v}_n^A \;\; h_A\mathbf{a}_n^A \;\; h_A\boldsymbol{\lambda}_n \;\; \mathbf{u}_n^B \;\; \mathbf{v}_n^B \;\; h_A\mathbf{a}_n^B\right]^T \tag{27}$$

and the matrix $\bar{\mathbf{M}}$ given by:

$$\bar{\mathbf{M}} = \textit{block diagonal}\left[\mathbf{I} \;\; \mathbf{I} \;\; h_A\mathbf{I} \;\; h_A\mathbf{I} \;\; \mathbf{I} \;\; \mathbf{I} \;\; h_A\mathbf{I}\right]^T \tag{28}$$

where $\mathbf{I}$ is the identity matrix (equal to 1 in the case of the split oscillator).

Let us begin with the study of the stability of the BGC-macro method. The absolute stability of the BGC-macro method is investigated by computing the seven eigenvalues of the amplification matrix $\mathbf{A}$ in the case of the split-oscillator. The eigenvectors are linearly independent for each repeated eigenvalues λ_i. Hence, the condition $|\lambda_i| \leq 1$, for $i = 1...7$, is sufficient to demonstrate the A-stability of the method. The coupling between two CH-α schemes (spectral radii $\rho_\infty = 0.8$ and $\rho_\infty = 0.5$ for the two subdomains at the coarse and fine time scales, respectively) is investigated as a function of the time step ratio ss by plotting the absolute values of the eigenvalues as a function of the reduced angular frequency $\Omega_B = \omega_B h_B$. Among the seven eigenvalues, only one pair is complex conjugate, giving the principal eigenvalues, whereas the five remaining are the spurious ones. In all investigated cases (any ρ_∞ and b_1), the BGC-macro method is found to be unconditionally stable when coupling unconditionally stable time integrators.

The local truncation error $\bar{\tau}_n$ is defined as

$$\bar{\tau}_n = \bar{\mathbf{A}}\bar{\mathbf{X}}(t_n) - \bar{\mathbf{X}}(t_{n+1}) \tag{29}$$

where $\bar{\mathbf{X}}(t_n)$ and $\bar{\mathbf{X}}(t_{n+1})$ correspond to the exact solutions of the state vector at the time t_n and t_{n+1}. From numerical calculations, the order of the truncation error can be assessed

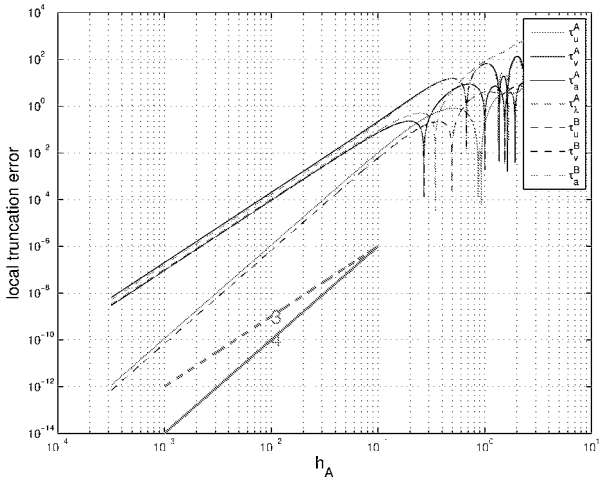


Figure 1: Local truncation error $\bar{\tau}$ of the split oscillator ($b_1 = 1$): CH-α ($\rho_\infty = 0.8$) coupled with CH-α ($\rho_\infty = 0.5$) with time step ratios $ss = 2$.

by computing the slope of $\ln(\bar{\tau}_n)$ as a function of the macro time step h_A. Indeed, the power k of the leading term of the local truncation error $\bar{\tau}_n = \alpha h_A^k + O(h_A^{k+1})$ (α being a constant) can be computed as: $k = \frac{\ln(\bar{\tau}_n(h_{2,A})) - \ln(\bar{\tau}_n(h_{1,A}))}{\ln(h_{2,A}) - \ln(h_{1,A})}$, where $\bar{\tau}_n(h_{2,A})$ and $\bar{\tau}_n(h_{1,A})$ are the numerical results of the local truncation error for two different small values of the macro time step h_A, with $h_{1,A} < h_{2,A}$. One obtains

$$\bar{\tau}_n = O(h_A^3) \tag{30}$$

for any values of the time step ratio ss. It is worth noting that the same result has been obtained in *Bonelli et al.* [2008] (Newmark schemes) and *Bursi et al.* [2010] (α-schemes) for the PM method, but only in the case of the same time step in both subdomains. Indeed, as soon as different time steps are adopted, the PM method only exhibits $\bar{\tau}_n = O(h_A^2)$.

For illustration, the local truncation error is plotted in Figure 1 for a time step ratio $ss = 2$. It can be seen that the power k of the leading term of the local truncation error depends on the quantity under consideration (displacement, velocity, acceleration and Lagrange multiplier). In details, we have: $\bar{\tau}_u^A = O(h_A^4)$, $\bar{\tau}_v^A = O(h_A^3)$, $\bar{\tau}_a^A = O(h_A^3)$, $\bar{\tau}_\lambda^A = O(h_A^3)$, $\bar{\tau}_u^B = O(h_A^4)$, $\bar{\tau}_v^B = O(h_A^3)$, $\bar{\tau}_a^B = O(h_A^3)$.

According to the Lax equivalence theorem (following the arguments of *Bonelli et al.* [2008] and *Bursi et al.* [2010]), a method will be convergent of the order k, if it is stable and consistent of the order k, *i.e.* $\bar{\tau}_n = O(h_A^{k+1})$. As a result, the BGC-macro method is convergent of the order 2 for any time step ratio.

It is important to underline that the second order accuracy of the α-schemes is preserved through coupling with the BGC-macro method, which is not the case for the algorithms proposed in the literature (*Bursi et al.* [2010]).

5 CONCLUSIONS

The paper presents a general methodology for building Heterogeneous Asynchronous Time Integrators based on energy considerations. It has been sought to cancel the interface pseudo-energy, originally introduced by Hughes in the energy method, for demonstrating the stability of hybrid time integration coupling methods related with homogeneous and heterogeneous time steps. The BGC-macro method solves the interface problem, whose Lagrange multipliers at the interface are the unknowns, at the macro time scale by deriving from the zero value of the interface pseudo-energy a kinematic constraint at the interface expressed in terms of velocity increments. A linear assumption on the variation of the Lagrange multipliers over the macro time step is adopted. From energy arguments, the BGC-macro method is stable. It has been confirmed by the convergence analysis through the spectral analysis of the amplification matrix. In addition to the stability, the BGC-macro method is found to be second-order accurate when dealing with Newmark and α-schemes.

References

Bonelli, A., O. Bursi, L. He, G. Magonette, and P. Pegon, Convergence analysis of a parallel interfield method for heterogeneous simulations with dynamic substructuring, *International Journal for Numerical Methods in Engineering*, *75*, 800–825, 2008.

Bonnet, P., M. Williams, and A. Blakeborough, Evaluation of numerical time-integration schemes for real-time hybrid testing, *Earthquake Engineering and Structural Dynamics 2008*, *37*, 1467–1490, 2008.

Brun, M., A. Batti, A. Limam, and A. Combescure, Implicit/explicit multi-time step co-computations for predicting reinforced concrete structure response under earthquake loading, *Soil dynamics and Earthquake Engineering*, *33*, 19–37, 2012.

Brun, M., A. Batti, A. Combescure, and A. Gravouil, External coupling software based on macro- and micro-time scales for explicit/implicit multi-time-step co-computations in structural dynamics, *Finite Elements in Analysis & Design*, *52*, 41–59, 2014.

Brun, M., A. Gravouil, A. Combescure, and A. Limam, Two feti-based heterogeneous time step coupling methods for newmark and alpha-schemes derived from the energy method, *Computer Methods in Applied Mechanics and Engineering*, *283*, 130–176, 2015.

Bursi, O., L. He, A. Bonelli, and P. Pegon, Novel generalized-α methods for interfield parallel integration of heterogeneous structural dynamic systems, *Journal of Computational and Applied Mathematics*, *234*, 2250–2258, 2010.

Chung, J., and G. Hulbert, A time integration algorithm for structural dynamics with improved numerical dissipation: the generalized-α method, *Journal of Applied Mechanics*, *60*, 371–375, 1993.

Combescure, A., and A. Gravouil, A numerical scheme to couple subdomains with different time-steps for predominantly linear transient analysis, *Computer methods in applied mechanics and engineering*, *191*, 1129–1157, 2002.

Combescure, D., and P. Pegon, α-operator splitting time integration technique for pseudodynamic testing - error propagation analysis, *Soil Dynamics and Earthquake Engineering*, *16*, 427–443, 1997.

Gravouil, A., and A. Combescure, A multi-time-step explicit-implicit method for non-linear structural dynamics, *International Journal for Numerical Methods in Engineering*, *50*, 199–225, 2001.

Gravouil, A., A. Combescure, and M. Brun, Heterogeneous asynchronous time integrators for computational structural dynamics, *International Journal of Numerical Methods in Engineering, special edition in honor of Ted Belytschko, in press*, 2015.

Hilber, H., T. Hughes, and R. Taylor, Improved numerical dissipation for time integration algorithms in structural dynamics, *Earthquake Engineering and Structural Dynamics*, *5*, 283–292, 1977.

Hughes, T., *The Finite Element Method: Linear Static and Dynamic Finite Element Analysis*, Prentice-Hall, Englewood Cliffs, NJ, 1987.

Masuri, S., A. Hoitink, X. Zhou, and K. Tamma, Algorithms by design: A new normalized time-weighted residual methodology and design of a family of energy-momentum conversing algorithms for non-linear structural dynamics, *International Journal for Numerical Methods in Engineering*, *79*, 1094–1146, 2009.

Pinto, A., P. Pegon, G. Magonette, and G. Tsionis, Pseudo-dynamic testing of bridges using non-linear substructuring, *Earthquake Engineering and Structural Dynamics*, *33*, 1125–1146, 2004.

Prakash, A., and K. Hjelmstad, A feti-based multi-time-step coupling method for newmark schemes in structural dynamics, *International Journal for Numerical Methods in Engineering 2004*, *61*, 2183–2204, 2004.

Wood, W., M. Bossak, and O. Zienkiewicz, An alpha modification of newmark's method, *International Journal for Numerical Methods in Engineering*, *15*, 1562–1566, 1981.

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis (Eds)

HYBRID SIMULATION OF HEAT TRANSFER PROBLEMS IN STRUCTURAL APPLICATIONS – COUPLED PROBLEMS 2015

GIUSEPPE ABBIATI*, ORESTE S. BURSI†, BOZIDAR STOJADINOVIC*, NICOLA TONDINI† AND CATHERINE WHYTE*

* Department of Civil, Environmental and Geomatic Engineering (D-BAUG), IBK, ETH Zurich
Stefano-Franscini-Platz 5, 8093 Zürich, Switzerland
e-mail: abbiati@ibk.baug.ethz.ch, whyte@ibk.baug.ethz.ch, stojadinovic@ibk.baug.ethz.ch

† Department of Civil, Environmental and Mechanical Engineering, University of Trento
Via Mesiano 77, 38123, Trento, Italy
email: oreste.bursi@unitn.it, nicola.tondini@unitn.it

Key words: Hybrid simulation, truss structures, thermal analysis, electric furnaces.

Abstract. The present paper presents all research activities focused on the development of a pure thermal hybrid simulator (THS). In detail, the need for a rigorous coupling is investigated, i.e. temperatures are sent from the numerical substructure (NS) to the physical substructure (PS) and interface heat fluxes are sent back from the PS to the NS. In this respect, a realistic benchmark case study is presented. It consists of a 2D truss bridge where a single truss element is "physically" substructured. In the current preliminary phase, an additional multiphysics FE code, COMSOL, is utilized to simulate the thermal response of this "experimental" substructure inside the electric furnace. In detail, two variants of the same case study are presented and characterized by two significantly different average thermal diffusivities. This approach provides a realistic insight into the capabilities of THS. Moreover, the present study paves the way for the implementation of a fully coupled thermomechanical hybrid simulation (TMHS), which account for indirect actions owing to restrained thermal deformations on the hot NS.

1 INTRODUCTION

Large-scale structural fire tests are rare because they are costly and require specialized facilities. As a result, most of the research regarding the behavior of structures in fire has been carried out on single structural components. These component tests do not provide insight into the thermomechanical interaction problem of the selected structural component with the remainder of the structure. In fact, statically indeterminate structural assemblies subjected to thermal loads experience indirect actions due to restrained thermal deformations. The hybrid simulation technique, extensively investigated in the mechanical domain, can be extended to the thermal field through temperature controlled furnaces and, thus, account for such a thermomechanical interaction. The first hybrid fire test (HFT) was performed by Mostafaei [1,2]. The PS was a first story column of a six-story reinforced concrete building, exposed to both axial and thermal loads in a vertical furnace. In order to simulate the fire development, a standard ASTM E119 fire curve was applied to both the column specimen and the numerical model of the remaining building, i.e. the NS implemented in the SAFIR FE Code [3]. However, the process was not

automated. A fully automated thermomechanical hybrid simulator (TMHS) developed by Whyte et al. [4] extended the scope of the well-known OpenFresco hybrid simulation framework [5] to include the temperature Degree-of-Freedom (DoF) and modeled the NS in the OpenSees FE code [6]. They validated their implementation on a Single-DoF two-element hybrid model. However, since OpenSees is a pure mechanical FE solver, the heat transfer dynamics was not accounted for on the numerical side. According to the state of the art, there remains a lack of a hybrid simulation (HS) environment where a substructure thermal coupling encompasses the whole domain, i.e. numerical and physical substructures. Along this line, the present paper describes research activities focused on the development of a pure thermal hybrid simulator (THS), which paves the way for simulating hot NSs for the purpose of TMHS.

2 GOVERNING EQUATIONS OF THERMOELASTICITY

First, the governing equations of thermoelasticity that define the coupled thermomechanical problem for a linear elastic isotropic continuum are briefly introduced [7]. The theory of thermoelasticity consists of i) the equilibrium equation;

$$\sigma_{ij,j} + X_i = \rho \ddot{u}_i \tag{1}$$

where σ_{ij} is the stress, X_i is the body force, ρ is the density and $\ddot{u}_i$ is the displacement.
ii) the heat transfer equation and iii) Fourier's heat conduction law:

$$\rho c_p \dot{\theta} - k\theta_{,ii} + \theta_0 \left(1 + \frac{\theta - \theta_0}{\theta_0}\right) \beta \dot{\varepsilon}_{ii} = Q \tag{2}$$

$$q_i = -k\theta_{,j} \tag{3}$$

where, k is the thermal conductivity, c_p is the specific heat at constant pressure, β represents the thermal expansion material constant, $\dot{\varepsilon}_{ij}$ the strain, Q is the heat source per unit volume, θ represents the temperature, θ_0 the reference temperature and q_i is the heat flux. Then, for a linear elastic behavior the constitutive equations read:

$$\varepsilon_{ij} = \frac{1}{2G}\left[\sigma_{ij} - \frac{\nu}{1+\nu}\sigma_{kk}\delta_{ij}\right] + \alpha \cdot (\theta - \theta_0)\,\delta_{ij} \tag{4}$$

where σ_{ij} and ε_{ij} are the generic stress and the strain tensor components, respectively; G is the tangent elastic modulus, whilst υ is the Poisson ratio and $\alpha = \beta/E$ is the thermal dilatation coefficient. The thermoelastic problem - Eqn. (1-4) - is solved by prescribing boundary conditions on the body surfaces and initial conditions. Several discretization techniques allow for approximating solutions of the set of coupled equations (1-4). For the sake of simplicity, all the following equations refer to a linear thermoelastic system. The resulting semi-discretized set of ordinary differential equations reads,

$$\begin{bmatrix} \mathbf{M}_{uu} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \ddot{\mathbf{u}} \\ \ddot{\boldsymbol{\theta}} \end{bmatrix} + \begin{bmatrix} \mathbf{C}_{uu} & \mathbf{0} \\ \mathbf{C}_{\theta u} & \mathbf{C}_{\theta\theta} \end{bmatrix} \begin{bmatrix} \dot{\mathbf{u}} \\ \dot{\boldsymbol{\theta}} \end{bmatrix} + \begin{bmatrix} \mathbf{K}_{uu} & \mathbf{K}_{u\theta} \\ \mathbf{K}_{\theta u} & \mathbf{K}_{\theta\theta} \end{bmatrix} \begin{bmatrix} \mathbf{u} \\ \boldsymbol{\theta} \end{bmatrix} = \begin{bmatrix} \mathbf{F}_u \\ \mathbf{F}_\theta \end{bmatrix} \tag{5}$$

In detail, matrix partitioning refers to displacement and temperature DoFs, i.e. $\mathbf{u}$ and $\boldsymbol{\theta}$, respectively. Accordingly, $\mathbf{M}_{uu}, \mathbf{K}_{uu}$ and $\mathbf{C}_{uu}$ are mass, stiffness and damping matrices; $\mathbf{K}_{\theta\theta}$ and $\mathbf{C}_{\theta\theta}$ are heat conductive and capacity matrices. $\mathbf{F}_u$ and $\mathbf{F}_\theta$ vectors represent mechanical and thermal loads, respectively, where positive thermal fluxes $\mathbf{F}_\theta$ supply energy to the system, whilst negative subtract energy from the system.
Since the out-of-diagonal block submatrix $\mathbf{K}_{u\theta}$ provides internal forces owing to restrained thermal deformations, it is crucial to account for thermomechanical structural interactions. $\mathbf{K}_{\theta u}$ accounts for thermal load changes due to mechanical behavior. This is the case when cover spalling in concrete structural elements occurs and reinforcement bars are directly exposed to the thermal action of fire or when the specimen undergoes large deformations in a compartment with non-uniform temperature, e.g. the floor deflects down towards the fire or a beam twists changing the exposure from the flange to the web. Finally, $\mathbf{C}_{\theta u}$ represents the heat generation due to strain rate, which is important for micro-scale applications, but negligible for large-scale problems. In this preliminary phase both contributions dictated by terms $\mathbf{C}_{\theta u}$ and $\mathbf{K}_{\theta u}$ were neglected, which means that the mechanical behaviour does not affect the temperature distribution in the solid body. At this stage, the following classification can be helpful:

- **Physics full coupling**: the thermal behavior affects the mechanical behavior and vice versa.
- **Physics partial coupling**: the thermal/mechanical behavior affects mechanical/thermal behavior but not vice versa.

Now, by applying the Hybrid Simulation (HS) technique, each matrix can be partitioned in pure Numerical-, pure Physical- and Boundary-DoFs, respectively, [8]. For brevity, the following simplified notation holds: N-DoFs, P-DoFs and B-DoFs, respectively. The complete derivation of the governing discretized equations entails that,

$$\mathbf{u} = \begin{bmatrix} \mathbf{u}^{N^T} & \mathbf{u}^{B^T} & \mathbf{u}^{P^T} \end{bmatrix}^T, \quad \boldsymbol{\theta} = \begin{bmatrix} \boldsymbol{\theta}^{N^T} & \boldsymbol{\theta}^{B^T} & \boldsymbol{\theta}^{P^T} \end{bmatrix}^T \tag{6}$$

Accordingly, a generic load vector $\mathbf{F}$ reads:

$$\mathbf{F}_u = \begin{bmatrix} \mathbf{F}_u^{N^T} & \mathbf{F}_u^{B^T} & \mathbf{F}_u^{P^T} \end{bmatrix}^T, \quad \mathbf{F}_\theta = \begin{bmatrix} \mathbf{F}_\theta^{N^T} & \mathbf{F}_\theta^{B^T} & \mathbf{F}_\theta^{P^T} \end{bmatrix}^T \tag{7}$$

All matrices must be expanded to all DoFs considered in the emulated system. In the HS practice B-DoFs coincide with P-DoFs [9]. The expanded stiffness and conductivity matrices yield,

$$\mathbf{K}_{uu}^N = \begin{bmatrix} \mathbf{K}_{uu}^N & \mathbf{K}_{uu}^N & \mathbf{0} \\ \mathbf{K}_{uu}^N & \mathbf{K}_{uu}^N & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}, \mathbf{K}_{uu}^P = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{K}_{uu}^P & \mathbf{K}_{uu}^P \\ \mathbf{0} & \mathbf{K}_{uu}^P & \mathbf{K}_{uu}^P \end{bmatrix}, \quad \mathbf{K}_{\theta\theta}^N = \begin{bmatrix} \mathbf{K}_{\theta\theta}^N & \mathbf{K}_{\theta\theta}^N & \mathbf{0} \\ \mathbf{K}_{\theta\theta}^N & \mathbf{K}_{\theta\theta}^N & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}, \mathbf{K}_{\theta\theta}^P = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{K}_{\theta\theta}^P & \mathbf{K}_{\theta\theta}^P \\ \mathbf{0} & \mathbf{K}_{\theta\theta}^P & \mathbf{K}_{\theta\theta}^P \end{bmatrix}, \tag{8}$$

The same criterion applies to remainder of (5). Accordingly, (5) can be split into NS and PS contributions.

$$\begin{bmatrix} \left(\mathbf{M}_{uu}^{N}+\mathbf{M}_{uu}^{P}\right) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}\begin{bmatrix} \ddot{\mathbf{u}} \\ \ddot{\boldsymbol{\theta}} \end{bmatrix}+\begin{bmatrix} \left(\mathbf{C}_{uu}^{N}+\mathbf{C}_{uu}^{P}\right) & \mathbf{0} \\ \mathbf{0} & \left(\mathbf{C}_{\theta\theta}^{N}+\mathbf{C}_{\theta\theta}^{P}\right) \end{bmatrix}\begin{bmatrix} \dot{\mathbf{u}} \\ \dot{\boldsymbol{\theta}} \end{bmatrix}+\dots$$
$$\dots\begin{bmatrix} \left(\mathbf{K}_{uu}^{N}+\mathbf{K}_{uu}^{P}\right) & \left(\mathbf{K}_{u\theta}^{N}+\mathbf{K}_{u\theta}^{P}\right) \\ \mathbf{0} & \left(\mathbf{K}_{\theta\theta}^{N}+\mathbf{K}_{\theta\theta}^{P}\right) \end{bmatrix}\begin{bmatrix} \mathbf{u} \\ \boldsymbol{\theta} \end{bmatrix}=\begin{bmatrix} \left(\mathbf{F}_{u}^{N}+\mathbf{F}_{u}^{P}\right) \\ \left(\mathbf{F}_{\theta}^{N}+\mathbf{F}_{\theta}^{P}\right) \end{bmatrix} \tag{9}$$

After simple manipulations of (9), the thermomechanical response of the PS enters the right hand side (RHS) of the balance equation as interface load **R**, i.e. force and heat flux, applied to the NS. With regard to the real-time (RT) case, where a rate dependent response of the PS is expected, (9) becomes,

$$\begin{bmatrix} \mathbf{M}_{uu}^{N} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}\begin{bmatrix} \ddot{\mathbf{u}} \\ \ddot{\boldsymbol{\theta}} \end{bmatrix}+\begin{bmatrix} \mathbf{C}_{uu}^{N} & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_{\theta\theta}^{N} \end{bmatrix}\begin{bmatrix} \dot{\mathbf{u}} \\ \dot{\boldsymbol{\theta}} \end{bmatrix}+\begin{bmatrix} \mathbf{K}_{uu}^{N} & \mathbf{K}_{u\theta}^{N} \\ \mathbf{0} & \mathbf{K}_{\theta\theta}^{N} \end{bmatrix}\begin{bmatrix} \mathbf{u} \\ \boldsymbol{\theta} \end{bmatrix}=\begin{bmatrix} \mathbf{F}_{u}^{N}-\mathbf{R}_{u}^{P} \\ \mathbf{F}_{\theta}^{N}-\mathbf{R}_{\theta}^{P} \end{bmatrix} \tag{10}$$

where,

$$\begin{bmatrix} \mathbf{R}_{u}^{P} \\ \mathbf{R}_{\theta}^{P} \end{bmatrix}=\begin{bmatrix} \mathbf{M}_{uu}^{P} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}\begin{bmatrix} \ddot{\mathbf{u}} \\ \ddot{\boldsymbol{\theta}} \end{bmatrix}+\begin{bmatrix} \mathbf{C}_{uu}^{P} & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_{\theta\theta}^{P} \end{bmatrix}\begin{bmatrix} \dot{\mathbf{u}} \\ \dot{\boldsymbol{\theta}} \end{bmatrix}+\begin{bmatrix} \mathbf{K}_{uu}^{P} & \mathbf{K}_{u\theta}^{P} \\ \mathbf{0} & \mathbf{K}_{\theta\theta}^{P} \end{bmatrix}\begin{bmatrix} \mathbf{u} \\ \boldsymbol{\theta} \end{bmatrix}-\begin{bmatrix} \mathbf{F}_{u}^{P} \\ \mathbf{F}_{\theta}^{P} \end{bmatrix} \tag{11}$$

However, the performance of the transfer system, i.e. furnaces and actuators, can limit the loading rate. Therefore, extended time scales are adopted in the laboratory where the test is slowed down with respect to RT testing. As a result, rate dependent components of the PS response must be accounted for numerically. This is the pseudodynamic (PsD) case and (9) turns into,

$$\begin{bmatrix} \left(\mathbf{M}_{uu}^{N}+\mathbf{M}_{uu}^{P}\right) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}\begin{bmatrix} \ddot{\mathbf{u}} \\ \ddot{\boldsymbol{\theta}} \end{bmatrix}+\begin{bmatrix} \left(\mathbf{C}_{uu}^{N}+\mathbf{C}_{uu}^{P}\right) & \mathbf{0} \\ \mathbf{0} & \left(\mathbf{C}_{\theta\theta}^{N}+\mathbf{C}_{\theta\theta}^{P}\right) \end{bmatrix}\begin{bmatrix} \dot{\mathbf{u}} \\ \dot{\boldsymbol{\theta}} \end{bmatrix}+\begin{bmatrix} \mathbf{K}_{uu}^{N} & \mathbf{K}_{u\theta}^{N} \\ \mathbf{0} & \mathbf{K}_{\theta\theta}^{N} \end{bmatrix}\begin{bmatrix} \mathbf{u} \\ \boldsymbol{\theta} \end{bmatrix}=\begin{bmatrix} \mathbf{F}_{u}^{N}-\mathbf{R}_{u}^{P} \\ \mathbf{F}_{\theta}^{N}-\mathbf{R}_{\theta}^{P} \end{bmatrix} \tag{12}$$

where,

$$\begin{bmatrix} \mathbf{R}_{u}^{P} \\ \mathbf{R}_{\theta}^{P} \end{bmatrix}=\begin{bmatrix} \mathbf{K}_{uu}^{P} & \mathbf{K}_{u\theta}^{P} \\ \mathbf{0} & \mathbf{K}_{\theta\theta}^{P} \end{bmatrix}\begin{bmatrix} \mathbf{u} \\ \boldsymbol{\theta} \end{bmatrix}-\begin{bmatrix} \mathbf{F}_{u}^{P} \\ \mathbf{F}_{\theta}^{P} \end{bmatrix} \tag{13}$$

The first block row of (10) and (12), without the coupling term $\mathbf{K}_{u\theta}$, has been deeply investigated for many years and extensively used in the seismic domain. Conversely, the second row-block still deserves particular attention. In fact, it determines the thermal coupling between NS and PS. In principle, compatibility of primal quantities -displacements, temperatures- and balance of dual quantities -forces, heat fluxes- must hold across the interface between the NS and the PS. This means that an accurate temperature control is needed for applying a prescribed temperature field at the interface of the PS and the corresponding heat fluxes should feedback to the NS. Nevertheless, some constraint can be potentially relaxed allowing for easier implementations. For the sake of clarity, a couple of definitions is given about strength of coupling between PSs and NSs:

- **Substructures full coupling**: displacements/temperatures are sent from the NS to PS and interface forces/heat fluxes are sent back from the PS to the NS.

- **Substructures partial coupling**: the same as full coupling forces/heat fluxes are not sent back to the NS from the PS.

In this respect, the paper focuses on the second row block of (9) from the hybrid simulation perspective, i.e. substructure thermal coupling. Therefore, a rigorous coupling among "physical" – here numerically emulated - and numerical subdomains is investigated through numerical simulations.

3 THE ILLINOIS BRIDGE REFERENCE CASE STUDY

From the perspective of the development of a full THS, a proof-of-concept case study was selected. The 5.6 m long, truss bridge mock-up model, which was tested at the Smart Structures Technology Laboratory (SSTL) of the University of Illinois at Urbana-Champaign [10], was taken as reference structural system, as illustrated in Figures 1 and 2. A typical fire scenario for such a structure can be a burning truck on the carriageway. The length of each bay of the truss is 0.4 m on each side. The thermal physics was considered only.

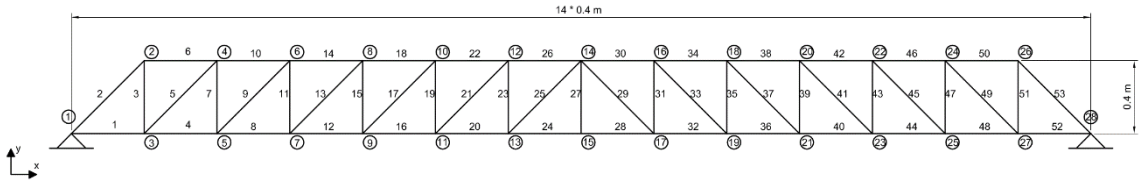


Fig. 1 - Scheme of the Matlab 2D FE model based on thermal truss elements with numbering enabled.

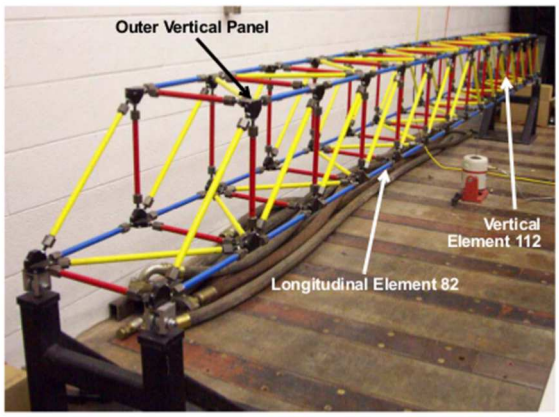


Fig. 2 - Mock-up model of the reference truss bridge case study.

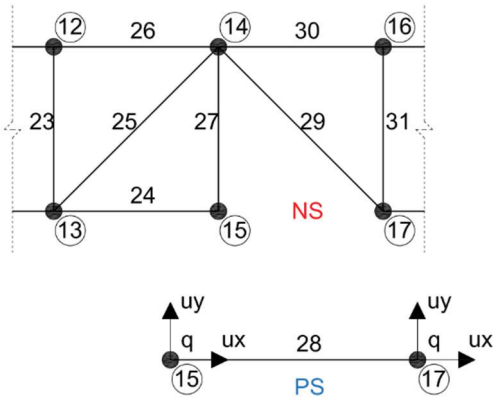
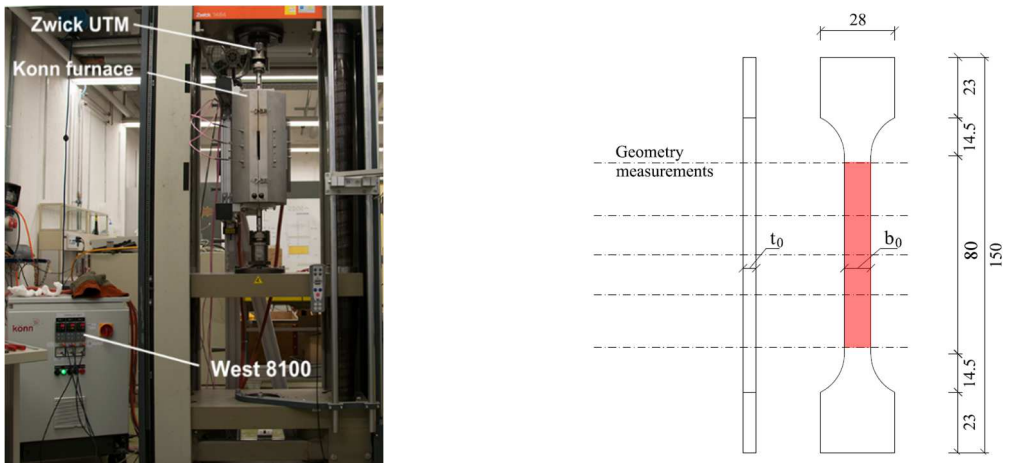


Fig. 3 - Substructuring scheme: Element #28 is substructured in the laboratory.

As can be appreciated from Figure 3, a single 2D truss element is substructured in the laboratory. A Könn STE-12 HR/350' electrical furnace (Könn Furnace in Figure 4a), which encapsulated the 1:5 scale specimen of the substructured truss of Figure 4b, controls temperatures using three West 8100+ Single Loop Controllers (West 8100). This provides the ability to control temperature in three zones. Although uniform temperature fields characterize the specimen in this preliminary study, thermal gradients can be potentially applied to the substructured element. The furnace is capable of heating at a rate of 60 degrees/minute up to 600 degrees and 40 degrees/minute up to 1100 degrees. A Zwick 1484 Universal Testing Machine (Zwick UTM) holds the specimen.



a b

Fig. 4 - a) Zwick UTM and Könn Furnace test setup; b) schematic of the dog-bone specimen including the geometric values in mm.

Temperature feedbacks are measured from both ends of the specimen. The Zwick UTM and Könn furnace had existing controllers with a Zwick/Roell testXpert II Software user interface [11]. An INDEL RT computer runs a Simulink model that handles both the controller as well as the THS software. Figure 4b shows the geometry of the test specimen. Since grips hold specimen's ends, the highlighted portion, is considered as the scaled substructured truss for the purpose of THS. The nominal width b_0 of the specimen measures 10 mm, whilst the nominal thickness t_0 measures 5 mm. The total length is 80 mm. Therefore, the specimen is consistent with the substructured truss with a geometric scaling factor δ equal to 5. A pair of thermocouples at each end of the specimen allows for estimating temperature gradients, and therefore feedback heat fluxes.

3.1 FE model of the reference case study

A linear 2D FE model of the bridge consisting of 53 thermal truss elements and 28 nodes was implemented in Matlab. According to Figure 1, structural nodes are numbered starting from left to right, with all the odd numbers at the lower chord, and all the even numbers at the upper chord, except node 28 which is at the right support. Each structural node number has a circle around it. All elements are characterized by a circular cross section with an inner diameter of 1.09 cm and an outer diameter of 1.71 cm except Elements #24 and #28, which have 2.50 x 5.00 cm rectangular solid cross sections. Uniform steel material properties were considered for the all bridge truss elements. Table 1 summarize steel material parameters characterizing the bridge that for simplicity were kept constant with temperature:

Table 1. Steel material parameters characterizing bridge trusses.

Description	Symbol	Value	Unit
Density	ρ	7800	[kg/m3]
Thermal conductivity	k	43	[W/m/K]
Thermal capacity	c_p	445	[J/kg/K]

In order to keep the problem linear, thermal radiation was neglected. Conversely, heat transfer by convection was considered at each node. An equivalent exchanging surface S of 0.80 m^2 was considered for all related elements. In order to investigate the influence of each heat transfer mechanism on the analyzed problem, two case studies were defined: Case Study #1 considers a convective coefficient h equal to 50 $W/(m^2K)$ and steel material properties; Case Study #2 analyses the problem with reduced density of 78 kg/m^3 and convective coefficient h equal to 1 $W/(m^2K)$. A significant reduction of density entails a much higher thermal diffusivity (from 11 mm^2/s typical of steel to 1200 mm^2/s typical of pyrolytic graphite) and a smaller Biot number for Case Study #2 than Case Study #1. This means that relative to Case Study #1, in Case Study #2 the resistance to heat transfer offered by conduction is less with respect to convection. The ANSYS FE code was taken as reference for the validation of the developed Matlab FE library.

3.2 Refined numerical modelling of the THS environment

HS was extensively applied for determining the seismic response of a prototype structure. In detail, the substructure, whose behavior is known, is modeled numerically - NS -, while the portion of the structure whose behavior is highly nonlinear or not well-understood is tested in the laboratory - PS -. The resulting hybrid model consists of numerical and physical subdomains that interact at each time step of the solution of the equation of motion for an applied dynamic excitation. The control software and the specimen actuation system enforce consistent boundary conditions at the interface between the substructures. The same philosophy applies to the thermal case. Here in the current preliminary phase, an axisymmetric COMSOL FE model simulates the thermal response of the "experimental" substructure inside the Könn Furnace. This approach allows for reliable estimates of actual heat fluxes coming from the truss specimen, which account for furnace leakage; thus, providing realistic insight into the capabilities of THS. The axisymmetric approach allows for reducing the total number of DoFs. Figure 5a reports the reference slice of the axisymmetric COMSOL FE model, whilst Figure 5b, shows the FE model in the revolved 3D domain.

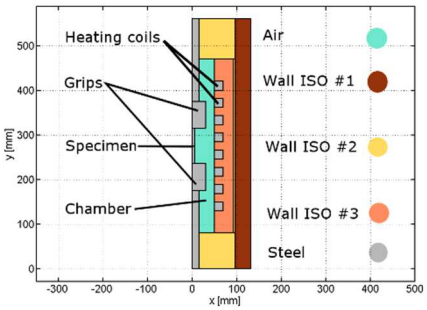


Fig. 5a - Scheme of the reference slice of the axisymmetric COMSOL FE model.

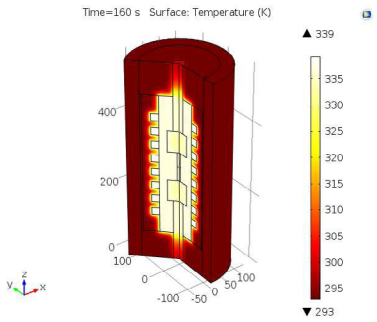


Fig. 5b - Temperature response of the COMSOL FE model in the revolved domain.

The PS in the 3D revolved domain, here numerically simulated, is characterized by the same lateral surface of the physical truss, which has a rectangular solid cross section. Therefore, an equivalent radius equal to 4.77 mm was considered. All furnace characteristics, included dimensions and material properties, were obtained from the manufacturer and are summarized in Table 2.

Table 2. Material parameters characterizing the COMSOL FE model of the furnace.

Material	Density $[kg/m^3]$	Specific heat $[J/(kgK)]$	Thermal conductivity $[W/(mK)]$
Wall ISO#1	300	1050	0.020
Wall ISO#2	128	1130	0.060
Wall ISO#3	880	1100	0.300
Steel (specimen and grips)	7800	466	43
Air (chamber)	1.25	1005	0.024

The thermal load was applied to the heating coil domain in terms of heat source per unit volume. A maximum total power of 6 kW was allowed according to the furnace datasheet. Since air is a very good thermal isolator, the most of heat transfer is caused by mass transportation of convective flows, which mix air from cold and hot regions of the chamber subdomain. Therefore, mixing is crucial to replicate a realistic furnace behavior. To this end, an isothermal constraint was applied to the air subdomain. In order to preserve linearity, the radiation contribution was neglected. Nevertheless, all approximations have proven to be satisfactory and the COMSOL FE model replicates the Könn Furnace response without a need for a Computational Fluid Dynamics (CFD) simulation in the temperature range of interest, i.e. 300-600 K. Figure 6 depicts a contour plot of the temperature response of the Könn Furnace subjected to maximum input power after 100 s. Black arrows represent heat flux streamlines. Figure 7 offers a close up view on the specimen domain, where interface temperatures and exchanged heat fluxes are indicated.

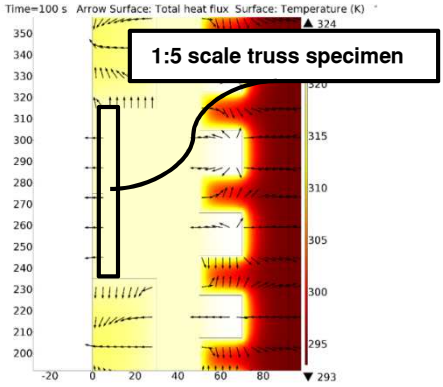


Fig. 6 - Contour plot of the temperature response of the furnace. Black arrows represent thermal flux streamlines.

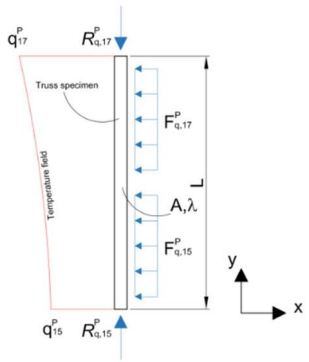


Fig. 7 - Schematic of thermal fluxes applied to the 1:5 scale truss specimen. Vector orientations must be intended as positive according to (8-11).

According to Figure 6, the heat fluxes $\mathbf{F}_{q,15}^{P}$ and $\mathbf{F}_{q,17}^{P}$ enter the specimen through the chamber and flows out to grips at both the specimen ends. Thus, negative heat fluxes $\mathbf{R}_{q,15}^{P}$ and $\mathbf{R}_{q,17}^{P}$ are expected at both ends of the specimen, which interface the PS to the NS.

In order to simulate the overall HS environment, the same PI control strategy of the Könn furnace was implemented and the developed COMSOL FE model simulated the plant. Since fire engineers are used to defining thermal loads in terms of time-temperature fire curves, a hybrid compatible Differential Algebraic Equation (DAE) solver was implemented. It is based on a dual formulation where a set of Lagrange multipliers can impose prescribed trajectories on derivatives of state quantities. Such an approach preserves the stability of the underlying time

stepping scheme. In this particular case the trapezoidal rule was selected as reference algorithm, and the following DAE was taken as a model problem,

$$\begin{aligned}\mathbf{C}_{\theta\theta}\dot{\boldsymbol{\theta}}+\mathbf{K}_{\theta\theta}\boldsymbol{\theta}&=\mathbf{F}_{\theta}(t)+\mathbf{L}_{\theta}^{T}\boldsymbol{\lambda}_{\theta}(t)\\ \mathbf{L}_{\theta}\dot{\boldsymbol{\theta}}&=\mathbf{c}_{\theta}(t)\end{aligned} \tag{14}$$

where the Lagrange multiplier vector $\boldsymbol{\lambda}_q$ consists of additional fictitious heat fluxes that force the temperature response to follow prescribed rates dictated by $\mathbf{c}_{\boldsymbol{q}}(t)$. $\mathbf{L}_q$ is a Boolean matrix that localizes constrained DoFs and, therefore, localizes Lagrange multipliers on the load vector.

4 NUMERICAL SIMULATION OF THS

Several numerical simulations of the proposed THS architecture were executed considering different extended time scales λ. The devised PI controller was selected to impose temperature at furnace nodes by modulating the input heat power of heating coils in the COMSOL FE model of the Könn Furnace. A thermal load case was defined based on the international standard ISO 834 temperature-time fire curve, which is defined in the Eurocode 1 Part 1-2 [12] as:

$$\theta_g = 20+345\cdot\log_{10}(8\cdot t+1) \tag{15}$$

where θ_g is gas or air temperature (°C) and t is time (min). For simplicity, the aforementioned ISO 834 time-temperature curve was applied to the Nodes #13, #15 and #17 of the hybrid model of the truss bridge as temperature history even though in reality the temperature of the element will not exactly follow the heating curve. SI units will be used in the following. Indeed temperature are expressed in K.

a b

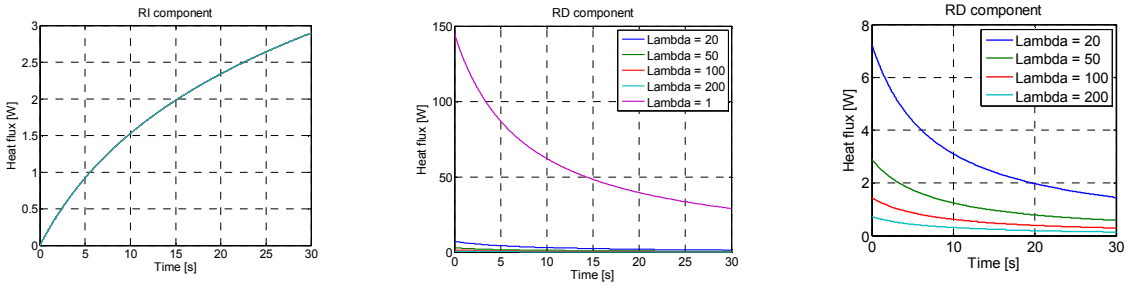
Fig. 8 - ISO 834 time - temperature curve: a) time vs. temperature; b) time vs. temperature rate

Since entailing temperature rate exceeds the Könn Furnace capabilities, which negates for values greater than 1 K/s, an extended time scale λ was considered in the simulated laboratory. Therefore, the test was slowed down of a factor λ and the PS experienced a simulation time step Δt in $\lambda\Delta t$ of wall clock time. In order to investigate the effects of time scaling, a 2-DoFs thermal truss model of the sole 1:5 scale steel specimen was implemented. Material parameters refer to steel; see Table 1 in this respect, whilst geometric values refer to the 1:5 scale truss specimen of Figure 4b. A convective coefficient h equal to 10 $W/(m^2K)$ was considered as a reasonable value. Time constants derived from the 2-DoFs model of the 1:5 scale specimen are 110.56 s and 605.80 s. Since temperature histories were imposed at both specimen ends,

estimations of both rate dependent and rate independent components of thermal fluxes, i.e. $\mathbf{R}_q^{P,RD}$ and $\mathbf{R}_q^{P,RI}$, were estimated,

$$\mathbf{R}_\theta^{P,RD} = \frac{\mathbf{C}_{\theta\theta}^P \dot{\boldsymbol{\theta}}}{\lambda} \qquad (16)$$
$$\mathbf{R}_\theta^{P,RI} = \mathbf{K}_{\theta\theta}^P \boldsymbol{\theta}$$

where $\mathbf{C}_{\theta\theta}^P$ and $\mathbf{K}_{\theta\theta}^P$, are the heat capacity and the conductivity matrices of the 2-DoFs model of the specimen, respectively. Superscripts RD and RI stand for rate dependent and rate independent, respectively. Since the experimental time scale λ affects temperature rates, it reflects on the corresponding heat flux $\mathbf{R}_q^{P,RD}$. Figure 9 reports time histories of both the components. In particular, Figure 9c focuses on rate dependent components for time scaling λ greater than one.



a b c

Fig. 9 - Time histories of feedback heat fluxes components: a) rate independent contribution $\mathbf{R}_q^{P,RI}$; b) rate dependent contribution $\mathbf{R}_q^{P,RD}$; close up view on rate dependent contribution $\mathbf{R}_q^{P,RD}$ in the case of extended time scales

From Figure 16 it is possible to observe that time scalings, which are compatible with the performance of the Könn Furnace, cancel rate dependent contributions, which must be taken into account numerically in the NS. This is the analogous to pseudodynamic (PsD) testing where rate dependent forces, e.g. inertia and viscous damping, are simulated numerically. Estimates of feedback heat fluxes retrieved from the COMSOL FE model were magnified according to the geometric scale factor δ, and accounted for in the thermal balance equation (10-11). A simple proof of this approach can be derived from the Fourier law in the one-dimensional case,

$$\Phi' = -\lambda A \delta^2 \frac{dT}{\delta dx} = \Phi \delta \qquad (17)$$

where Φ' and Φ are heat fluxes related to the full scale model and the scaled specimen, respectively. As can be appreciated from Figure 10, which reports feedback heat fluxes retrieved from the COMSOL, transient dynamics occurs at the beginning of the simulation for all values of time scale λ. Their lengths are constant in the laboratory time and of the same order of magnitude of the lower thermal time constant of the sole specimen, which was estimated at 110 s. Therefore, they can be ascribed to a local transient thermal response of the

furnace/specimen. Figure 11 report the temperature responses of Node #17, which corresponds to the upper node of the specimen, for the Case Study #1.

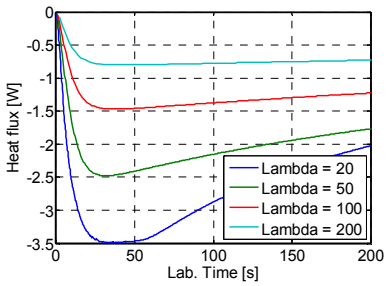


Fig. 10 - Feedback heat flux measured at the upper end of the 1:5 scale specimen on the COMSOL FE model vs. laboratory time (wall clock) for the Case Study #1

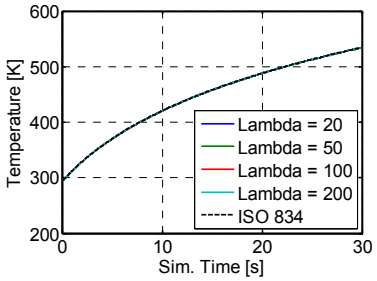
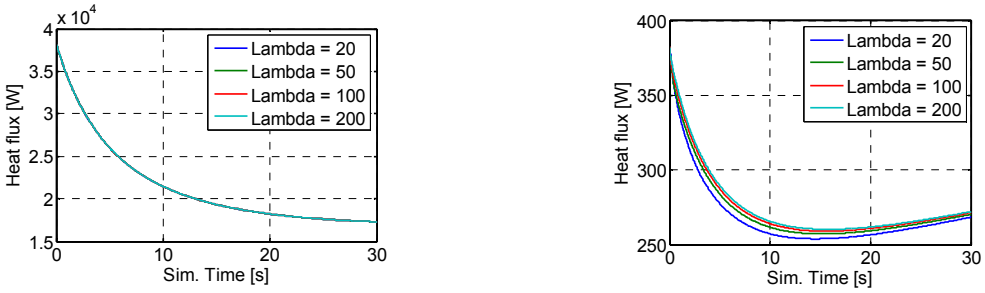


Fig. 11 - Temperature response of Node #15 for different values of time scale λ for the Case Study #1.

As can be appreciated from Figure 11, the devised THS allowed for applying the correct temperature path and simulating the global system response in a consistent manner. Potential aftermath of substructures partial and full coupling can be deducted more clearly by Lagrange multipliers behavior. In greater detail, Figure 12 reports the time histories of the Lagrange multiplier applied at Node #15, where the ISO 834 was prescribed for both the case studies.



a b

Fig. 12 - Time histories of Lagrange multipliers at Node #15 for different values of time scale λ for: a) Case Study #1; and b) Case Study #2.

In detail, Figure 12a, which refers to Case Study #1, describes that when thermal diffusivity is on the order of 12 mm^2/s and h/k = 1.16 m/m, specimen feedback heat fluxes are negligible because no effects can be appreciated on Lagrange multipliers that impose the temperature path on Node #15 as fictitious heat fluxes. As a consequence, a substructure partial coupling can be applied. Conversely, from Figure 12b, which refers to Case Study #2, when thermal diffusivity is on the order of 1200 mm^2/s and h/k = 0.02 m/m, feedback heat fluxes affect the dynamic solution and should be taken into account.

5 CONCLUSIONS AND FUTURE PERSPECTIVES

The present paper summarizes all research activities focused on the development of a pure THS. The full set of coupled equations defining the thermomechanical problem was discussed from the HS perspective and particular care was given to the heat transfer physics. The study highlights the feasibility of the THS by applying an arbitrary temperature path to both case

studies and simulating the global system response in a consistent manner. Results have shown that specimen feedback heat fluxes can be neglected in the global heat balance when an is element characterized by thermal diffusivity on the order of 12 mm^2/s - Case Study #1 – and $h/k \approx 1$. Therefore, a substructure partial coupling can be applied. Conversely, when thermal diffusivity is higher (1200 mm^2/s) and conduction offers very little resistance to heat transfer with respect to convection ($h/k << 1$) - Case Study #2 -, feedback fluxes affect the dynamic solution and must be taken into account using a substructure full coupling. A forthcoming experimental campaign will allow for updating the developed simulation environment and tuning an effective testing protocol. The presented approach provides a realistic insight into the capabilities of THS and paves the way for simulating hot NSs for the purpose of TMHS.

REFERENCES

[1] Mostafaei, H. (2013a). Hybrid fire testing for assessing performance of structures in fire -Application. Fire Safety Journal, 56, 30–38. doi:10.1016/j.firesaf.2012.12.003

[2] Mostafaei, H. (2013b). Hybrid fire testing for assessing performance of structures in fire -Methodology. Fire Safety Journal, 58, 170–179. doi:10.1016/j.firesaf.2013.02.005

[3] Franssen, J. M. (2005). SAFIR: A thermal/structural program for modeling structures under fire. Engineering Journal-American Institute of Steel Construction Inc, 42(3).

[4] Whyte, C. A., K.R. Mackie, and B. Stojadinovic, (2014). Hybrid Simulation of Thermomechanical Structural Response. ASCE Journal of Structural Engineering (in review).

[5] OpenFresco, (2014). Open Framework for Experimental Setup and Control. http://openfresco.neesforge.nees.org.

[6] OpenSees, (2014). Open System for Earthquake Engineering Simulation. http://opensees.berkeley.edu.

[7] Eslami, M., & Richard, B. (2008). Thermal Stresses: Advanced Theory and Applications.

[8] Shing, P. B. (2008), "Real-time hybrid testing techniques" in Modern testing techniques for structural systems Bursi O.S., Wagg D.J.E. (eds.), CISM International Centre for Mechanical Sciences, vol 502. Springer, Vienna.

[9] Bursi, O.S., Buelga, A.G., Vulcan, L., Neild, S.A. and Wagg, D.J. (2008), "Novel coupling Rosenbrock-based algorithms for real-time dynamic substructure testing", Earthquake Engineering and Structural Dynamics, 37:339-360.

[10] Yong Gao, Billie F. Spencer, Jr., (2008). Structural Health Monitoring Strategies for Smart Sensor Networks Report Series Report No. NSEL-011 Department of Civil and Environmental Engineering University of Illinois at Urbana-Champaign.

[11] Pauli, J., Somaini, D., Knobloch, M., Fontana, M., (2012). "Experiments on steel columns under fire conditions." Institute of Structural Engineering Technical Report 340, Swiss Federal Institute of Technology, Zurich, Switzerland.

[12] European Committee for Standardization (CEN). (2002). "EN1991-1-2, Eurocode 1: Actions on structures – Part 1-2: General actions – Actions on structures exposed to fire," Brussels.

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis (Eds)

HYBRID SIMULATION OF COMPLEX STRUCTURAL SYSTEMS BASED ON PARTITIONED TIME INTEGRATION SCHEMES

Giuseppe Abbiati*, Enrico Cazzador†, Oreste S. Bursi†, Pierre Pegon~, Francisco Javier Molina~ and Fabrizio Paolacci§.

* Department of Civil, Environmental and Geomatic Engineering (D-BAUG), IBK, ETH Zurich
Stefano-Franscini-Platz 5, 8093 Zürich, Switzerland
e-mail: abbiati@ibk.baug.ethz.ch

† Department of Civil, Environmental and Mechanical Engineering, University of Trento
Via Mesiano 77, 38123, Trento, Italy
e-mail: oreste.bursi@unitn.it, enrico.cazzador@unitn.it

§ Department of Engineering, University of Roma Tre
Via Vito Volterra 62, Rome, Italy
e-mail: fabrizio.paolacci@uniroma3.it

~ European Commission, Joint Research Centre (JRC), Institute for the Protection and Security of the Citizen,
European Laboratory for Structural Assessment Unit, Via Enrico Fermi 2749, 21027 Ispra (VA), Italy
e-mail: pierre.pegon@jrc.ec.europa.eu

Key words: Hybrid Simulations, Finite element, Partitioned algorithms, Structural dynamics, Reinforced concrete old viaduct.

Abstract. In the last two decades, the increasing complexity of engineering systems boosted the development of very efficient simulation methods based on partitioning. In view of coupling dynamic parts of hybrid systems, the finite element tearing and interconnecting approach emerged as the most promising technique. Nonetheless, there is still a lack of a comprehensive study of algorithmic performances from the experimental perspective. In this view, the present paper sheds light on the application of two well-known parallel partitioning methods for the purpose of the simulation of hybrid models. Thus, an existent reinforced concrete bridge is chosen as a benchmark case study. In order to perform hybrid simulations, a novel coupling software was devised. It allowed for combining two physical piers to the numerical model of the remaining part of the bridge. As a result, successful tests were conducted at the ELSA laboratory of the Joint Research Centre of Ispra (Italy).

1 INTRODUCTION

In view of coupling dynamic parts of hybrid systems, partitioning methods are more than ever appealing and arouse a lot of interest in the context of hybrid experimental/numerical simulation techniques [1]. In fact, they are prone to combine a Physical Substructure (PS), which is the key region of interest of the emulated heterogeneous system, and a Numerical Substructure (NS) collecting all well-known subparts. During a heterogeneous simulation, a compliant time integrator calculates the interaction between NS and PS by solving the equation of motion of the entire system. Since Finite Element Tearing and Interconnecting (FETI) approaches allow for the concurrent solution of involved subdomains with large time step ratios, they guarantiee

a smooth stream of displacement commands on the PS, regardless the coarser time grid on the NS. In this context, [2] proved that velocity continuity on interface DoFs is the key solution to obtain unconditionally stable algorithms as long as all individual subdomains satisfy their own stability requirements. Along this line, Pegon and Magonette developed the PM method [3], which paved the way for parallel implementations of the FETI tailored to hybrid models. Interface energy dissipation and one order of accuracy loss in the subcycling case were stressed as major drawbacks. Moreover, the PM method needs for a starting procedure. Brun et al. proposed an improved variant of the PH algorithm [4], namely the modified PH method [5], which allows for computations in subdomains in a completely concurrent manner with no need for initialization and energy conserving.

Mixed implicit/explicit multi-time partitioning methods tailored to Newmark schemes are ripe for a profitable application to hybrid simulations. Nevertheless, there is still a lack of a comprehensive study of algorithmic performances from an experimental perspective. Along this line, this paper presents algorithmic implementations of both the PM and modified PH methods tailored to both numerical/physical nonlinear systems. A twelve-pier reinforced concrete bridge is introduced as reference case study. Then, numerical simulations of the dynamic response of the bridge emphasize pros and cons of both algorithms. Finally, the experimental implementation of the PM method, which allows for conducting hybrid simulations of aforementioned bridge at the ELSA Laboratory of the Joint Research Centre of Ispra (Italy) is presented.

2 PARTITIONED INTEGRATORS FOR SUBDOMAIN COUPLING IN HYBRID SIMULATIONS

Both the PM and the modified PH algorithms are presented in this section. In detail, all implementations refer to the following set of coupled of nonlinear dynamic equations,

$$\begin{cases} \mathbf{M}^A\ddot{\mathbf{u}}^A + \mathbf{R}^A\left(\mathbf{u}^A, \dot{\mathbf{u}}^A\right) + \mathbf{L}^{A^T}\boldsymbol{\Lambda} = \mathbf{F}^A\left(t\right) \\ \mathbf{M}^B\ddot{\mathbf{u}}^B + \mathbf{R}^B\left(\mathbf{u}^B, \dot{\mathbf{u}}^B\right) + \mathbf{L}^{B^T}\boldsymbol{\Lambda} = \mathbf{F}^B\left(t\right) \end{cases} \tag{1}$$

$$\mathbf{L}^A\dot{\mathbf{u}}^A + \mathbf{L}^B\dot{\mathbf{u}}^B = \mathbf{0} \tag{2}$$

where, with regard to the generic subdomain k, $\mathbf{M}^k$ is the mass matrix; $\mathbf{u}^k, \dot{\mathbf{u}}^k$ and $\ddot{\mathbf{u}}^k$ are displacement, velocity and acceleration vectors, $\mathbf{R}^k$ is the restoring force and $\mathbf{F}^k(t)$ represents the external time-varying load. $\mathbf{L}^k$ is a Boolean matrix that collocates interface DoFs on the related subdomain. According to [2], in order to obtain unconditionally stable algorithms as long as all individual subdomains satisfy their own stability requirements, velocity continuity on interface DoFs is pursued. For the sake of clarity, zero tangent stiffness $\mathbf{K}^k$ and damping $\mathbf{C}^k$matrices read,

$$\mathbf{K}^k = \left.\frac{\partial \mathbf{R}^k}{\partial \mathbf{u}^k}\right|_{\mathbf{u}^k, \dot{\mathbf{u}}^k = \mathbf{0}}, \mathbf{C}^k = \left.\frac{\partial \mathbf{R}^k}{\partial \dot{\mathbf{u}}^k}\right|_{\mathbf{u}^k, \dot{\mathbf{u}}^k = \mathbf{0}} \tag{3}$$

Both coupling method are supposed to handle a pair of arbitrary Newmark schemes with their own characteristic parameters, i.e. β_A, γ_A, and β_B, γ_B, for subdomains A and B, respectively [6].

2.1 The PM method

The task sequence of the PM method, analyzed in depth in [7] is sketched in Figure 1. It clearly recalls the GC method [2].However, a time step equal to $2\Delta t_A$ is exploited in subdomain A in order to anticipate information on the subdomain B at the beginning of a new time step t_{n+1}.

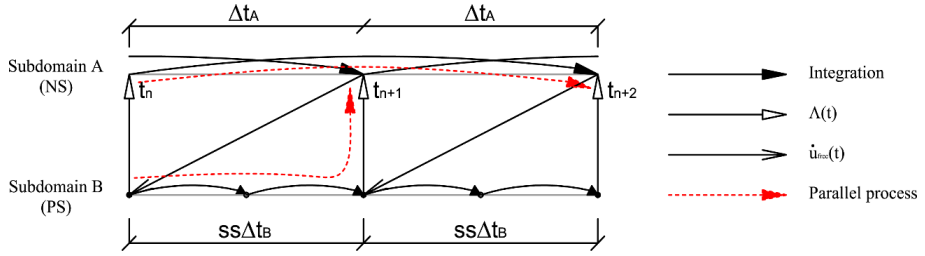


Fig. 1 - Task sequence of the PM method algorithm.

The dashed line describes the ongoing parallel process in the two subdomains. Note that the synchronized exchange of information enables the parallel computation of both subdomains. The PM method is briefly summarized hereinafter.

Subdomain A

if n == 1

Initialization step on subdomain A from $\boldsymbol{t_n}$ to $\boldsymbol{t_{n+1}}$

$$\tilde{\mathbf{u}}_{n+1}^{A,free} = \mathbf{u}_n^A + \dot{\mathbf{u}}_n^A \Delta t^A + \ddot{\mathbf{u}}_n^A \left(1/2 - \beta^A\right)\left(\Delta t^A\right)^2$$

$$\tilde{\dot{\mathbf{u}}}_{n+1}^{A,free} = \dot{\mathbf{u}}_n^A + \ddot{\mathbf{u}}_n^A \left(1-\gamma^A\right)\left(\Delta t^A\right)$$

$$\ddot{\mathbf{u}}_{n+1}^{A,free} = \mathbf{D}_1^{A^{-1}}\left(\mathbf{F}_{n+1}^A - \mathbf{R}^A\left(\tilde{\mathbf{u}}_{n+1}^{A,free}, \tilde{\dot{\mathbf{u}}}_{n+1}^{A,free}\right)\right)$$

$$\dot{\mathbf{u}}_{n+1}^{A,free} = \tilde{\dot{\mathbf{u}}}_{n+1}^{A,free} + \ddot{\mathbf{u}}_{n+1}^{A,free}\gamma^A\left(\Delta t^A\right)$$

$$\mathbf{u}_{n+1}^{A,free} = \tilde{\mathbf{u}}_{n+1}^{A,free} + \ddot{\mathbf{u}}_{n+1}^{A,free}\beta^A\left(\Delta t^A\right)^2$$

end

Solution of the free problem on subdomain A from $\boldsymbol{t_n}$ to $\boldsymbol{t_{n+2}}$

$$\tilde{\mathbf{u}}_{n+2}^{A,free} = \mathbf{u}_n^A + \dot{\mathbf{u}}_n^A 2\Delta t^A + \ddot{\mathbf{u}}_n^A \left(1/2 - \beta^A\right)\left(2\Delta t^A\right)^2$$

$$\tilde{\dot{\mathbf{u}}}_{n+2}^{A,free} = \dot{\mathbf{u}}_n^A + \ddot{\mathbf{u}}_n^A \left(1-\gamma^A\right)\left(2\Delta t^A\right)$$

$$\ddot{\mathbf{u}}_{n+2}^{A,free} = \mathbf{D}_2^{A^{-1}}\left(\mathbf{F}_{n+2}^A - \mathbf{R}^A\left(\tilde{\mathbf{u}}_{n+2}^{A,free}, \tilde{\dot{\mathbf{u}}}_{n+2}^{A,free}\right)\right)$$

Subdomain B

for = j 1:1:ss

Solution of the free problem on subdomain B from $\boldsymbol{t_{n+\frac{j-1}{ss}}}$ to $\boldsymbol{t_{n+\frac{j}{ss}}}$

$$\tilde{\mathbf{u}}_{n+\frac{j}{ss}}^{B,free} = \mathbf{u}_{n+\frac{j-1}{ss}}^B + \dot{\mathbf{u}}_{n+\frac{j-1}{ss}}^B \Delta t^B + \ddot{\mathbf{u}}_{n+\frac{j-1}{ss}}^B \left(1/2 - \beta^B\right)\Delta t^{B^2}$$

$$\tilde{\dot{\mathbf{u}}}_{n+\frac{j}{ss}}^{B,free} = \dot{\mathbf{u}}_{n+\frac{j-1}{ss}}^B + \ddot{\mathbf{u}}_{n+\frac{j-1}{ss}}^B \left(1-\gamma^B\right)\Delta t^B$$

$$\ddot{\mathbf{u}}_{n+\frac{j}{ss}}^{B,free} = \mathbf{D}^{B^{-1}}\left(\mathbf{F}_{n+\frac{j}{ss}}^B - \mathbf{R}^B\left(\tilde{\mathbf{u}}_{n+\frac{j}{ss}}^{B,free}, \tilde{\dot{\mathbf{u}}}_{n+\frac{j}{ss}}^{B,free}\right)\right)$$

$$\dot{\mathbf{u}}_{n+\frac{j}{ss}}^{B,free} = \tilde{\dot{\mathbf{u}}}_{n+\frac{j}{ss}}^{B,free} + \ddot{\mathbf{u}}_{n+\frac{j}{ss}}^{B,free}\gamma^B\Delta t^B$$

$$\mathbf{u}_{n+\frac{j}{ss}}^{B,free} = \tilde{\mathbf{u}}_{n+\frac{j}{ss}}^{B,free} + \ddot{\mathbf{u}}_{n+\frac{j}{ss}}^{B,free}\beta^B\Delta t^{B^2}$$

Interpolation of free quantities on subdomain A

$$\dot{\mathbf{u}}_{n+\frac{j}{ss}}^{A,free} = \left(1-\frac{j}{ss}\right)\dot{\mathbf{u}}_n^{A,free} + \left(\frac{j}{ss}\right)\dot{\mathbf{u}}_{n+1}^{A,free}$$

Calculation of interface Lagrange multipliers

$$\boldsymbol{\Lambda}_{n+\frac{j}{ss}} = \mathbf{H}^{AB^{-1}}\left(\mathbf{L}^A\dot{\mathbf{u}}_{n+\frac{j}{ss}}^{A,free} + \mathbf{L}^B\dot{\mathbf{u}}_{n+\frac{j}{ss}}^{B,free}\right)$$

Solution of the link problem on subdomain B

$$\ddot{\mathbf{u}}_{n+\frac{j}{ss}}^{B,link} = -\mathbf{D}^{B^{-1}}\mathbf{L}^{B^T}\boldsymbol{\Lambda}_{n+\frac{j}{ss}}$$

$$\dot{\mathbf{u}}_{n+2}^{A,\,free} = \dot{\tilde{\mathbf{u}}}_{n+2}^{A,\,free} + \ddot{\mathbf{u}}_{n+2}^{A,\,free}\gamma^{A}\left(2\Delta t^{A}\right)$$

$$\mathbf{u}_{n+2}^{A,\,free} = \tilde{\mathbf{u}}_{n+2}^{A,\,free} + \ddot{\mathbf{u}}_{n+2}^{A,\,free}\beta^{A}\left(2\Delta t^{A}\right)^{2}$$

Solution of the link problem on subdomain A

$$\ddot{\mathbf{u}}_{n+1}^{A,link} = -\mathbf{D}_{m}^{A^{-1}}\mathbf{L}^{A^{T}}\boldsymbol{\Lambda}_{n+1}$$

$$\dot{\mathbf{u}}_{n+1}^{A,link} = \ddot{\mathbf{u}}_{n+1}^{A,link}\gamma^{A}\left(m\Delta t^{A}\right)$$

$$\mathbf{u}_{n+1}^{A,link} = \ddot{\mathbf{u}}_{n+1}^{A,link}\beta^{A}\left(m\Delta t^{B}\right)^{2}$$

Coupled solution on subdomain A

$$\ddot{\mathbf{u}}_{n+1}^{A} = \ddot{\mathbf{u}}_{n+1}^{A,\,free} + \ddot{\mathbf{u}}_{n+1}^{A,link}$$

$$\dot{\mathbf{u}}_{n+1}^{A} = \dot{\mathbf{u}}_{n+1}^{A,\,free} + \dot{\mathbf{u}}_{n+1}^{A,link}$$

$$\mathbf{u}_{n+1}^{A} = \mathbf{u}_{n+1}^{A,\,free} + \mathbf{u}_{n+1}^{A,link}$$

$$\dot{\mathbf{u}}_{n+\frac{j}{ss}}^{B,link} = \ddot{\mathbf{u}}_{n+\frac{j}{ss}}^{B,link}\gamma^{B}\Delta t^{B}$$

$$\mathbf{u}_{n+\frac{j}{ss}}^{B,link} = \ddot{\mathbf{u}}_{n+\frac{j}{ss}}^{B,link}\beta^{B}\Delta t^{B^{2}}$$

Coupled solution on subdomain B

$$\ddot{\mathbf{u}}_{n+\frac{j}{ss}}^{B} = \ddot{\mathbf{u}}_{n+\frac{j}{ss}}^{B,\,free} + \ddot{\mathbf{u}}_{n+\frac{j}{ss}}^{B,link}$$

$$\dot{\mathbf{u}}_{n+\frac{j}{ss}}^{B} = \dot{\mathbf{u}}_{n+\frac{j}{ss}}^{B,\,free} + \dot{\mathbf{u}}_{n+\frac{j}{ss}}^{B,link}$$

$$\mathbf{u}_{n+\frac{j}{ss}}^{B} = \mathbf{u}_{n+\frac{j}{ss}}^{B,\,free} + \mathbf{u}_{n+\frac{j}{ss}}^{B,link}$$

end

where,

$$\mathbf{D}_{m}^{A} = \mathbf{M}^{A} + \mathbf{C}^{A}\gamma^{A}m\Delta t^{A} + \mathbf{K}^{A}\beta^{A}\left(m\Delta t^{A}\right)^{2} \tag{4}$$

$$\mathbf{D}^{B} = \mathbf{M}^{B} + \mathbf{C}^{B}\gamma^{B}\Delta t^{B} + \mathbf{K}^{B}\beta^{B}(\Delta t^{B})^{2} \tag{5}$$

Since the PM method needs to be initialized, m is equal to 1 when n is equal to 0, and the PM method is equivalent to the GC method, which is a staggered procedure. After the initialization step, i.e. n > 0, m is equal to 2 and the two-step forward prediction allows for parallel implementations. For the purpose of HS, the restoring force of the PS $\mathbf{R}^{B}$is measured on physical specimens, after applying the trial displacement $\tilde{\mathbf{u}}_{n+\frac{j}{ss}}^{B,free}$. Moreover, based on previous equations, the Steklov-Poincare' operator $\mathbf{H}^{AB}$, can be quantified as follows:

$$\mathbf{L}^{A}\dot{\mathbf{u}}_{n+\frac{j}{ss}}^{A} + \mathbf{L}^{B}\dot{\mathbf{u}}_{n+\frac{j}{ss}}^{B} = \mathbf{L}^{A}\left(\dot{\mathbf{u}}_{n+\frac{j}{ss}}^{A,\,free} + \dot{\mathbf{u}}_{n+\frac{j}{ss}}^{A,link}\right) + \mathbf{L}^{B}\left(\dot{\mathbf{u}}_{n+\frac{j}{ss}}^{B,\,free} + \dot{\mathbf{u}}_{n+\frac{j}{ss}}^{B,link}\right) \tag{6}$$

$$\mathbf{L}^{A}\left(\dot{\mathbf{u}}_{n+\frac{j}{ss}}^{A,\,free}\right) + \mathbf{L}^{B}\left(\dot{\mathbf{u}}_{n+\frac{j}{ss}}^{B,\,free}\right) = \left(\mathbf{L}^{A}\mathbf{D}^{A^{-1}}\mathbf{L}^{A^{T}}\gamma^{A}\left(m\Delta t^{A}\right) + \mathbf{L}^{B}\mathbf{D}^{B^{-1}}\mathbf{L}^{B^{T}}\gamma^{B}\Delta t^{B}\right)\boldsymbol{\Lambda}_{n+\frac{j}{ss}} = \mathbf{H}^{AB}\boldsymbol{\Lambda}_{n+\frac{j}{ss}} \tag{7}$$

So, the $\Lambda_{n+\frac{j}{ss}}$ can be explicitly evaluated to cope with the synchronization requirements of HS. It is evident that the proposed method nicely suites the requirements of HS.

2.2 The modified PH method

The task sequence of the modified PH method [5] is sketched in Figure 2. As for the PM method, it clearly recalls the PH method [4]. However, the evaluation of the Lagrange multiplier vector Λ is done at the end of each macro time step Δt^A.

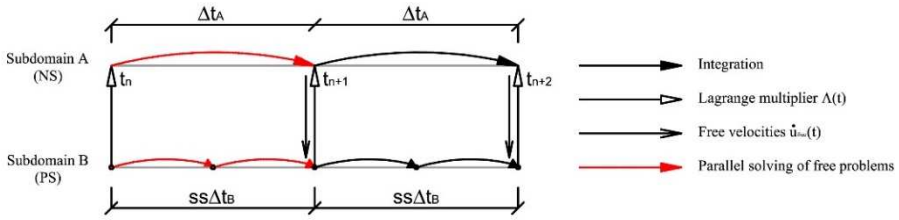


Fig. 2 - Task sequence of the modified PH method algorithm.

The red line describes the ongoing parallel process in the two subdomains. Note that the synchronized exchange of information enables the parallel computation of both subdomains. The modified PH method is briefly summarized herein.

Subdomain A

Solution of the free problem on subdomain A from $\boldsymbol{t_n}$ to $\boldsymbol{t_{n+1}}$

$$\tilde{\mathbf{u}}_{n+1}^{A,free} = \mathbf{u}_n^A + \dot{\mathbf{u}}_n^A \Delta t^A + \ddot{\mathbf{u}}_n^A \left(1/2 - \beta^A\right) \Delta t^{A^2}$$

$$\tilde{\dot{\mathbf{u}}}_{n+1}^{A,free} = \dot{\mathbf{u}}_n^A + \ddot{\mathbf{u}}_n^A \left(1 - \gamma^A\right) \Delta t^A$$

$$\ddot{\mathbf{u}}_{n+1}^{A,free} = \mathbf{D}^{A^{-1}} \left(\mathbf{F}_{n+1}^A - \mathbf{R}^A \left(\tilde{\mathbf{u}}_{n+1}^{A,free}, \tilde{\dot{\mathbf{u}}}_{n+1}^{A,free} \right) \right)$$

$$\dot{\mathbf{u}}_{n+1}^{A,free} = \tilde{\dot{\mathbf{u}}}_{n+1}^{A,free} + \ddot{\mathbf{u}}_{n+1}^{A,free} \gamma^A \Delta t^A$$

$$\mathbf{u}_{n+1}^{A,free} = \tilde{\mathbf{u}}_{n+1}^{A,free} + \ddot{\mathbf{u}}_{n+1}^{A,free} \beta^A \Delta t^{A^2}$$

Subdomain B

$$\mathbf{u}_n^{B,free} = \mathbf{u}_n^B, \dot{\mathbf{u}}_n^{B,free} = \dot{\mathbf{u}}_n^B$$

for j = 1:1:ss

Solution of the free problem on subdomain B from $\boldsymbol{t_n}$ to $\boldsymbol{t_{n+1}}$ in $\boldsymbol{ss}$ substeps

$$\tilde{\mathbf{u}}_{n+\frac{j}{ss}}^{B,free} = \mathbf{u}_{n+\frac{j-1}{ss}}^{B,free} + \dot{\mathbf{u}}_{n+\frac{j-1}{ss}}^{B,free} \Delta t^B + \ddot{\mathbf{u}}_{n+\frac{j-1}{ss}}^{B,free} \left(1/2 - \beta^B\right) \Delta t^{B^2}$$

$$\tilde{\dot{\mathbf{u}}}_{n+\frac{j}{ss}}^{B,free} = \dot{\mathbf{u}}_{n+\frac{j-1}{ss}}^{B,free} + \ddot{\mathbf{u}}_{n+\frac{j-1}{ss}}^{B,free} \left(1 - \gamma^B\right) \Delta t^B$$

$$\ddot{\mathbf{u}}_{n+\frac{j}{ss}}^{B,free} = \mathbf{D}^{B^{-1}} \left(\mathbf{F}_{n+\frac{j}{ss}}^A - \mathbf{R}^B \left(\tilde{\mathbf{u}}_{n+\frac{j}{ss}}^{B,free}, \tilde{\dot{\mathbf{u}}}_{n+\frac{j}{ss}}^{B,free} \right) \right)$$

$$\dot{\mathbf{u}}_{n+\frac{j}{ss}}^{B,free} = \tilde{\dot{\mathbf{u}}}_{n+\frac{j}{ss}}^{B,free} + \ddot{\mathbf{u}}_{n+\frac{j}{ss}}^{B,free} \gamma^B \Delta t^B$$

$$\mathbf{u}_{n+\frac{j}{ss}}^{B,free} = \tilde{\mathbf{u}}_{n+\frac{j}{ss}}^{B,free} + \ddot{\mathbf{u}}_{n+\frac{j}{ss}}^{B,free} \beta^B \Delta t^{B^2}$$

end

Calculation of interface Lagrange multipliers

$$\Lambda_{n+1} = \mathrm{H}^{AB^{-1}} \mathrm{V}_{n+1}$$

$$\mathbb{V}_{n+1}^A = \left[\ddot{\mathbf{u}}_{n+1}^{A,free^T} \quad \dot{\mathbf{u}}_{n+1}^{A,free^T} \quad \mathbf{u}_{n+1}^{A,free^T} \right]^T \qquad \mathbb{V}_{n+\frac{j}{ss}}^B = \left[\ddot{\mathbf{u}}_{n+\frac{j}{ss}}^{B,free^T} \quad \dot{\mathbf{u}}_{n+\frac{j}{ss}}^{B,free^T} \quad \mathbf{u}_{n+\frac{j}{ss}}^{B,free^T} \right]^T$$

$$\mathrm{V}_{n+1} = \left[\mathbb{V}_{n+\frac{1}{ss}}^{B^T} \quad \cdots \quad \mathbb{V}_{n+\frac{j}{ss}}^{B^T} \quad \cdots \quad \mathbb{V}_{n+1}^{B^T} \quad \mathbb{V}_{n+1}^{A^T} \right]^T$$

Calculation of link solutions

$$\mathrm{W}_{n+1} = \mathrm{Y}^{AB} \Lambda_{n+1}$$

$$\mathrm{W}_{n+1}=\left[\begin{array}{ccccccc}\mathbb{W}^{B^T}_{n+\frac{1}{ss}} & \cdots & \mathbb{W}^{B^T}_{n+\frac{j}{ss}} & \cdots & \mathbb{W}^{B^T}_{n+1} & \mathbb{W}^{A^T}_{n+1}\end{array}\right]^T$$

$$\mathbb{W}^{A}_{n+1}=\left[\begin{array}{ccc}\ddot{\mathbf{u}}^{A,link^T}_{n+1} & \dot{\mathbf{u}}^{A,link^T}_{n+1} & \mathbf{u}^{A,link^T}_{n+1}\end{array}\right]^T \quad \mathbb{W}^{B}_{n+\frac{j}{ss}}=\left[\begin{array}{ccc}\ddot{\mathbf{u}}^{B,link^T}_{n+\frac{j}{ss}} & \dot{\mathbf{u}}^{B,link^T}_{n+\frac{j}{ss}} & \mathbf{u}^{B,link^T}_{n+\frac{j}{ss}}\end{array}\right]^T$$

Calculation of the coupled solution

$$\mathrm{U}_{n+1}=\mathrm{V}_{n+1}+\mathrm{W}_{n+1}$$

$$\mathrm{U}_{n+1}=\left[\begin{array}{ccccccc}\mathbb{U}^{B^T}_{n+\frac{1}{ss}} & \cdots & \mathbb{U}^{B^T}_{n+\frac{j}{ss}} & \cdots & \mathbb{U}^{B^T}_{n+1} & \mathbb{U}^{A^T}_{n+1}\end{array}\right]^T$$

$$\mathbb{U}^{A}_{n+1}=\left[\begin{array}{ccc}\ddot{\mathbf{u}}^{A^T}_{n+1} & \dot{\mathbf{u}}^{A^T}_{n+1} & \mathbf{u}^{A^T}_{n+1}\end{array}\right]^T \quad \mathbb{U}^{B}_{n+\frac{j}{ss}}=\left[\begin{array}{ccc}\ddot{\mathbf{u}}^{B^T}_{n+\frac{j}{ss}} & \dot{\mathbf{u}}^{B^T}_{n+\frac{j}{ss}} & \mathbf{u}^{B^T}_{n+\frac{j}{ss}}\end{array}\right]^T$$

where,

$$\mathbf{D}^A=\mathbf{M}^A+\mathbf{C}^A\gamma^A\Delta^A_t+\mathbf{K}^A\beta^A\Delta t^{A^2} \tag{8}$$

$$\mathbf{D}^B=\mathbf{M}^B+\mathbf{C}^B\gamma^B\Delta^B_t+\mathbf{K}^B\beta^B\Delta t^{B^2} \tag{9}$$

Derivation of both coupling operators H^{AB} and Y^{AB} are briefly summarized. They can be easily derived on the linearized system of coupled equations (1-2) rearranged in matrix form. To this end, the Newmark algorithm is condensed to the following matrix equation,

$$\mathbb{M}^k\mathbb{U}^k_{n+1}=\mathbb{F}^k_{n+1}-\mathbb{N}^k\mathbb{U}^k_n \tag{10}$$

where,

$$\mathbb{M}^k=\left[\begin{array}{ccc}\mathbf{M}^k & \mathbf{0} & \mathbf{K}^k\\ -\gamma^k\Delta t^k\mathbf{I} & \mathbf{I} & \mathbf{0}\\ -\beta^k\Delta t^{k^2}\mathbf{I} & \mathbf{0} & \mathbf{I}\end{array}\right] \qquad \mathbb{N}^k=\left[\begin{array}{ccc}\mathbf{0} & \mathbf{0} & \mathbf{0}\\ -\left(1-\gamma^k\right)\Delta t^k\mathbf{I} & -\mathbf{I} & \mathbf{0}\\ -\left(1/2-\beta^k\right)\Delta t^{k^2}\mathbf{I} & -\Delta t^k\mathbf{I} & -\mathbf{I}\end{array}\right] \tag{11}$$

$$\mathbb{F}^k_{n+1}=\left[\begin{array}{c}\mathbf{F}^k_{n+1}\\ \mathbf{0}\\ \mathbf{0}\end{array}\right] \qquad \mathbb{U}^k_{n+1}=\left[\begin{array}{c}\ddot{\mathbf{u}}^k_{n+1}\\ \dot{\mathbf{u}}^k_{n+1}\\ \mathbf{u}^k_{n+1}\end{array}\right] \tag{12}$$

Accordingly, for a linearized system, the modified PH method can be expressed as,

$$\begin{bmatrix} \mathbb{M}^B & & & & & (1/ss)\mathbb{L}^B \\ \mathbb{N}^B & \mathbb{M}^B & & & & (2/ss)\mathbb{L}^B \\ & \ddots & \ddots & & & \vdots \\ & & \mathbb{N}^B & \mathbb{M}^B & & \mathbb{L}^B \\ & & & \mathbb{N}^B & \mathbb{M}^A & \mathbb{L}^A \\ & & & \mathbb{B}^B & \mathbb{B}^A & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbb{U}^B_{n+\frac{1}{ss}} \\ \mathbb{U}^B_{n+\frac{2}{ss}} \\ \vdots \\ \mathbb{U}^B_{n+1} \\ \mathbb{U}^A_{n+1} \\ \mathbf{\Lambda}_{n+1} \end{bmatrix} = \begin{bmatrix} \mathbb{F}^B_{n+\frac{1}{ss}} - \mathbb{N}^B\mathbb{U}^B_n - (1-1/ss)\mathbb{L}^B\mathbf{\Lambda}_n \\ \mathbb{F}^B_{n+\frac{2}{ss}} - (1-2/ss)\mathbb{L}^B\mathbf{\Lambda}_n \\ \vdots \\ \mathbb{F}^B_{n+1} \\ \mathbb{F}^A_{n+1} - \mathbb{N}^A\mathbb{U}^A_n \\ \mathbf{0} \end{bmatrix} \tag{13}$$

Boolean matrices $\mathbb{L}^k$ and $\mathbb{B}^k$ localize Lagrange multipliers and coupling DoFs, respectively. Since the coupling of velocities was proved to favor the stability of the global scheme, they read,

$$\mathbb{L}^k = \begin{bmatrix} \mathbf{L}^{k^T} & \mathbf{0} & \mathbf{0} \end{bmatrix}^T, \quad \mathbb{B}^k = \begin{bmatrix} \mathbf{0} & \mathbf{L}^k & \mathbf{0} \end{bmatrix} \tag{14}$$

The linear system (13) can be solved using a bordered system approach. Subdomain matrix blocks can be gathered together and the system expressed as:

$$\begin{bmatrix} \mathrm{M} & \mathrm{L} \\ \mathrm{B} & 0 \end{bmatrix} \begin{bmatrix} \mathrm{U}_{n+1} \\ \Lambda_{n+1} \end{bmatrix} = \begin{bmatrix} \mathrm{F}_{n+1} \\ 0 \end{bmatrix} \tag{15}$$

Since the global solution is the sum of free and link quantities,

$$\mathrm{U}_{n+1} = \mathrm{V}_{n+1} + \mathrm{W}_{n+1} \tag{16}$$

where,

$$\mathrm{V}_{n+1} = \mathrm{M}^{-1}\mathrm{F}_{n+1} \tag{17}$$

$$\mathrm{W}_{n+1} = -\mathrm{M}^{-1}\mathrm{L}\Lambda_{n+1} = \mathrm{Y}^{AB}\Lambda_{n+1} \tag{18}$$

The second block row of (24) can be split as follows,

$$\mathrm{B}\mathrm{U}_{n+1} = \mathrm{B}\mathrm{V}_{n+1} + \mathrm{B}\mathrm{W}_{n+1} = 0 \tag{19}$$

Accordingly, Lagrange multipliers λ_{n+1} can be easily expressed through the Steklov-Poincaré operator H^{AB}, which is the dual Shur complement of the block matrix reported in (15),

$$\Lambda_{n+1} = -\left(\mathrm{B}\mathrm{M}^{-1}\mathrm{L}\right)^{-1}\mathrm{B}\mathrm{V}_{n+1} = \mathrm{H}^{AB^{-1}}\mathrm{V}_{n+1} \tag{20}$$

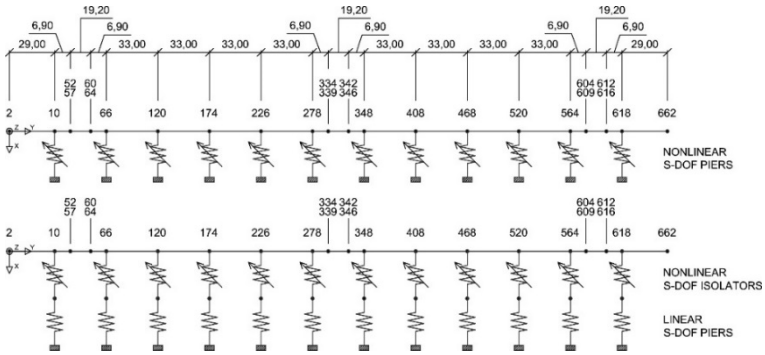
In this way, the relevant energy at the interface is preserved.

3 THE RIO TORTO BRIDGE CASE STUDY

The Rio Torto viaduct is a RC structure built during the sixties and located between Florence and Bologna in Italy. Twelve portal piers support the thirteen-span bay deck of each independent roadway as shown in Figure 3. End spans measure 29 m, whilst internal spans 33 m. Six Gerber saddles are present in the middle of the bridge and close to both abutments. Additional details can be found in [8, 9].

Fig. 3. Lateral view of the Rio Torto viaduct.

Solid and hollow circular cross section columns characterize short and tall piers, respectively. Relevant diameters measure 1200 mm and 1600 mm. The taller Pier #7 measured 41.34 m, whilst the shorter Pier #12 measured 13.80 m. Gerber saddles were removed in the isolated configuration where a pair of Concave Sliding Bearing (CSB) isolator were interposed between the deck and each pier. A comprehensive set of time history analyses of a refined FE model of the bridge was conducted in the OpenSees environment at both the SLS and ULS. With reference to the *as built* configuration, all piers showed hysteretic dissipation already at SLS. Conversely, time history analyses conducted in the *isolated* case proved that isolator pairs dissipated the most of hysteretic energy whilst piers remained in the linear range. A linear response of the deck was observed for both conditions. Moreover, numerical simulations highlighted that frame piers experienced in-plane deformations. In order to comply with the computational and control performance of the experimental facility for complex HSs, a reduced 88-DoFs substructured model of the viaduct was assembled for the *as built* and the *isolated* case [8, 9].In greater detail, the internal constraint setting was simplified and out-of-plane displacements of piers were fixed, whilst relative rotations among the deck and piers were released. Figure 4 depicts schematics of bridge configuration with node numbering, dimensions and frequencies of the four lowest eigenmodes. Nonlinear reduced S-DoF piers, provided transversal stiffness to the deck, whose cross sectional characteristics were $A = 4.63\ m^2, Ixx = 51.90\ m^4, Izz = 3.45\ m^4, Iyy = 0.10\ m^4$.



a

Mode	Freq. [Hz]
1	0.3674
2	0.3771
3	0.3922
4	0.4607

b

Mode	Freq. [Hz]
1	0.5010
2	0.6606
3	0.9392
4	1.3642

Fig. 4. Plan view of the reduced nonlinear models of the Rio Torto viaduct in: a) as built configuration; b) Isolated configuration. Dimensions in m;

As can be appreciated from Figure 4, Pier #9 and #11 were substructured in the laboratory together with related pairs of CSB isolators. The deck, remaining piers and related isolators were simulated numerically by means of reduced state space models. In greater detail, nonlinear S-DoF reduced models of piers were based on a modified version of the well-known Bouc-Wen mode. They were capable of reproducing the softening behavior predicted by the aforementioned refined OpenSees FE model. To this end, the factor $1/(1 + \alpha x^2)$ was added to modify the linear component of the tangent stiffness and a ρ parameter was introduced to simulate the degradation of the linear stiffness k. The resulting formulation reads,

$$\begin{cases} r + c \cdot \dot{x} + m \cdot \ddot{x} = -f \cdot \ddot{u}_g(t) + p(t) \\ \dot{r} = \left[\rho \cdot k / \left(1 + \alpha \cdot x^2\right) - (\beta \cdot sgn(\dot{x} \cdot r) + \gamma) \, | r |^n \right] \cdot \dot{x} \end{cases} \tag{21}$$

where β, γ and n are parameters of the Bouc-Wen model. In order to decrease the number of nonlinear parameters, γ and n were set to zero and to one, respectively [10]. Because, reduced nonlinear springs were not capable of reproducing piers' behavior at their full operating range, a set of parameters was identified for each different limit state. In order to force a linear response of piers without damage, β and ρ were set to zero, respectively, in the *isolated* case. Single Friction Pendulum Bearing OpenSees elements embed a physical model that is able to replicates the slip mechanism of CSB devices. Because bilinear shapes characterize inherent hysteretic loops, the state space model proposed by [11] was selected to reproduce the NS of isolator elements. Although the effect of variable vertical loads was neglected in the NSs, simplified bilinear hysteretic models well reproduced the response of all OpenSees isolators. Identified values of all nonlinear parameters of both reduced isolators and piers can be found in [8, 9].

4 NUMERICAL COMPARISON OF PROPOSED METHODS

In the following, the algorithmic performance of the aforementioned algorithm are compared in a pure numerical setting. From the HS perspective, the behavior of link solutions is crucial. In fact, their magnitude directly reflects on smoothness of actuators trajectories. Along this line, the following analyses try to shed light on both link quantities, which force velocity continuity at the interface between the NS and PS. Figure 5 depicts a schematic of the isolated bridge model with particular emphasis to the substructuring scheme.

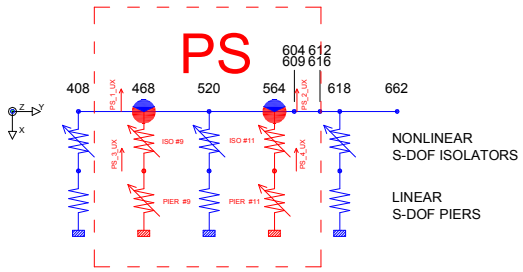


Fig. 5. Close up view on the substructuring scheme of the reduced nonlinear models of the Rio Torto viaduct in the Isolated configuration.

As can be appreciated from Figure 5, both interface Nodes #468 and #564 gather mass from both the deck and the related CSB isolator element. The last contributions is very small and entails high frequency eigenmodes on each 2-DoFs PS when detached from the overall bridge, i.e. in the decoupled configuration, hereinafter. In principle, resulting faster dynamics increase free displacements $\mathbf{u}^B_{free}$ on the physical side, and therefore, that link displacements $\mathbf{u}^B_{link}$ must compensate for within the subcycling time. As a consequence, actuators, which span free solutions, can potentially experience sharp trajectories that reduce the quality of the test. In order to mitigate such effect, part of the mass on the deck side -NS- is moved on the isolator side -PS-. Since both contributions share the same nodes, the overall dynamics remains unchanged and eigenfrequencies of the decoupled PS are pleasantly reduced. As a result, the PS benefits of smaller link displacements $\mathbf{u}^B_{link}$, and therefore, actuators experience smoother

trajectories. The mass fraction parameters m_f was introduced to modulate the mass transfer among NSs and PSs as follows,

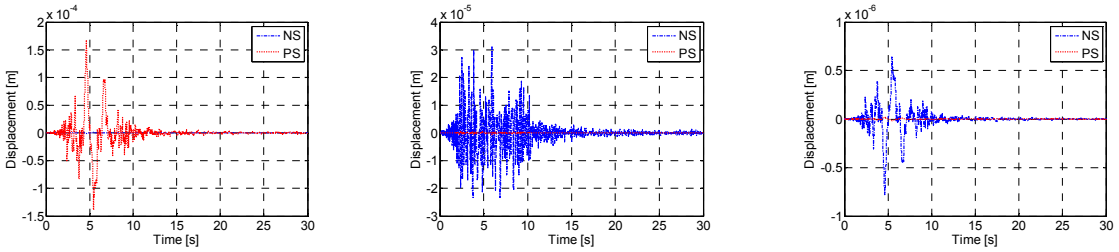
$$\mathbf{M}_i^{B'} = \mathbf{M}_i^B + m_f \cdot \mathbf{M}_i^A, \mathbf{M}_i^{A'} = \left(1 - m_f\right) \cdot \mathbf{M}_i^A \tag{22}$$

where the subscript i stands for interface; accordingly, mass matrices of (22) must be intended confined to interface DoFs. Therefore, $\mathbf{M}^{A'}$ and $\mathbf{M}^{B'}$ matrices characterize the interface DoFs masses of the PS and the NS, respectively, during HS. For the sake of clarity, Table 2 summarizes the eigenfrequencies of the decoupled 4-DoFs PS for different values of the mass fraction parameter.

Table 2 - Eigenfrequency of the decoupled PS

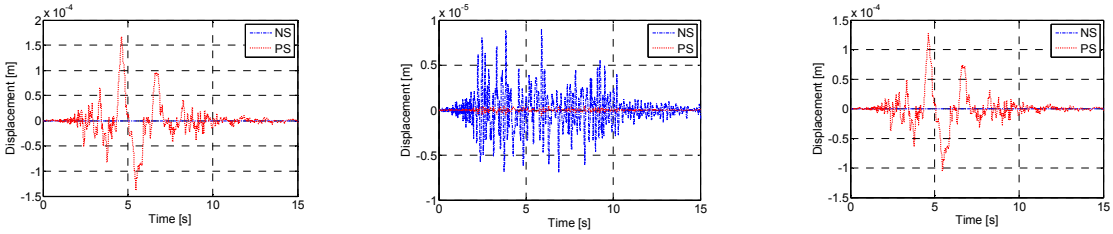
	Mass fraction m_f			
Mode	0.001	0.005	0.05	0.95
1	2.36	2.32	1.96	0.76
2	3.56	3.50	2.98	1.15
3	135.82	61.81	23.03	13.69
4	174.67	79.45	29.53	17.51

As can be appreciated from Table 2, a very small contribution of 0.005 already halves isolator eigenfrequencies. Clearly, such high frequency modes are not present in the PS of the as built configuration. Therefore, the isolated bridge is considered to compare algorithmic performance. In this preliminary study, coarse time step quantities are analyzed. Future investigations will shed light also on subcycling solutions. The same coarse time step was considered for both algorithms, i.e. $\Delta t^A = 1/1024s$. Moreover, a Newmark implicit scheme - $\gamma = 0.5$ and $\beta = 0.25$ - was applied to both subdomains. In order to estimate errors, a reference solution was calculated considering a monolithic implementation of the Newmark implicit scheme with a time integration step $\Delta t = \Delta t^A/100$. All reported results refer to the transversal displacement DoF of Node #468, which connects Pier #9 and related isolator to the deck. In detail, Figures 8 and 10 compares link displacement histories obtained from both the PM and the modified PH methods for different subcycling and mass fraction setting.



a) b) c)

Fig.6 - PM method: link displacement solutions in the case of: a) $\boldsymbol{ss = 1, m_f = 0.001}$; b) $\boldsymbol{ss = 1, m_f = 0.95}$; and c) $\boldsymbol{ss = 100, m_f = 0.001}$.



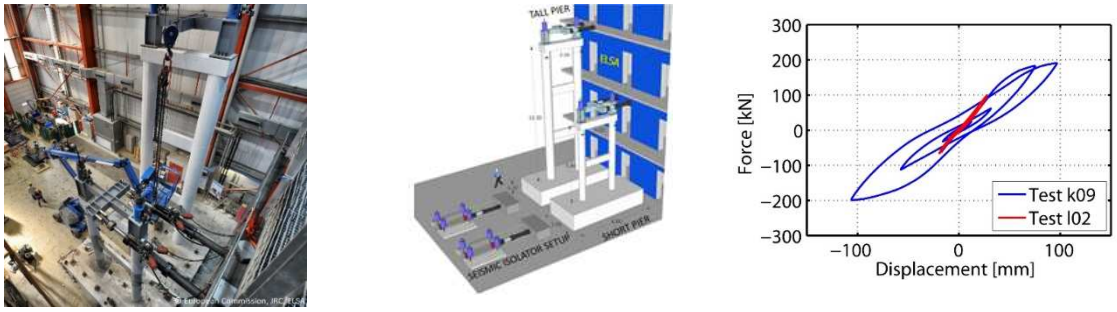
a) b) c)

Fig. 7 - Modified PH method: link displacement solutions in the case of: a) $\boldsymbol{ss = 1, m_f = 0.001}$; b) $\boldsymbol{ss = 1, m_f = 0.95}$; and c) $\boldsymbol{ss = 100, m_f = 0.001}$.

As can be appreciated from Figures 6a, 6b, 7a and 7b, given the same subcycling ss the mass fraction parameter affects $link$ displacements that prevail where the mass is smaller, and both the PM and the modified PH algorithms show the same behavior. Conversely, as can be observed in Figures 6c and 7c, only the PM method can benefit of subcycling for reducing $link$ displacements on the PS.

5 THE EXPERIMENTAL FRAMEWORK FOR HYBRID SIMULATION

A comprehensive set of HSs of the viaduct was conceived. The PM method was selected as parallel partitioned time integrator. Its inherent subcycling capabilities permit the synchronization of the two separated integration processes that involve both NSs and PSs with a coarse time step Δt_A and a fine Δt_B and, respectively. As a result, smooth actuator trajectories were achieved. In this particular case, an extended time scale $\lambda = 200$ was assumed together with a subcycling parameter ss equal to 250. As result, $\Delta t_A = 2.5$ ms was applied to NSs, whilst $\Delta t_B = 0.01$ ms was selected for PSs. Thus, actuator displacement commands were generated on the PS at the controller rate of 500 Hz. An additional computational driver ran the Cast3m FE model of the NS. 2.5 scale mock-up models of Piers #9 and #11 were experimentally substructured in the ELSA laboratory. Since gravity loads did not play an important role on piers, Procedure #2 proposed by [12] was selected for specimen scaling. Therefore, scale factors S and S^2 were applied to displacements and forces, respectively. Hence, both stress and strain quantities were preserved. Several HSs were carried out. Figure 8 depicts the experimental setup and the hysteretic response of Pier #9 during Test k09 and Test l02, which correspond to the as built and the isolated configuration at ULS, respectively.



a) b) c)

Fig. 8. a) 1:2.5 mock-up scale specimens of Piers #9 and #11; b) vertical loading system of frame piers. c) Comparison of hysteretic loops of HS k09 and l02, i.e. as-built and isolated configurations at ULS, respectively, for Pier #9

As can be appreciated from Figure 8c, HSs proved the effectiveness of the retrofitting scheme based on CSB isolators.

6 CONCLUSIONS

Mixed implicit/explicit multi-time partitioning methods are ripe for a profitable application to HS. From this perspective, the present paper presents a preliminary study on the performance of the PM and modified PH methods tailored to hybrid systems. A twelve-pier reinforced concrete was introduced as reference case study. Comparisons were done in term of link displacement quantities, which must entail smooth actuator trajectories. During HS umerical

simulations and experimental tests proved that both algorithms are suitable to HS. In this respect particular care must be devoted to the partitioning of mass between PS and NS. Future investigation will be addressed in the frequency domain, where spectral distributions of errors, kinematic quantities and Lagrange multipliers can be compared to eigenfrequencies of both coupled and decoupled substructures.

7 ACKNOWLEDGMENTS

The financial support from the SERIES (Seismic Engineering Research Infrastructures for European Synergies) Project, funded within the 7th Framework Programme of the European Commission [FP7/2007-2013] under grant agreement n° 227887, is greatly appreciated. Any opinions, findings and conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect those of SERIES sponsors.

REFERENCES

[1] Pegon P. Continuous PsD Testing With Substructuring. In O. S. Bursi and D. J. Wagg, editors, Modern Testing Techniques for Structural Systems, Dynamics and Control. CISM-Springer Wien NewYork, 2008.

[2] A. Gravouil, A. Combescure . Multi-time-step explicit – implicit method for non-linear structural dynamics, pp 199–225, 2001.

[3] P. Pegon and G. Magonette Continuous PSD testing with nonlinear substructuring: presentation of a stable parallel inter-field procedure. Technical Report 1.02.167, E.C., JRC, ELSA, Ispra, Italy, 2002.

[4] Brun, M., Batti, a., Combescure, a., & Gravouil, a. (2014). External coupling software based on macro- and micro-time scales for explicit/implicit multi-time-step co-computations in structural dynamics. Finite Elements in Analysis and Design, 86, 101–119. doi:10.1016/j.finel.2014.04.005.

[5] A. Prakash, K. D. Hjelmstad. A FETI-based multi-time-step coupling method for Newmark schemes in structural dynamics. International Journal for Numerical Methods in Engineering, 61(13), pp 2183–2204, 2004.

[6] Newmark, N. M. (1959) A method of computation for structural dynamics. Journal of Engineering Mechanics, ASCE, 85 (EM3) 67-94.

[7] A. Bonelli, O.S. Bursi, L. He, G. Magonette, P. Pegon. Convergence analysis of a parallel interfield method for heterogeneous simulations with dynamic substructuring. International Journal for Numerical Methods in Engineering, 75(7), pp 800–825, 2008.

[8] Abbiati G., Bursi O.S., Caperan P., Di Sarno L., Molina F.J., Paolacci F., Pegon P., 2014. Hybrid simulations of a multi-span RC viaduct with plain bars and sliding bearings. Earthquake Engineering and Structural Dynamics – under review

[9] Paolacci, F. et al. 3-D Assessment of the seismic vulnerability of an old RC viaduct with frame piers and study of the effectiveness of base isolation through PsD testing (RETRO), Eur Report 26471 EN , 2014. JRC., European Union.

[10] Abbiati G. Dynamic Substructuring of Complex Hybrid Systems Based on Time-Integration, Model Reduction and Model Identification Techniques", PhD Thesis, Department of Civil, Env. & Mechanical Engineering, University of Trento, 2014.

[11] Mostaghel, N. Analytical Description of Pinching, Degrading Hysteretic Systems. Journal of Engineering Mechanics, 1999, 2:125, 216-224.

[12] Kumar, S., Itoh, Y., Saizuka, K., and Usami, T. Pseudodynamic Testing of Scaled Models. Journal of Structural Engineering, 1997, 123(4), 524–526.

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis (Eds)

PREDOMINANT PERIODS OF MULTI-DEGREE-OF-FREEDOM-SYSTEM ANALYSIS AND DYNAMIC SOIL-STRUCTURE INTERACTION FOR BUILDING STRUCTURES

MIZUO INUKAI[*], TOSHIHIDE KASHIMA[†], TAIKI SAITO[††] AND TATSUYA AZUHATA[†]

[*] International Institute of Seismology and Earthquake Engineering(IISEE),
Building Research Institute (BRI)
Tsukuba, Ibaraki, 305-0802, Japan
e-mail: inkm@kenken.go.jp, Web page: http://iisee.kenken.go.jp

[†] IISEE, BRI
Tsukuba, Ibaraki, 305-0802, Japan
e-mail: kashima@kenken.go.jp, azuhata@kenken.go.jp

[††] Toyohashi University of Technology
Toyohashi, Aichi, 441-8580, Japan
Email: tsaito@ace.tut.ac.jp, Web page: http:// http://www.tut.ac.jp/

Key words: Predominant Period, MDOF, Soil-Structure Interaction, Dynamic, Building.

Abstract. Coupled problems of the multi-degree-of-freedom-system (MDOF) or Soil-Structure Interaction (SSI) are usually translated to a series of the single-degree-of-freedom (SDOF) equations. In this paper, the predominant periods of MDOF analysis are analyzed without SDOF and the dynamic amplification factors for a sample data of a model building are shown. The analysis method is the assumption that, as well as SDOF analysis, the MDOF analysis is applied to the predominant periods by inelastic analysis. That can give the dynamic amplification factors of the MDOF. At the results, it is cleared that the predominant period of the MDOF is close to be the period by the eigenvalues and in the short period or in the high level modes, the dynamic amplification factors are high which should not be neglected.

Moreover, soil-structure interaction with SDOF are also analyzed by elastic analysis for the model building. Some effects to the building structure model by the rocking on the ground are discussed in the sway-rocking models.

1 INTRODUCTION

In the single-degree-of-freedom (SDOF) analysis, the mass, the damping coefficient and the stiffness can give the predominant periods and the response values. In the multi-degree-of-freedom-system (MDOF) analysis, after the mass matrix [M] , the damping coefficient matrix [C] and the stiffness matrix [K] without any external forces can give the eigenvalues, the

modes and others, the MDOF model is usually translated to a series of the SDOF models and the response values are analyzed. In the MDOF analysis, the predominant periods of the MDOF system with external forces are not usually discussed.

Therefore, in this paper, the predominant periods of the MDOF of the shear system of 3 mass with external forces of the periodic motions are analyzed. This analysis can give the response values of each mass in any periods of external forces, the predominant periods of each mass, by elastic and inelastic analysis. The dynamic amplification matrix is proposed.

Moreover, soil-structure interaction (SSI) with the SDOF is also analyzed by elastic analysis for the model building. Some effects of SSI to the building structure model are discussed in the sway-rocking model. The external force vector including moment is proposed.

2 PREDOMINANT PERIOD OF MDOF ANALYSIS

The predominant periods of the MDOF analysis are analyzed and the dynamic amplification ratios are shown. In this analysis, as well as the periodic motion $\ddot{y}$ ($=p\sin\omega t$) in the external force of SDOF analysis in equation (1), the periodic motion in the external force of MDOF analysis is applied in equation (2).

$$m\ddot{x} + c\dot{x} + kx = -mp\sin\omega t \tag{1}$$

$$[M]\{\ddot{x}\} + [C]\{\dot{x}\} + [K]\{x\} = -[M]\{\ddot{y}\} \tag{2}$$

Where

$[M]$, $[C]$, $[K]$: n-by-n square matrix

$\{\ddot{x}\}, \{\dot{x}\}, \{x\}$: n-by-1 column matrix of response values of acceleration, velocity or displacement of MDOF

$\{\ddot{y}\}$: n-by-1 column matrix for external force ($\ddot{y}=p\sin\omega t$) (cm/sec^2)

p : constant value (cm/sec^2)

ω : angular frequency of the periodic motion (rad./sec)

The system of the response values is also applied to be the periodic motion in equation (3). Equation (3) is the particular solution of equation (2) by the method of undetermined coefficient which has the assumption that $\{x\}$ on i-th storey is linear combinations of shaking functions $<\{x_i\}= \{c_{1i}\sin\omega t + c_{2i}\cos\omega t\}>$ (equation (3)), in which $\{c_{1i}\}$ and $\{c_{2i}\}$ are the undetermined coefficients. Equation (3) is substituted for equation (2) to make eqaution (4). Equation (4) is an identity in any time and can be solved to equation (5), (6), (7) or (8) which give the solution $\{x_i\}$ for the reponse displacement of MDOF.

The each element of $\{x_i\}$in equation (6) includes the periodic motion in $\{\ddot{y}_{\phi i}\}$ calculated by the shaking functions with the angular frequency ω and the phase lag ϕ_i on i-th storey. The dynamic amplification matrix $[A]$, which doesn't include any shaking function, is calculated by $[M]$, $[C]$, $[K]$ and ω. The multiple of $[A]$ and $[M]$ can be amplitudes of the periodic motion $\{\ddot{y}_{\phi i}\}$. Therefore, equation (6) gives the predominant periods of the MDOF by the spectral analysis of ω.

$$\begin{Bmatrix} \vdots \\ x_i \\ \vdots \end{Bmatrix} = \begin{Bmatrix} \vdots \\ c_{1i}\sin\omega t + c_{2i}\cos\omega t \\ \vdots \end{Bmatrix} \tag{3}$$

$$\begin{aligned} &-\omega^2[M](\{c_{1i}\}\sin\omega t + \{c_{2i}\}\cos\omega t) + \omega[C](\{c_{1i}\}\cos\omega t - \{c_{2i}\}\sin\omega t) \\ &+[K](\{c_{1i}\}\sin\omega t + \{c_{2i}\}\cos\omega t) = -[M]\{p\}\sin\omega t \end{aligned} \tag{4}$$

$$\begin{cases} \{c_{1i}\} = [A_1][M]\{p\} \\ \{c_{2i}\} = [A_2][M]\{p\} \end{cases} \tag{5}$$

$$\{x_i\} = [A][M]\{\ddot{y}_{\phi i}\} \tag{6}$$

where

$[A]$: n-by-n square matrix. $[A]$ is a part of the particular solution and also the dynamic amplification matrix, which is calculated by $[M]$, $[C]$, $[K]$ and ω

$[A_1]$, $[A_2]$: n-by-n square matrix. A part of $[A]$

$\{\ddot{y}_{\phi i}\}$: n-by-1 column matrix for a periodic motion part of the particular solution with the phase lag ϕ_i on i-th storey, $\{\ddot{y}_{\phi i}=p\sin(\omega t+\phi_i)\}$ (cm/sec^2)

$$\begin{Bmatrix} \vdots \\ x_i \\ \vdots \end{Bmatrix} = \begin{Bmatrix} \vdots \\ \sqrt{c_{1i}^{\,2} + c_{2i}^{\,2}}\sin(\omega t + \phi_i) \\ \vdots \end{Bmatrix} \tag{7}$$

$$\begin{cases} \sin\phi_i = \dfrac{c_{2i}}{\sqrt{c_{1i}^{\,2} + c_{2i}^{\,2}}} \\ \cos\phi_i = \dfrac{c_{1i}}{\sqrt{c_{1i}^{\,2} + c_{2i}^{\,2}}} \end{cases} \tag{8}$$

In the analysis of SDOF, the complementary solution of equation (1), by the assumption that the external force is zero, can give the predominant period $\omega_{sgl} = \sqrt{k/m}$ and the particular solution by the method of undetermined coefficient of equation (1) can give the dynamic amplification ratio $\{\mu_S\}$ explicitly by the ratio of the dynamic amplification to the static one in equation (9). The spectral relationship between ω/ω_{sgl} and μ_S of equation (9) is well known as a curve line with one peak.

$$\{\mu_S\} = \frac{k}{\sqrt{(k - m\omega^2)^2 + (c\omega)^2}} \tag{9}$$

However, in the analysis of MDOF, the dynamic amplification matrix $[A]$ of equation (6) doesn't explicitly show the dynamic amplification ratio, likely equation (9) of SDOF. Moreover, the complementary solution of equation (2), can give the eigenvalues and the eigenvectors. These eigenvalues are considered to be the square of the angular frequency in the solutions and usually the high level modes are negleted in the analysis.

In order to know what the eigenvalues are in the analysis of MDOF, one example of a stucture data is as follows.

3 EXAMPLE FOR PREDOMINANT PERIOD OF MDOF ANALYSIS

3.1 Sant'Agostino in L'Aquila, Italy

The structure is a 2 storeys masonry sotructure heritage with a lantern storey. İt is assumed to be the shear system of 3 mass.

In the 2009 Italy L'Aquila Earthquake, Abruzzo, Italy, many monuments are damaged. The complex monument and church of Sant'Agostino in L'Aquila, Italy is one of these monuments and the very valuable chance was given to see the damage of it in L'Aquila. The typical damage of Sant'Agostino is the lantern fallen down to the roof of the next building in the West and severe damages in the ellipse dome or the walls. In this analysis, some structural identification was made by the shear system of 3 mass (Figure 1.) [1].

The inelastic dynamic response analysis needs the weight, the stiffness and the shear coefficient of each storey. The size of Sant'Agostino were measured based on the Google satellite map and plan or elevation figures in some references of Sant'Agostino. The Guideline for the Construction Technical Law [2], Italy is also referred for the material properties. The specific gravity is assumed to be 19 (kN/m^3) and the shear stiffness is assumed to be 23.0 (kN/cm^2). The period of the 1st storey is assumed to be 1.0 (sec) and the structural identification for the appropriate period and shear coefficient of each storey were made to follow the collapse of the lantern fallen down in the West. According to the drawings, the height is considered to be 41m. The weight of each storey is based on the wall areas of planning and the height of each storey to calculate the wall volume of each storey, decreasing of the openings, multiplying the specific gravity. The periods of each storey are based on the period of the 1st storey and the assumption of the linear relationship between the period T and the height of other storeys. The shear coefficients of each storey are decided by the assumption that the storey drift angles of the 1st and 2nd storey at the yielding point are nearly 1/1,000 (rad.) and the one of the 3rd storey (the lantern storey) is nearly 1/500. These storey drift angles, the periods or the stiffness give the shear coefficient.

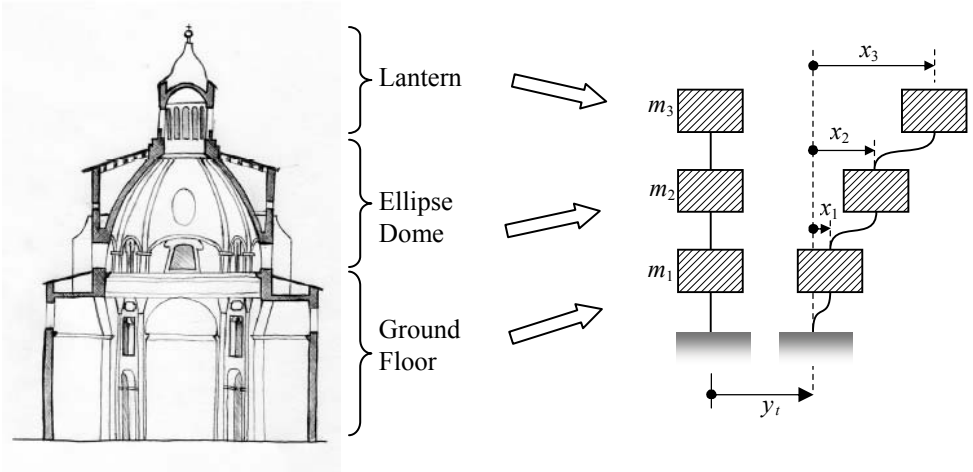


Figure 1: Cross section transverse in nearly EW direction and analysis model for shear system of 3 mass

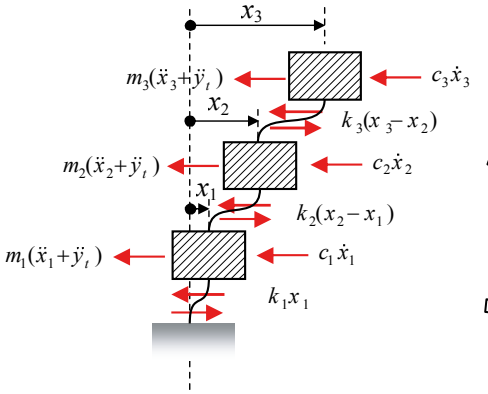


Table 1: Example values of stiffness and period of each storey for shear system of 3 mass

Storey	Stiffness k_i (kN/cm)	Period T_i (sec)
3	234	0.40
2	2,004	0.80
1	6,069	1.00

Figure 2: Analysis model for shear system of 3 mass

$$\begin{bmatrix} m_3 & 0 & 0 \\ 0 & m_2 & 0 \\ 0 & 0 & m_1 \end{bmatrix}\begin{Bmatrix} \ddot{x}_3 \\ \ddot{x}_2 \\ \ddot{x}_1 \end{Bmatrix}+\begin{bmatrix} c_3 & 0 & 0 \\ 0 & c_2 & 0 \\ 0 & 0 & c_1 \end{bmatrix}\begin{Bmatrix} \dot{x}_3 \\ \dot{x}_2 \\ \dot{x}_1 \end{Bmatrix}+\begin{bmatrix} k_3 & -k_3 & 0 \\ -k_3 & k_3+k_2 & -k_2 \\ 0 & -k_2 & k_2+k_1 \end{bmatrix}\begin{Bmatrix} x_3 \\ x_2 \\ x_1 \end{Bmatrix}=-\sin\omega t\begin{bmatrix} m_3 & 0 & 0 \\ 0 & m_2 & 0 \\ 0 & 0 & m_1 \end{bmatrix}\begin{Bmatrix} 1 \\ 1 \\ 1 \end{Bmatrix} \tag{10}$$

where

Period of i-th storey : $T_i = 2\pi\sqrt{\dfrac{\sum_{j=i}^{n} m_j}{k_i}}$ [sec]

Damping coefficient of i-th storey : $c_i = 2h\omega_i m_i = 2hm_i\sqrt{\dfrac{k_i}{\sum_{j=i}^{n} m_j}}$ $\left[\text{kN}\dfrac{\text{sec}}{\text{cm}}\right]$

The matrix equation of motion in forced shakings with external dampings is shown in equation (10). The external forces are the periodic motions.

Figure 2 shows the analysis model. Table 1 shows the stiffness k_i and the period *Ti* of each 3 storeys. According to the notes of equation (10), the period of i-th storey is calculated by the stiffness k_i of i-th storey and the summation of the mass upper than i-th storey. The damping coefficient of i-th storey is calculated by the damping factor h , the stiffness k_i and the mass m_i of i-th storey and the summation of the mass upper than i-th storey. The damping factor h is 0.05.

Equation (10) and (6) can give the dynamic amplification $\{x_{DA}\}$ and the static amplification $\{x_{SA}\}$ in equation (11-1) and equation (11-2). The ratio $\{\mu_D\}$ of the dynamic amplification to the static one in each storey is shown in equation (12).

Figure 3 shows the spectral analysis results for ω (ω=2π/T) of the relationship between the dynamic amplification factor μ_D and the period T (sec) for the spectral analysis. Figure 3 also shows the periods calculated by the eigenvalues which are given by equation (13). Equation (13) means the matrix equation of damped free motion with external damping.

$$\{x_{DA}\}=[A][M]\{1\}=\begin{Bmatrix} x_{DA3} \\ x_{DA2} \\ x_{DA1} \end{Bmatrix} \quad (11\text{-}1), \qquad \{x_{SA}\}=[K]^{-1}[M]\{1\}=\begin{Bmatrix} x_{SA3} \\ x_{SA2} \\ x_{SA1} \end{Bmatrix} \tag{11-2}$$

$$\{\mu_D\}=\begin{Bmatrix}\mu_{D3}\\ \mu_{D2}\\ \mu_{D1}\end{Bmatrix}=\begin{Bmatrix}x_{DA3}/x_{SA3}\\ x_{DA2}/x_{SA2}\\ x_{DA1}/x_{SA1}\end{Bmatrix} \tag{12}$$

$$\begin{bmatrix}m_3 & 0 & 0\\ 0 & m_2 & 0\\ 0 & 0 & m_1\end{bmatrix}\begin{Bmatrix}\ddot{x}_3\\ \ddot{x}_2\\ \ddot{x}_1\end{Bmatrix}+\begin{bmatrix}c_3 & 0 & 0\\ 0 & c_2 & 0\\ 0 & 0 & c_1\end{bmatrix}\begin{Bmatrix}\dot{x}_3\\ \dot{x}_2\\ \dot{x}_1\end{Bmatrix}+\begin{bmatrix}k_3 & -k_3 & 0\\ -k_3 & k_3+k_2 & -k_2\\ 0 & -k_2 & k_2+k_1\end{bmatrix}\begin{Bmatrix}x_3\\ x_2\\ x_1\end{Bmatrix}=-\begin{bmatrix}m_3 & 0 & 0\\ 0 & m_2 & 0\\ 0 & 0 & m_1\end{bmatrix}\begin{Bmatrix}0\\ 0\\ 0\end{Bmatrix} \tag{13}$$

According to Figure 3, the spectra of μ_D has the curves lines with usually 3 peaks on each storey. The periods calculated by the 3 eigenvalues are "0.39(s), 0.64(s) and 1.11(s)".

At the results, it is cleared that the predominant periods of MDOF are close to be the periods by the eigenvalues and in the short period or in the high level modes, the dynamic amplification ratios are at most 10. The dynamic amplification $\{x_{DA}\}$ at the predominant periods are usually different from the eigenvectors (Figure 4).

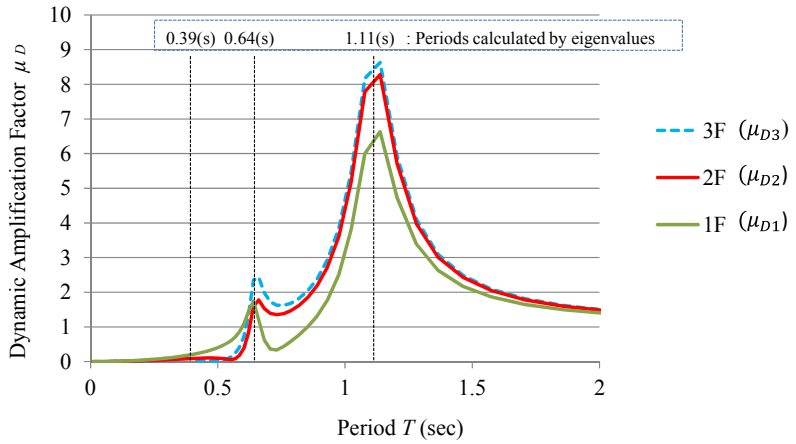
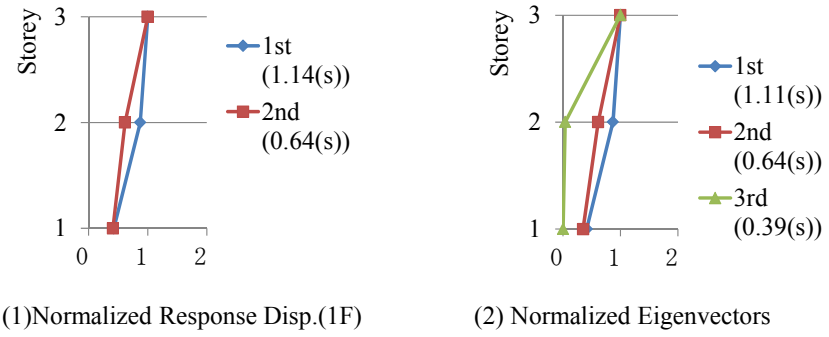


Figure 3: Spectra of dynamic amplification ratio μ_D
(when the stiffness is the initial stiffness of the hysteresis property.) (Ref. : Section 4.2)



(1)Normalized Response Disp.(1F) (2) Normalized Eigenvectors

Figure 4: Comparison of normalized response displacement (1F) and normalized eigenvectors by the value on the 3rd storey, at the initial stiffness of the hysteresis property. (Ref. : Section 4.2)

Almost eigenvalues are complex numbers ($a \pm b$i, where a and b are real numbers, i: is imaginary number). But during the stiffness degrading, some of eigenvalues become real numbers. When the eigenvalues are real numbers, the motion is called as overdamping motion.

The number of these eigenvalues is even, the issued frequency f (=1/T)(Hz) is zero and the eigenvector of the 1st eigenvalue is similar to the one of the 2nd eigenvalue. All elements of the eigenvectors are also similar.

A part of the inelastic analysis results of Table 3.1 (shear coefficients at the yielding points qCyi are 0.07, 0.09 and 0.11), Section 4.2 is as follows. When the stiffness is decreased to 1/2,200 of the initial stiffness, the number of the eigenvalues which are real numbers is 2. The static amplification $\{x_{SA}\}$ in equation (11-2) is nearly 2.5×10^4 (cm) and the dynamic amplification $\{x_{DA}\}$ in equation (11-1) in the period T =0.83 (sec), is from 19 ~ 15 (cm), when the eternal force vector is ${}^t\{1\ \ 1\ \ 1\}$ (transposed matrix) (g). Therefore the ratio $\{\mu_D\}$ in equation (12) is nearly 0.001 (Figure 5). The ratio $\{\mu_D\}$ is less than 1, because the stiffness on the 1st storey is decreased in the stiffness matrix [K] and the motion is an overdamping motion on the 1st storey.

Figure 6 shows the comparison of normalized response displacement (3F) and normalized eigenvectors by the value on 3rd storey, when the eigenvalue is a real number and the qk_3 is decreased. Figure 6.(1) shows the response displacement on the 1st storey is larger than the one on the upper storeys in the 2nd and 3rd mode. It seems to match the overdamping motion on the 1st storey and the higher level modes should not be neglected.

However, Figure 6.(2) seems to be similar to Figure 4.(2) at the the initial stiffness.

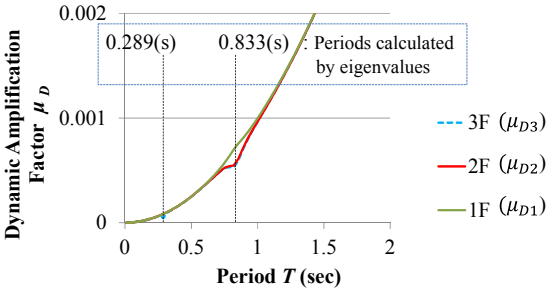
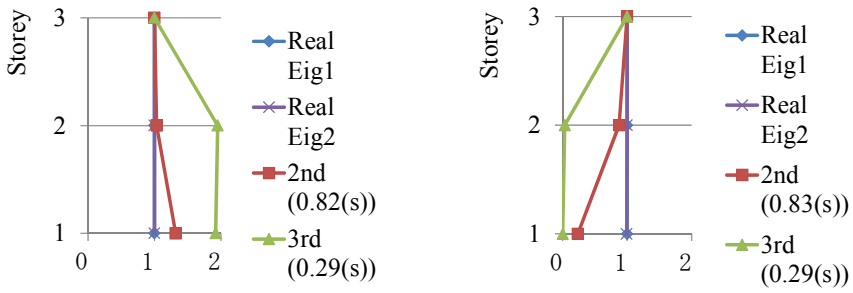


Figure 5: Spectra of dynamic amplification ratio μ_D
(when the eigenvalues are real numbers and qk_3 = (1/1000)(qCy/δy) (Ref. : Section 4.2)



(1)Normalized Response Disp.(3F) (2) Normalized Eigenvectors

Figure 6: Comparison of normalized response displacement (3F) and normalized eigenvectors by the value on the 3rd storey, when the eigenvalues are real numbers and qk_3 = (1/1000)(qCy/δy) (Ref. : Section 4.2)

4 RESPONSE VALUES OF MDOF BY ACCELERATION RECORDS

4.1 Acceleration records

The earthquake waves are shown in Table 2. Not only [2009 AQV EW, Itaca], but also [2009 L'Aquila Parking Entrance (AQK) EW, Itaca] and [2009 L'Aquila Castle (AQU) EW, Itaca] are used because the observation point of AQK is closer to Sant'Agostino and AQU is closest to it. In order to make the structural identification to clarify the physical phenomenon of the lantern fallen down to the West, the EW direction of these earthquake waves were used.

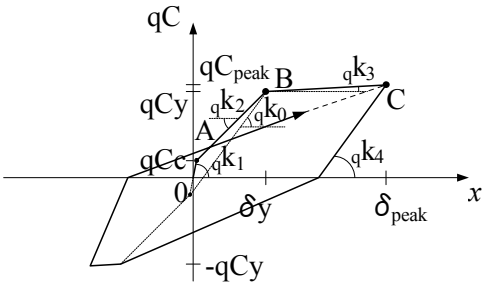
Table 2: Earthquake waves in EW direction of L'Aquila Main shock (Magnitude 5.8)

Date and Time in local time	Station Code	Address (L'Aquila, Abruzzo, Italy)	Peak Acceleration in EW (cm/sec^2)	Epicentral Distance (km)	Distance to Sant'Agostino in L'Aquila (km)
06 Apr. 2009 03:32:39	AQV	Center of Valley Aterno	662.6	4.9	5.5
	AQK	L'Aquila Parking Entrance	323.8	5.7	0.4
	AQU*)	L'Aquila Castle	258	6.0	1.0

Note) *) The acceleration records of AQU are corrected by subtracting the average of all data.

4.2 Inelastic dynamic analysis

Figure 7 shows the hysteresis property of Tri-linear Model for the inelastic dynamic analysis. In this analysis, Takeda Model is applied which has the oriented point before the yielding point is the crack point in the opposite. The stiffness ${}_qk_4$ of unloading after the



Takeda Model (The oriented point before the yielding point is the crack point in the opposite. The stiffness of unloading after the yielding point ; ${}_qk_4 = {}_qk_0 \left(\frac{\delta y}{\delta \text{ peak}} \right)^{0.4}$)

$${}_qk_1 = 2.2 \cdot \frac{qCy}{\delta y} \qquad {}_qk_2 = \frac{{}_qk_1}{3}$$

$${}_qk_3 = \frac{1}{1000} \frac{qCy}{\delta y}$$

$$qCc = 0.4 \cdot qCy, \quad \frac{qCy}{\delta y} = \frac{ky}{mg} = {}_qk_y$$

Notes)

(1) qC : Shear Coefficient

$$\left(qC = \frac{\text{Storey Shear Force (Q)}}{\text{Weight } (m \cdot g)} \right)$$

(2) qCy : Shear Coefficient at the yielding point of A (qCy = 0.07, 0.11, etc.)

(3) qC_{max} : Shear Coefficient at the peak of B

(4) Q : Storey Shear Force (kN) ($= k \cdot x$)

(5) ${}_qk_1$, ${}_qk_2$, ${}_qk_3$: Stiffness Coefficient (1/cm)

$$\left({}_qk_i = \frac{\text{Initial Stiffness } k_i}{\text{Weight } (m \cdot g)} \right)$$

(6) Stiffness Degrading Ratio : 1/1000

$$\left({}_qk_3 = \frac{1}{1000} \times {}_qk_y \right)$$

(7) h : Damping factor (= 0.05)

(8) g : Gravity Acceleration (=980 (cm/sec^2))

Figure 7: Hysteresis property of Tri-linear Model

yielding point has some relationship with the stiffness ${}_{q}k_{0}$ between the yielding point and the crack point in the opposite side (${}_{q}k_{4} = {}_{q}k_{0}(\delta y/\delta \text{peak})^{0.4}$).

The other symbols of qCc, ${}_{q}k_{1}$, ${}_{q}k_{2}$ and ${}_{q}k_{3}$ are calculated by equations in Figure 7. Equations of qCc=0.4・qCy and $({}_{q}k_{2})=({}_{q}k_{1})/3$ are referred by the reference [3].

Ty is calculated by the mass *m* and the secant stiffness *ky* in equation (14).

$$Ty = 2\pi\sqrt{\frac{m}{ky}} = \frac{2\pi}{\sqrt{{}_{q}k_{y}\cdot g}} = 2\pi\sqrt{\frac{\delta y}{qCy\cdot g}} \tag{14}$$

4.3 Response values of MDOF by inelastic dynamic analysis

In the structural identification, some cases of the period and the shear coefficient on each storey are calculated and the one of these results are shown in Table 3.1 ~ Table 3.3. Table 3.1 shows that all of 3 cases (AQV, AQK and AQU) identified the larger displacement x3 on the lantern storey to the West. Table 3.2 and Table 3.3 show the results if the lantern storey would be retrofitted, when the qCy_3 is increased from 0.11 to 0.15 or 0.16. This retrofitting method would have valuable effects which would reduce the response displacement x3 on the lantern storey less than that before retrofitting.

This Section shows the analysis results for the effects of the shear coefficient qCy_3.

Table 3.1: Structural properties and analysis results before retrofitting

Structure Properties of each Storey						ACC. Records (EW)	Analysis Results of each Storey		
Period (s)			Shear Coefficient (-)				Response Relative Displacement (cm)		
*Ty*1	*Ty*2	*Ty*3	qCy1	qCy2	qCy3		*x*1	*x*2	*x*3
1.00	0.80	0.40	0.07	0.09	0.11	AQV	-4.15	-4.15	+8.05
						AQK	-2.83	-4.79	-10.88
						AQU	1.84	2.60	+2.75

Notes)(1) *x*1, *x*2, *x*3 : the positive (+) means the displacement to the East and negative (-) to the West.
(2) Hysteresis property is the Tri-Takeda Model. Ref. : Figure 7.
(3) Period *Tyi* is calculated by the secant stiffness and the mass on i-th storey.
($Tyi=2\pi\sqrt{m_i/kyi}$)
(4) These Notes are also applied to the Table 3.2 and Table 3.3.

Table 3.2: Structural properties and analysis results if retrofitting (qCy_3 → 0.15)

Structure Properties of each Storey						ACC. Records (EW)	Analysis Results of each Story		
Period (s)			Shear Coefficient (-)				Response Relative Displacement (cm)		
*Ty*1	*Ty*2	*Ty*3	qCy1	qCy2	qCy3		*x*1	*x*2	*x*3
1.00	0.80	0.40	0.07	0.09	0.15	AQV	-4.17	-4.00	+14.23
						AQK	-2.89	-3.91	+5.89
						AQU	1.80	-2.56	+5.73

Table 3.3: Structural properties and analysis results if retrofitting (qCy_3 → 0.16)

Structure Properties of each Storey						ACC. Records (EW)	Analysis Results of each Story		
Period (s)			Shear Coefficient (-)				Response Relative Displacement (cm)		
*Ty*1	*Ty*2	*Ty*3	qCy1	qCy2	qCy3		*x*1	*x*2	*x*3
1.00	0.80	0.40	0.07	0.09	0.16	AQV	-4.11	-4.27	+1.41
						AQK	-2.84	-4.51	- 3.07
						AQU	1.79	-2.60	+0.81

5 PREDOMINANT PERIOD OF SOIL-STRUCTURE INTERACTION

5.1 Sway Rocking (SR) model

The predominant period of soil-structure interaction is also analyzed.

The sway rocking (SR) model is applied. Figure 8 shows overview of sway-rocking model for soil-structure interaction with SDOF.

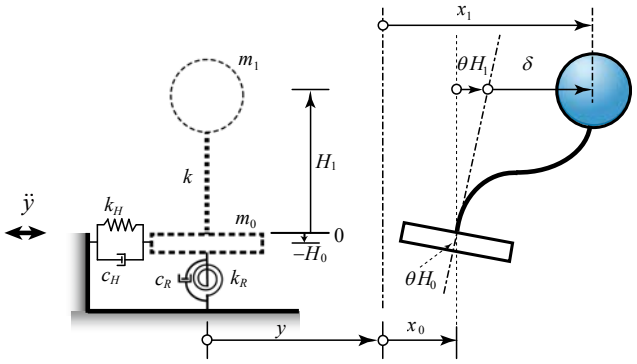


Figure 8: Overview of sway-rocking model for soil-structure ınteraction

Figure 8 shows SDOF, sway and rocking between the foundation and the ground. This SR model has a mass of m_0 and m_1 (kN/g) for mass of the structure and mass of the foundation. Stiffness of the structure and sway stiffness are k, k_H (kN/cm). Rocking stiffness is k_R (kN·cm/rad.). Damping coefficient of the structure is $c = 2h\sqrt{m_1 k}$ (kN·sec/cm), the one of the sway in soil-structure interaction (SSI) is $c_H = 2hm_0\sqrt{k_H/(m_1+m_0)}$ (kN·sec/cm) and the one of the rocking in SSI is $c_R = 2h\sqrt{I\,k_R}$ (kN·cm·sec/rad.). Equivalent height at the center of gravity of the structure and the foundation are H_1 and H_0 (cm). I $(=m_1H_1^2 + m_0H_0^2)$ (kN·cm^2/g /rad.) is the moment of inertia at the ground surface for the 2 mass of SDOF and the foundation. h is damping facor (h = 0.02). g is gravity acceleration (g = 980 (cm/sec^2)).

When the external force $\ddot{y}$ (cm/sec^2) in acceleration records takes place on the ground level, SR model has the horizontal displacement of the ground y (cm), the horizontal sway displacement x_0 (cm), the horizontal displacement of SDOF δ (cm) and the rotation angle of rocking θ (rad.). The relative horizontal displacement to the ground of SDOF x_1 (cm) is calculated by equation (15). The storey shear force of SDOF is $k\,\delta$ $(=k\,(x_1 - x_0 - \theta H_1))$ (kN).

$$x_1 = x_0 + \theta H_1 + \delta \tag{15}$$

$$\begin{cases} m_1\ddot{x}_1 + c\,\dot{x}_1 + k(x_1 - x_0 - \theta H_1) = -m_1\ddot{y} \\ m_0\ddot{x}_0 + c_H\dot{x}_0 - k(x_1 - x_0 - \theta H_1) + k_H(x_0 - \theta H_0) = -m_0\ddot{y} \\ I\ddot{\theta} + c_R\dot{\theta} + k_R\theta = -m_1\ddot{y}H_1 - m_0\ddot{y}\,(-H_0) \end{cases} \tag{16}$$

These displacements have their velocities and accelerations. The motion equation is equation (16). In equation (16), the left sides are internal forces or moments and the right sides are external forces or moments. The 3rd equation in equation (16) has the external moment $\{m_1H_1 + m_0(-H_0)\}\ddot{y}$ (kN·cm) to SR model by the ground acceleration $\ddot{y}$, which

means the equilibrium of moments for the 2 mass around the ground surface and means that the 3rd equation includes implicitly the inertia forces, the damping forces and the storey shear forces in the 1st and 2nd equation in equation (16). Equation (16) is also described by equation (17) in matrix which has forced shakings and external dampings.

$$\begin{bmatrix} m_1 & 0 & 0 \\ 0 & m_0 & 0 \\ 0 & 0 & I \end{bmatrix} \begin{Bmatrix} \ddot{x}_1 \\ \ddot{x}_0 \\ \ddot{\theta} \end{Bmatrix} + \begin{bmatrix} c & 0 & 0 \\ 0 & c_H & 0 \\ 0 & 0 & c_R \end{bmatrix} \begin{Bmatrix} \dot{x}_1 \\ \dot{x}_0 \\ \dot{\theta} \end{Bmatrix} + \begin{bmatrix} k & -k & -kH_1 \\ -k & k+k_H & kH_1 - k_H H_0 \\ 0 & 0 & k_R \end{bmatrix} \begin{Bmatrix} x_1 \\ x_0 \\ \theta \end{Bmatrix} = -\begin{bmatrix} m_1 & 0 & 0 \\ 0 & m_0 & 0 \\ 0 & 0 & I \end{bmatrix} \begin{Bmatrix} 1 \\ 1 \\ \{m_1 H_1 + m_0(-H_0)\}/I \end{Bmatrix} \ddot{y} \qquad (17)$$

The right side of equation (17) shows the unit external force vector ${}^t\{\ 1\ \ 1\ \ \{m_1 H_1 + m_0(-H_0)\}/I\ \}$ (transposed matrix). Similar unit external force vector in equation (10) shows that all elements are 1 (${}^t\{1\ \ 1\ \ 1\}$). But the 3rd equation of equation (17) is described for the response angle θ (rad.) or the moment by the horizontal force, the mass matrix [*M*] for the response acceleration vector is same as the one for the external force vector and the rightest value is $\ddot{y}$ (cm/sec^2). Therefore, this third element of the unit external force vector is "$\{m_1 H_1 + m_0(-H_0)\}/I$". This unit is "rad./cm" and 1 (rad./cm) is not appropriate for this third element. It is understandable when m_0 is 0 and H_0 is 0, the moment of inertia *I* would be described to be $m_1 H_1^2$ and this third element is calculated to be $m_1 H_1 / I = m_1 H_1 / m_1 H_1^2 = 1/H_1$ (rad./cm), which is a reciprocal of an arm length of moment.

5.2 Example of elastic analysis for SR model

The predominat periods are also analyzed for SR model by elastic analysis.

The model building is still Sant'Agostino in L'Aquila and it is modeled to SDOF of Figure 8. SR model needs many values in equation (17). All values are listed in Table 4.

The period T_1 of SDOF of Figure 8 is assumed to be 1.14(sec), according to Figure 4, (1). Therefore, the stiffness k of SDOF is 4.69x10^3 (kN/cm) ($k = (2\pi/T_1)^2 \cdot m_1$).

k_H and k_R are assumed to be 1.47x10^6 (kN/cm) and 2.77x10^{11} (kN·cm/rad.), according to the test results in Japan before 1962 [4].

The height is modified to the equivalent height at the center of gravity for SDOF (H_1=1,130(cm)). Figure 1 doesn't show the foundation, but it is assumed to have the foundation of 100 (cm) depth, the equivalent height (H_0=50(cm)), the specific gravity (=15 (kN/m^3)) and the mass (m_0=27.4 (kN/*g*)).

Moreover, the period of the sway T_H is 0.0697 (sec) ($T_H = 2\pi\sqrt{(m_1+m_0)/k_H}$).

The period of the rocking T_R is 0.167 (sec) ($T_R = 2\pi\sqrt{I/k_R}$).

Table 4: Data for SR model

m_1	m_0	I	k	k_H	k_R	c	c_H	c_R	H_1	H_0	h
154	27.4	1.95x10^8	4.69x10^3	1.47x10^6	2.77x10^{11}	34.0	98.8	2.94x10^8	1,130	50	0.02

Figure 9 shows that the predominat periods of SR model are almost same as the periods calculated by the eigenvalues. The reason is predicted probably that the stiffness are far from other stiffness ($k \ll k_H \ll k_R$). All curves lines of SDOF, Sway and Rocking have the peaks at their periods and the curve of Sway has another peak at the same period as SDOF. The new period 0.0271 (sec) is very close to the period of only Sway without SDOF (T_H' = $2\pi\sqrt{m_0/k_H}$ =0.0271 (sec)). The maximum of μ_D is nearly 25 in elastic analysis.

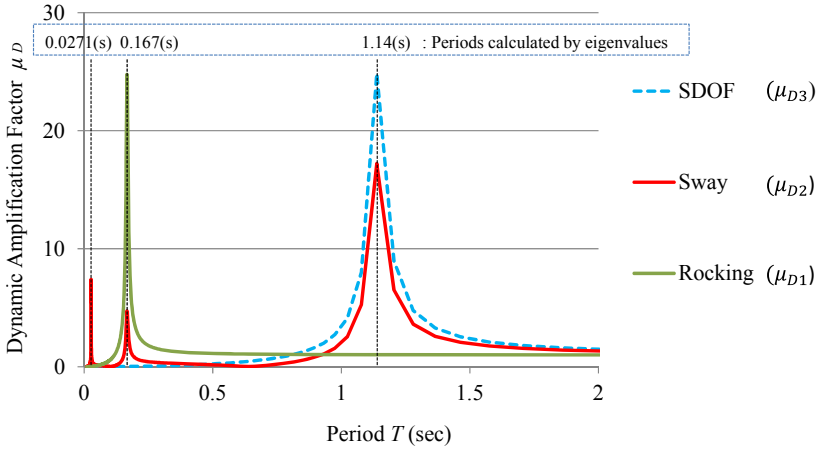


Figure 9: Spectra of dynamic amplification ratio μ_D in elastic dynamic analysis.

6 COUCLUSIONS

In this paper, the predominant periods of MDOF analysis are executed without SDOF and the dynamic amplification factors for a sample data of a model building are shown. The effects of soil-structure interaction are also analyzed in elastic analysis for the model building.

The main points of these results are as follows;

(1) According to the dynamic amplification ratio μ_D in the sample data of MDOF, the predominant periods of the MDOF for a model building are close to be the periods by the eigenvalues. The dynamic amplification ratios μ_D of MDOF are at most 10.

(2) The normalized response displacement of MDOF at the stiffness degrading shows that the higher level modes should not be neglected.

(3) In MDOF inelastic analysis, when some eigenvalues are real numbers, the motion seems to be overdamping.

(4) For the effects of soil-structure interaction (SSI) with SDOF, the elastic analysis of sway-rocking (SR) model, the predominant periods are close to the periods by the eigenvalues.

(5) The maximum of the dynamic amplification ratio μ_D of SR model in elastic analysis is nearly 25.

REFERENCES

[1] M. Inukai. Response Values of Earthquake Waves and Structural Assessment of Cultural Heritage, *Proceeding of The 15th World Conference on Earthquake Engineering* (2012), International Association for Earthquake Engineering, Paper ID. : 2012-2336.

[2] Application Guideline Draft on 7 March 2008 for Construction Technical Law by Ministral Directive on 14 Jan. 2008, *Higher Council of Public Works*, Italy (in Italian).

[3] M. Inukai, T. Kashima, T. Saito. Response Values in Hysteresis Properties by Acceleration Records, *Proceeding Usb of Second European Conference on Earthquake Engineering and Seismology*, Istanbul, Aug. 2014.

[4] I. Kaneko, K. Nakagawa, M. Nagatomo and T. Ohta. Vibrational Characteristics of Building Existing in Japan (Part 1), *Proceeding of Conference in Kanto Branch, Architectural Institute of Japan 1962*, pp.21-24 (in Japanese).

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis (Eds)

VALIDATION OF OPENFRESCO-BASED THERMOMECHANICAL HYBRID SIMULATION TO ADDRESS AN EARTHQUAKE-FIRE COUPLED PROBLEM – COUPLED PROBLEMS 2015

CATHERINE A. WHYTE[*], KEVIN R. MACKIE[†], GIUSEPPE ABBIATI[*] AND BOZIDAR STOJADINOVIC[*]

[*]Department of Civil, Environmental and Geomatic Engineering (D-BAUG)
Institute of Structural Engineering (IBK)
Swiss Federal Institute of Technology (ETH), Zürich
Stefano-Franscini-Platz 5, 8093 Zürich, Switzerland
e-mail: whyte@ibk.baug.ethz.ch, abbiati@ibk.baug.ethz.ch, stojadinovic@ibk.baug.ethz.ch

[†]Department of Civil, Environmental, and Construction Engineering
University of Central Florida
Engr II 402, 4000 Central Florida Blvd., Orlando, FL 32816-2450
email: kmackie@mail.ucf.edu

Key words: Thermomechanical, Hybrid Simulation, Earthquake, Fire.

Abstract. A new hybrid simulation platform for addressing thermomechanical coupled problems has been introduced in OpenFresco. This middleware resides between a numerical substructure (NS) and a physical substructure (PS) in a hybrid simulation and provides the means of communicating between the two. Whereas previously, this communication was restricted to mechanical loads, the new OpenFresco thermomechanical hybrid simulation (TMHS) capability additionally provides thermal degrees of freedom and temperature loads in the hybrid model. TMHS was implemented at the ETH Zürich IBK Structural Testing Laboratory. It provides a platform for addressing mechanical-fire coupled problems, and in particular, earthquake-fire problems.

The test presented herein demonstrates the capability of TMHS to simulate structural response to multi-hazard scenarios. The hybrid model consists of two elements. The NS is modeled in OpenSees. The PS is enclosed in a furnace placed in a universal testing machine. The hybrid model is first exposed to a ground motion excitation, applied mechanically by the universal testing machine, followed by a fire load specified by a fire cure and applied by the furnace. After completion of the fire loading and some cooling, a ground motion aftershock is applied to the hybrid model. The entire loading sequence is fully automated, so no user interaction is necessary except to open the doors of the furnace for the cooling phase. Demonstrating this successful investigation of the earthquake-fire coupled problem opens the possibilities for future investigations with more complex models and larger-scale tests.

1 INTRODUCTION

Knowledge about the behavior of structures in fire is severely limited by a lack of adequate testing of complete structural systems. Due to the need for expensive and highly specialized facilities to perform such tests, only a few full-scale tests [e.g. 1,2] or large-scale tests [e.g. 3,4,5] have been performed. Instead, the majority of our knowledge about structural performance in fire is the result of single-component tests [e.g. 6,7,8,9,10,11,12], exposed to standard fire curves. These fire curves do not represent real fires [13], and single-component test do not provide information about the interactions that occur between structural members in a full structure [14].

Structural fire design relies on prescriptive codes which are based on material behavior at elevated temperatures, observed in component tests: thus, the risk of structural system damage or collapse is not evaluated directly. There is great interest to adopt performance-based design approaches for fire [15]. However, high costs and access to sophisticated facilities will continue to impede extensive large-scale testing. The ability to develop performance-based fire design practices will rely on improved testing methods that assess global structural behavior, use of more realistic fire loads that include the post-fire cooling phase [16], and studies of the post-earthquake fire scenario, which is a common multi-hazard occurrence but is not addressed in the state-of-the-art design practices [17]. Hybrid simulation [18], which partitions a structure into numerical and physical substructures (NS and PS), and evaluates response of the combined hybrid model to some external excitation, was designed predominately to evaluate seismic loads. With its flexibility to accommodate different NS and PS configurations, ease of testing multi-hazard scenarios, including safely testing to the point of structural collapse, it is the ideal tool to improve our understanding of structural response in a variety of fire and combined earthquake-fire scenarios.

The thermomechanical hybrid simulation (TMHS) method has been implemented in the Open-source Framework for Experimental Setup and Control (OpenFresco) [19] hybrid simulation software framework, and verified and validated using a small-scale model in an electric furnace at the Swiss Federal Institute of Technology (ETH), Zürich [20].

A TMHS earthquake-fire-earthquake test of a structure exposed to an earthquake, followed by a fire curve, partial cooling, and then an earthquake aftershock is presented herein as a demonstration of the capabilities of this method for simulating multi-hazard scenarios.

2 BACKGROUND INFORMATION

The hybrid simulation testing method provides an opportunity to perform large-scale, coupled mechanical and fire tests, without the need for highly specialized laboratory facilities. It was originally developed [21,22,23] for determining the response of a structure to dynamic excitations, especially seismic. The portion of the prototype structure whose behavior is well-understood is modeled as the NS using standard finite elements, while the portion of the structure whose behavior is highly nonlinear or not well-understood is modeled as the PS using experimental elements. These experimental elements interface with physical specimens in the laboratory through laboratory test setups. The NS and PS interact throughout the simulation via computer control software and the specimen actuation system, which enforce consistent boundary conditions at the interface between the substructures, as the coupled numerical-physical system responds to an excitation.

The structural response of the hybrid model to thermomechanical loads over the time domain of interest is described by the following differential equations (Equation 1).

$$\begin{bmatrix} M_{11}^{NN} & 0 & \vdots & 0 & 0 \\ 0 & M_{11}^{PP} & \vdots & 0 & 0 \\ \cdots & \cdots & \cdot & \cdots & \cdots \\ 0 & 0 & \vdots & 0 & 0 \\ 0 & 0 & \vdots & 0 & 0 \end{bmatrix} \begin{Bmatrix} \ddot{u}^N \\ \ddot{u}^P \\ \cdots \\ \ddot{\theta}^N \\ \ddot{\theta}^P \end{Bmatrix} + \begin{bmatrix} C_{11}^{NN} & C_{11}^{NP} & \vdots & 0 & 0 \\ C_{11}^{PN} & C_{11}^{PP} & \vdots & 0 & 0 \\ \cdots & \cdots & \cdot & \cdots & \cdots \\ C_{21}^{NN} & C_{21}^{NP} & \vdots & C_{22}^{NN} & C_{22}^{NP} \\ C_{21}^{PN} & C_{21}^{PP} & \vdots & C_{22}^{PN} & C_{22}^{PP} \end{bmatrix} \begin{Bmatrix} \dot{u}^N \\ \dot{u}^P \\ \cdots \\ \dot{\theta}^N \\ \dot{\theta}^P \end{Bmatrix} + \begin{bmatrix} K_{11}^{NN} & K_{11}^{NP} & \vdots & K_{12}^{NN} & K_{12}^{NP} \\ K_{11}^{PN} & K_{11}^{PP} & \vdots & K_{12}^{PN} & K_{12}^{PP} \\ \cdots & \cdots & \cdot & \cdots & \cdots \\ K_{21}^{NN} & K_{21}^{NP} & \vdots & K_{22}^{NN} & K_{22}^{NP} \\ K_{21}^{PN} & K_{21}^{PP} & \vdots & K_{22}^{PN} & K_{22}^{PP} \end{bmatrix} \begin{Bmatrix} u^N \\ u^P \\ \cdots \\ \theta^N \\ \theta^P \end{Bmatrix} = \begin{bmatrix} F_u^N(t) \\ F_u^P(t) \\ \cdots \\ F_\theta^N(t) \\ F_\theta^P(t) \end{bmatrix} \quad (1)$$

N in the superscript refers to the NS and P in the superscript refers to the PS. The variables u, $\dot{u}$, and $\ddot{u}$ are the mechanical displacement, velocity, and acceleration, respectively. θ, and $\dot{\theta}$ are the temperature and rate of temperature change, respectively. M_{11} is the mass matrix. C_{11} is the mechanical damping matrix. C_{21} is the heat generation due to strain rate, which is important for micro-scale applications but is neglected for large-scale civil structural tests. C_{22} is the thermal capacity matrix. K_{11} is the mechanical stiffness matrix. K_{12} represents the internal forces due to restrained thermal deformations. K_{21} represents the thermal load changes due to mechanical displacements. K_{22} is the thermal conductivity matrix. $F_u(t)$ are the time-dependent mechanical external forces. $F_\theta(t)$ are the time-dependent thermal fluxes.

Fully coupled physics is defined as the situation when thermal behavior affects the mechanical behavior and vice versa. For most problems, full coupling only involves the K_{12} off-diagonal term. K_{21} is important in cases when the specimen undergoes large deformations in a compartment with a non-uniform temperature, such that the mechanical behavior of the specimen changes its thermal exposure (particularly when the structural deformations result in new ventilation, e.g. when a window breaks or integrity of a fire wall is lost).

The substructures are mechanically fully coupled when the displacements are sent from the NS to the PS and interface forces are measured and returned from the PS back to the NS (displacement control). The feedback restoring forces are used to compute the displacements in the next time step, so the displacement sequences applied to the PS are not known a-priori. The substructures are fully coupled thermally when temperatures are sent from the NS to the PS and thermal fluxes are measured and returned from the PS back to the NS (temperature control). The substructures are fully coupled thermomechanically when they are fully coupled mechanically and thermally. The substructures are partially coupled thermomechanically when the substructures are fully coupled mechanically or thermally but not both.

The first application of hybrid simulation to investigate structural response in fire was performed by Mostafaei [24,25]. A standard fire test of a single column PS was combined with a NS of the remainder of the structure to enable examination of the behavior of the whole structure exposed to both mechanical and fire loads. However, this implementation involved human interaction for data transfer between the numerical and physical substructures. TMHS advances the state-of-the-art by providing computer controlled interaction between the NS and PS. This automated infrastructure is critical for working with complicated models, eliminating the possibility of human error in data transfer, maintaining synchronization of the mechanical and thermal loading patterns in the event of a hardware delay in the laboratory, and enabling real-time testing capabilities.

3 TMHS EARTHQUAKE-FIRE-EARTHQUAKE SIMULATION

3.1 TMHS Framework

TMHS uses the Open System for Earthquake Engineering Simulation [26] software framework for modeling the NS. The OpenFresco hybrid simulation software framework provides the communication link between the NS and the PS, including the degree of freedom (DOF) transformations between the NS and the test setup in the laboratory and the interface with the test setup control system. Because OpenSees and OpenFresco were both originally developed for purely mechanical problems, the frameworks were extended to include the temperature DOFs at the nodes of the hybrid model [20]. The DOF transformation and experimental control for the thermal problem are both new additions to OpenFresco.

The earthquake-fire-earthquake TMHS presented here was done using a hybrid model with a partial coupling of the physics and a partial coupling of the substructures. Because OpenSees solves purely mechanical problems, the NS is not exposed to thermal loads. Developments for thermal capabilities in OpenSees are underway by researchers at Edinburgh [27,28]. Displacement control is used, meaning that the hydraulic actuator applies displacements to the PS and measures restoring forces. The temperatures are controlled on the PS, but the thermal fluxes are not measured and returned back to the NS. This constitutes a partial coupling of the substructures.

In the earthquake-fire-earthquake test, the equation of motion is solved in each time step using a hybrid simulation specific (but mechanical-only) Newmark Implicit integrator with an increment reduction factor of 0.4 and a fixed number of 4 iterations per time step [29]. The fixed number of iterations per time step is important for real-time hybrid simulation such that each simulation time step is applied in the same amount of clock time. The number of iterations necessary to achieve the required tolerance in each simulation time step is preselected based on the performance of the hybrid test in full-simulation (PS modeled as another NS on a separate processor) mode.

3.2 Experimental Design

The earthquake-fire-earthquake test is conducted on a simple single DOF hybrid model. A single beam element NS is implemented in OpenSees (A) and a single truss element PS is modeled physically (B), as shown in Figure 1a. The fundamental vibration period of the hybrid model was set at 1 s and instantiated by calculating the necessary mass and assigning this mass entirely to the NS. A 5% mass proportional Rayleigh damping is used. The PS is a Grade S355 structural steel RHS 120-60-3.6 dogbone-shaped specimen (the narrow portion of the dogbone shape is 75 mm long, 10 mm wide, and 3.6 mm thick), tested in a combined Zwick 1484 Universal Testing Machine (UTM), which applies axial displacements and measures reaction forces, and a Könn STE-12 HR/350’ (0.1 m radius x 0.5 m tall) electrical furnace, which encapsulates the specimen and controls the temperature at specimen mid-height (Figure 1b). Though the TMHS framework is designed to enable real-time testing, the low heating power of the Könn furnace cannot achieve real-time temperature loadings. An extensometer measures the strain of the physical specimen directly. The length, L, in Figure 1a, is 40 mm to match the gage length of the extensometer probes. Three thermocouples measure temperature on the surface of the specimen. An Indel CPU board stand-alone master

(SAM) is used as the digital signal processor (DSP), responsible for generating displacement and temperature command signals and receiving feedbacks from both the UTM and furnace.

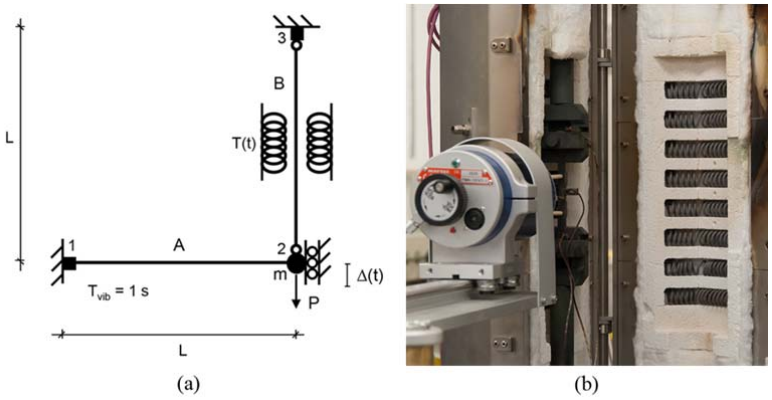


Figure 1: (a) Hybrid model; (b) Physical specimen in the furnace

In this earthquake-fire-earthquake test, PS response was maintained in the elastic range because this test was part of a series of proof tests and the specimen was reused many times. Furthermore, the UTM is configured for tensile loads only, so compression of the PS had to be prevented. Performing the test in displacement control in this scenario requires careful design and coordination of the mechanical and thermal load patterns. As the specimen heats in a displacement controlled environment, its expansion has the effect of reducing the restoring force. The mechanical load pattern must counteract this behavior to maintain some tension on the specimen at all times, but not too much tension such that the specimen yields.

With these requirements in mind, the earthquake-fire-earthquake test was performed in the following way. Initially, a small force ramp of 5 kN was applied to the single DOF in the hybrid model over 30 simulation seconds and 120 steps. Then the 1940 El Centro NS ground motion was applied using a very small scaling factor (0.0005) to avoid yielding the PS. This ground motion record has 1600, 0.02 s long steps: thus, it took 32 simulation seconds. Next, another small force ramp of 5 kN was applied to the single DOF in the hybrid model. Simultaneously, the fire curve was commanded to the experimental element in the hybrid model (and thus sent to the PS). This was performed over 30 simulation seconds and 120 steps. The fire curve was the international standard ISO 834 temperature-time curve, scaled to reach 200 deg C at 30 s, and defined in the Eurocode 1 Part 1-2 [30] as:

$$\Theta_g = 20 + 345\log_{10}(8t + 1) \tag{2}$$

where Θ_g is temperature (˚C) in the fire compartment and t is time (min). After the peak of the fire curve was attained, the furnace was turned off and its doors were opened, with the intent to apply linear time-temperature cooling pattern over 90 simulation seconds. A negative 5 kN force ramp was applied simultaneously to unload the specimen and avoid yielding it as it cools and contracts. Finally, the 1940 El Centro NS ground motion scaled with the same scale factor of 0.0005 was applied to the hybrid model with the furnace door open to represent an aftershock striking a partly cooled structure.

3.3 Results

The displacement command history and temperature command history for the earthquake-fire-earthquake test are shown in Figure 2.

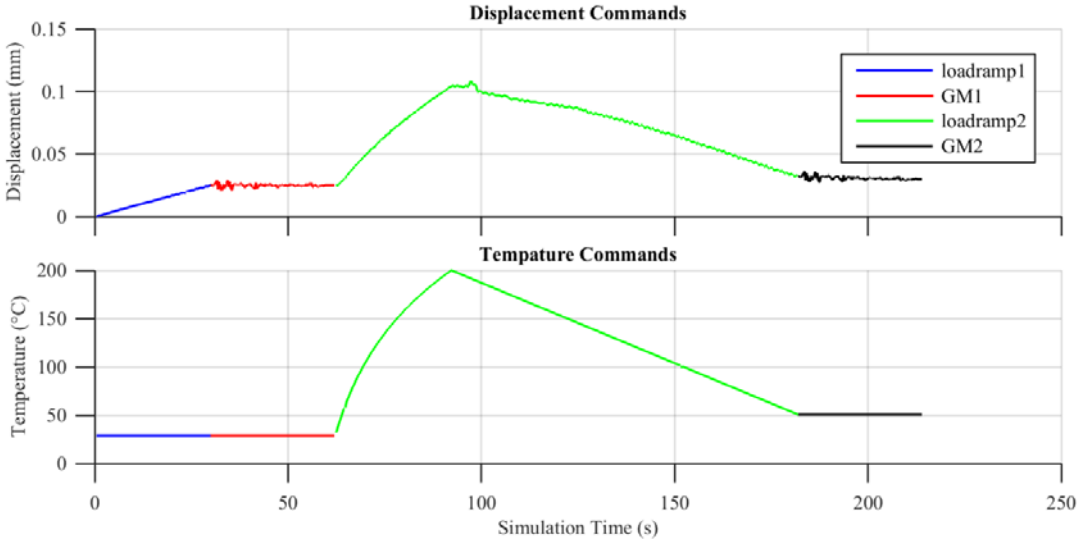


Figure 2: Displacement and temperature commands

The force feedback history and temperature feedback history are shown in Figure 3. The control of the temperature at the start of the cooling phase has a small error. The cooling phase should follow a linear decreasing time-temperature ramp. However, when the furnace doors were opened, the temperature dropped quickly. OpenFresco temperature error tolerance was set at 5˚C, but the temperature dropped too fast below the tolerance range. By temporarily allowing a higher temperature error tolerance, the discrepancy was resolved, and the remainder of the cooling ramp was performed as planned.

The displacement control errors and temperature control errors are shown in Figure 4. The displacement control errors are all on the order of 10^{-3} mm. The temperature control errors are generally within +/-5˚C, except for the error when the furnace doors were opened. At the end of the test, the temperature continued to drop as the aftershock earthquake was applied. This is a realistic scenario for continued cooling while an earthquake is applied to a partly cooled structure.

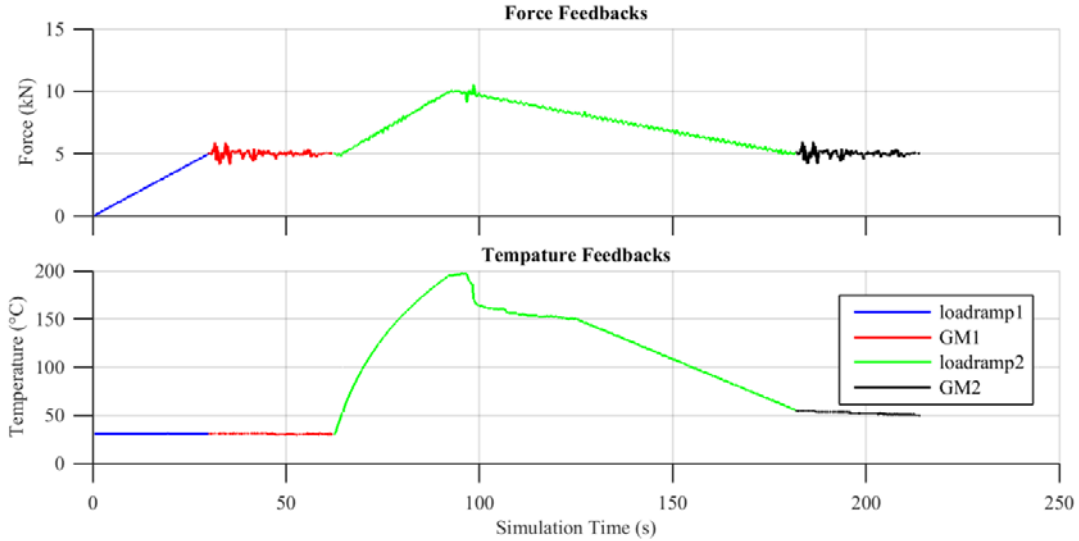


Figure 3: Force and temperature feedbacks

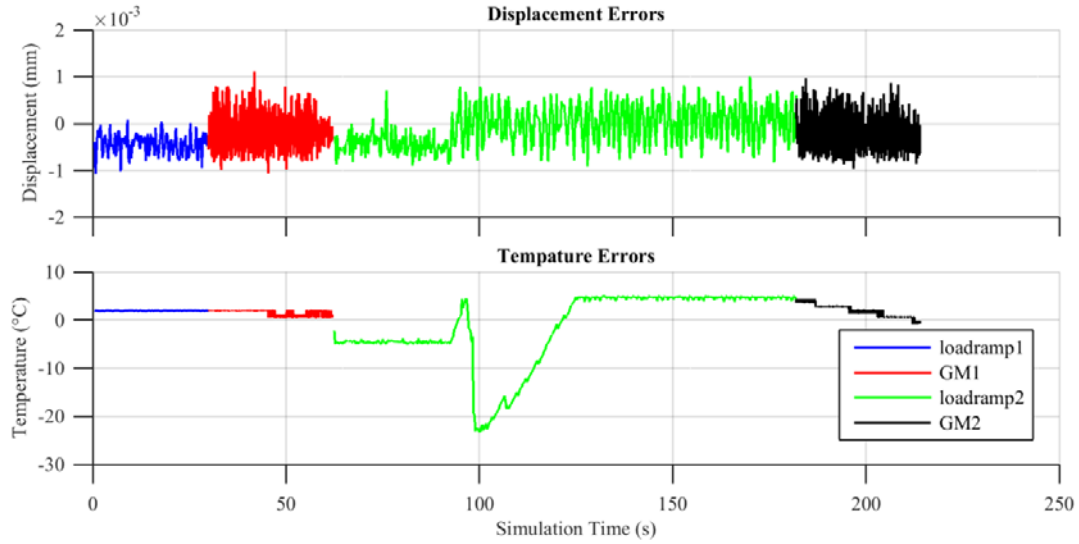


Figure 4: Displacement and temperature control errors

The force-displacement hysteresis for the single DOF of the hybrid model is shown in Figure 5. The forces are the sum of the restoring forces from the NS and PS. The specimen remains linear for the entirety of the test. This is confirmed by the same slope for the first load ramp (load ramp 1), the initial earthquake (GM1), and the final earthquake (GM2). The green line (load ramp 2) plots both the combined effects of the increasing load ramp (+5 kN) and the thermal load (fire curve), followed by the simultaneous decreasing load ramp (-5 kN) and cooling phase. The lower slope of this green line is the result of the specimen resistance dropping as it heats and expands (and the reverse during unloading and cooling).

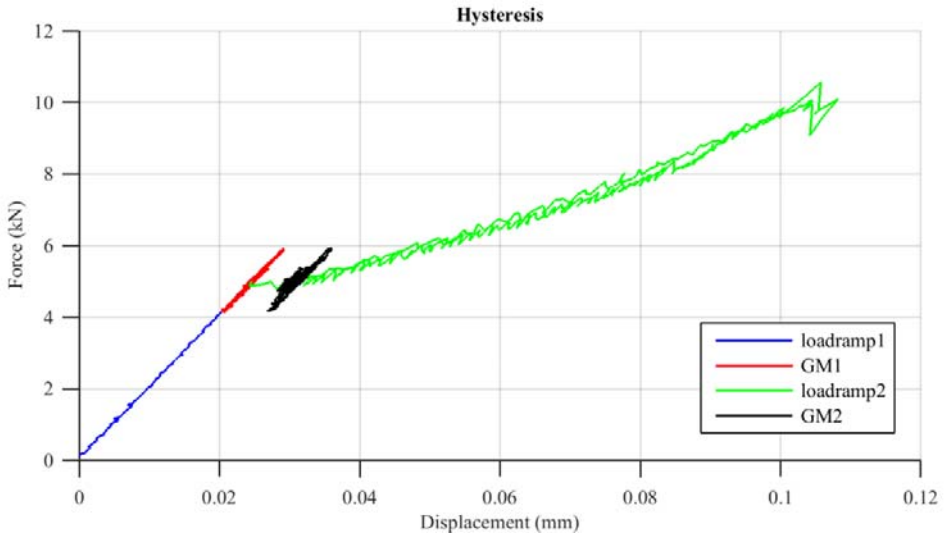


Figure 5: Hybrid model force-displacement response

4 CONCLUSIONS

The lack of knowledge about the behavior of entire structures in fire scenarios is a result of a lack of adequate testing. Very few full-scale structural fire tests have been performed due to the need for highly specialized facilities and high costs. Instead, tests are largely performed on single structural components, exposed to standard fire curves. Fire ratings are the outcome of such tests, which feed prescriptive fire design codes.

As in seismic design, fire design is moving towards performance-based standards. By designing for specific performance levels in hazard scenarios, the overall behavior of the structure will be considered from the outset of the design process. The risk to structures in fire scenarios will be known and quantifiable.

Development of performance-based fire design hinges on the development of new testing methods that can assess the global behavior of structures in fire. The thermomechanical hybrid simulation framework addresses these needs. A TMHS proof test, presented in this paper, demonstrated the capabilities for testing multi-hazard scenarios. The response of a hybrid model to an earthquake, followed by a fire, a partial cooling phase, and an aftershock earthquake was successfully simulated.

There are many opportunities for future work. The NS portion of the hybrid model used in the proof test was “cold”. By interfacing OpenFresco with finite element software that is specifically developed for thermomechanical hybrid simulation, such as SAFIR [31], more powerful numerical models can be instantiated and full coupling of the physics can be enforced. By measuring thermal flux feedback from the PS, full coupling of the substructures can be implemented. The low heater power of the Könn furnace and its poor controller were major impediments to the TMHS tests. With a more powerful furnace or other powerful heating elements, real-time TMHS is possible. The developments in OpenFresco presented herein support such real-time simulations.

11 ACKNOWLEDGMENTS

The authors are grateful to Mr. Dominik Werne and Mr. Martin Neuenschwander for their help with the experimental testing equipment. Funding for this project was provided by ETH Zürich. Any opinions, findings, and conclusions expressed herein are those of the authors and do not necessarily reflect the view of ETH Zürich.

REFERENCES

[1] British Steel. (1999). "The behavior of multi-storey steel frame buildings in fire." Rotherham: British Steel. 82 pages.

[2] Chen, M., Pantoli, E., Wang, X., Espino, E., Mintz, S., Conte, J., Hutchinson, T., Marin, C., Meacham, B., Restrepo, J., Walsh, K., Englekirk, R., Faghihi, M., and Hoehler, M. (2012). "Design and construction of a full-scale 5-story base isolated building outfitted with nonstructural components for earthquake testing at the UCSD-NEES facility." *Proceedings of the ASCE Structures Congress*. Chicago, Illinois.

[3] Chen, J., Young, B., and Uy, B. (2006). "Behavior of High Strength Structural Steel at Elevated Temperatures", ASCE Journal of Structural Engineering, 132(12), 1948-1954.

[4] Kirby, B.R. (1997). "Large scale fire tests: the British Steel European collaborative research programme on the BRE 8-storey frame." *Proceedings of the Fifth International Symposium on Fire Safety Science*, Melbourne, Australia.

[5] Kitano, T., et al. (2000). "Large scale fire tests of a 4-story type car park, Part 1: The behavior of structural frame exposed to the fire at the deepest part of the first floor." *Proceedings of the Fourth Asia-Oceania Symposium on Fire Science and Technology*, Tokyo, Japan.

[6] Dwaikat, M.M.S., Kodur, V.K.R., Quiel, S.E., and Garlock, M.E.M. (2010). "Experimental behavior of steel beam-columns subjected to fire-induced thermal gradients." *Journal of Construction Steel Research*, 67, 30-38.

[7] Franssen, J.M., et al. (1994). "A simple model for the fire resistance of axially-loaded members according to Eurocode 3." *Journal of Construction Steel Research*, 35, 49-69.

[8] Garner, L. and Baddoo, N.R. (2006). "Fire testing and design of stainless steel structures." *Journal of Construction Steel Research*, 62, 532-543.

[9] Lie, T.T. (1989). "Fire resistance of reinforced concrete columns: a parametric study." *Journal of Fire Protection Engineering*, 1(4), 121-130.

[10] Martins, A.M.B. and Rodrigues, J.P.C. (2010). "Fire resistance of reinforced concrete columns with elastically restrained thermal elongation." *Engineering Structures*, 32, 3330-3337.

[11] Tondini N., Hoang V.L., Demonceau J.-F. and Franssen J.-M. (2013). "Experimental and numerical investigation of high-strength steel circular columns subjected to fire loading." Journal of Constructional Steel Research, 80: 57-81.

[12] Tondini N., Rossi B. and Franssen J.-M. (2013). "Experimental investigation on ferritic stainless steel columns in fire." Fire Safety Journal, 62: Part C, 238-248.

[13] Almand, K. H. (2012). "Structural fire resistance experimental research." SpringerBriefs in Fire, Fire Protection Research Foundation, DOI: 10.1007/978-1-4614-8112-6_5.

[14] Wang, Y.C. (2002). "Steel and composite structures: Behaviour and design for fire

safety." Spon Press, New York, NY.
[15] Hamilton, S. R. (2011). "Performance-based fire engineering for steel framed structures: a probabilistic methodology," *Dissertation*, Stanford University.
[16] Dimia, M. S., Guenfoud, M., Gernay, T., and Franssen, J. M. "Collapse of concrete columns during and after the cooling phase of a fire." *Journal of Fire Protection Engineering*, 21(4), 245-263.
[17] Usmami, A. S. (2008). "Research priorities for maintaining structural fire resistance after seismic damage." Proceedings of the 14th World Conference on Earthquake Engineering, Beijing, China, October 12-17.
[18] Saouma, V.E., and Sivaselvan, M.V. (2008). "Hybrid simulation: Theory, implementation and applications." Taylor and Francis, London, UK.
[19] OpenFresco (2014). "Open Framework for Experimental Setup and Control." <http://openfresco.neesforge.nees.org> (Sept. 9, 2014).
[20] Whyte, C. A., Mackie, K.R. and Stojadinovic, B. (2014). "Hybrid Simulation of Thermomechanical Structural Response." *ASCE Journal of Structural Engineering* (in review).
[21] Takanashi, K., Udagawa, K., Seki, M., Okada, T. and Tanaka, H. (1975), "Non-linear Earthquake Response Analysis of Structures by a Computer-Actuator On-line System", Bull. Of Earthquake Resistant Structure Research Center, Institute of Industrial Science, University of Tokyo, No. 8, 1975.
[22] Mahin, S.A. and Williams, M.E. (1981), "Computer Controlled Seismic Performance Testing", *Second ASCE-EMD Specialty Conference on Dynamic Response of Structures*, Atlanta, GA.
[23] McClamroch, N.H., Serakos, J. and Hanson, R.D. (1981), "Design and Analysis of the Pseudo-Dynamic Test Method", Technical report UMEE 81R3, University of Michigan, Ann Arbor.
[24] Mostafaei, H. (2013a). Hybrid fire testing for assessing performance of structures in fire -Application. Fire Safety Journal, 56, 30–38. doi:10.1016/j.firesaf.2012.12.003.
[25] Mostafaei, H. (2013b). Hybrid fire testing for assessing performance of structures in fire -Methodology. Fire Safety Journal, 58, 170–179. doi:10.1016/j.firesaf.2013.02.005.
[26] OpenSees (2014). "Open System for Earthquake Engineering Simulation." <http://opensees.berkeley.edu> (Sept. 9, 2014).
[27] Jiang, J., Jiang, L., Kotsovinos, P., Zhang, J., Usmani, A., McKenna, F., and Li, G. (2014). "OpenSees software architecture for the analysis of structures in fire." Journal of Computing in Civil Engineering, 10.1061/(ASCE)CP.1943-5487.0000305, 04014030.
[28] Jiang, J. and Usmani, A. (2013). "Modeling of steel frame structures in fire using OpenSees." *Computers and Structures*, 118, 90-99.
[29] Schellenberg, A. (2008) "Advanced Implementation of Hybrid Simulation", *Ph.D. Thesis*, University of California, Berkeley.
[30] European Committee for Standardization (CEN). (2002). "EN1991-1-2, Eurocode 1: Actions on structures – Part 1-2: General actions – Actions on structures exposed to fire," Brussels.
[31] Franssen, J. M. (2005). SAFIR: A thermal/structural program for modeling structures under fire. Engineering Journal-American Institute of Steel Construction Inc, 42(3).

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis(Eds)

NUMERICAL SIMULATION OF GEOCHEMICAL COMPACTION WITH DISCONTINUOUS REACTIONS

Abramo Agosti*, Luca Formaggia*, Bianca Giovanardi* and Anna Scotti*

*MOX, Dipartimento di Matematica "F. Brioschi"
Politecnico di Milano
Via Bonardi 9, 20133 Milano, Italy
e-mail: mox@mate.polimi.it, web page: http://mox.polimi.it/

Key words: Compaction processes, discontinuous RHS ODE, mineral dissolution/precipitation.

Abstract. The present work deals with the numerical simulation of porous media subject to the coupled effects of mechanical compaction and reactive flows that can significantly alter the porosity due to dissolution, precipitation or transformation of the solid matrix. These chemical processes can be effectively modelled as ODEs with discontinuous right hand side, where the discontinuity depends on time and on the solution itself. Filippov theory can be applied to prove existence and to determine the solution behaviour at the discontinuities. From the numerical point of view, tailored numerical schemes are needed to guarantee positivity, mass conservation and accuracy. In particular, we rely on an event-driven approach such that, if the trajectory crosses a discontinuity, the transition point is localized exactly and integration is restarted accordingly.

1 Introduction

We propose a model to describe and simulate the compaction process under a progressive burial of a layer of sediments in which a mineral can dissolve in the water flow and precipitate on the grains of the rock. On one hand, compaction is due to the burial of the layer, which makes the overburden increase. On the other, the presence of the precipitated mineral can affect the solid matrix porosity, since the dissolving mineral may leave some void spaces, whereas the precipitated mineral may fill them. Moreover, an alteration of porosity implies a variation of permeability that affects the pressure of the water flowing through the rock layer. Again, fluid pressure counteracts compaction reducing the effective vertical stress, which has an effect on porosity. Finally, the chemical reactions that cause minerals to dissolve and precipitate are influenced by the fluid flow, which transports the solute.

The result is a nonlinear system of strongly coupled equations. The flow is assumed to obey Darcy's law. We use a simplified description of chemical reactions to model the

precipitation and dissolution of a mineral, such as quartz, see [7, 11, 2]. The effect of the fluid-solid conversion on porosity is accounted for by modifying Athy's constitutive law for the porosity [1], following [15, 16]. Again, we follow the strategy proposed in [15] and recast the governing equations in a Lagrangian frame. We simulate the sedimentation process by providing a sedimentation rate that causes the overburden to increase.

The system is solved by splitting the stronger coupling between pressure and compaction from the solution of the advection-diffusion-reaction equation for the transported solute. One main issue is that the equation that models the reaction of the mineral is a differential equation with discontinuous right hand side [10]. This led us to introduce another splitting between the advection-diffusion part of this equation and the reaction one, to take advantage of the available *ad hoc* methods to manage the discontinuity. In particular, among the possible strategies such as time step adaptation [8], or regularization of the right hand side [4], we chose to adopt an event-driven approach where the crossing of the discontinuity is exactly localized before restarting integration [5].

The paper is structured as follows. In Section 2 we present the mathematical model and we recast it into a Lagrangian frame. Section 3 deals with the numerical approach. In particular, we dwell on the application of a method to treat ODEs with discontinuous right hand side to our specific case. Finally in Section 4 the numerical solution obtained in one test configuration is analyzed, while conclusions are drawn in Section 5.

2 The model

We consider a layer of a sedimentary rock subject to a progressive burial. We assume that the porous medium is saturated with water. We are interested in modeling the behavior of a reactive material that can be advected by the flow when dissolved in water, and precipitate on the grains surface in the rocks. On the other hand, the precipitated mineral may dissolve into the fluid. It is clear that this mechanism, together with the increasing load due to the burial of the layer, affects the porosity ϕ of the sediments, leading to both geochemical and mechanical compaction.

2.1 The domain

We consider a two-dimensional model of the aforementioned processes, simulated in a vertical cross section of a sedimentary layer. Due to compaction, the domain of interest Ω evolves during the simulation. However, it is more convenient to cast the coupled problem of fluid flow, compaction and chemical reactions in a fixed geometry. For this reason, under the assumption that compaction acts only vertically, we define a fixed domain $\hat{\Omega}$, obtained from $\Omega = \Omega(t)$, as its completely compacted configuration once removed the reactive (dissolvable) part of the rock.

More precisely, following [15] and [16], we assume that one can define at any point $\boldsymbol{x} \in \Omega(t)$ and at any time the field $C = C(\boldsymbol{x}, t)$, which represents the ratio between the volume of the reactive part of the rock and the initial rock volume, so that the map

$\boldsymbol{\varphi}_t : \hat{\Omega} \to \Omega(t)$, $(x, \xi) \mapsto (x, z(\xi, t))$, is well defined and depends on C and ϕ. The deformation gradient $\boldsymbol{F}$ associated to this map is

$$\boldsymbol{F} := \nabla \boldsymbol{\varphi}_t = \begin{bmatrix} 1 & 0 \\ \partial z/\partial x & \partial z/\partial \xi \end{bmatrix}, \tag{1}$$

and it can be shown that the Jacobian is

$$J := \det(\boldsymbol{F}) = \frac{\partial z}{\partial \xi} = \frac{1 - C_0 + C}{(1 - C_0)(1 - \phi)} > 0, \tag{2}$$

where $C_0(\boldsymbol{x}) = C(\boldsymbol{x}, 0)$.

Note that the time derivative of the map coincides with the velocity $\boldsymbol{u}_s$ of the sediments and that, due to the hypothesis of vertical compaction, one has $\boldsymbol{u}_s = u_{sz}\boldsymbol{e}_z$, being $\boldsymbol{e}_z$ the unit vector of the z-axis.

2.2 Governing equations

In this section we present the equations governing the porous matrix evolution, fluid flow and chemical reactions in the fixed domain $\hat{\Omega}$. See [9] for a detailed derivation of the model.

The generic scalar field f in Ω becomes $\hat{f} = f \circ \boldsymbol{\varphi}_t$ in $\hat{\Omega}$, and the generic vector field $\boldsymbol{v}$ is transformed through the Piola transformation, that is $\hat{\boldsymbol{v}} = \hat{J}\, \hat{\boldsymbol{F}}^{-1} \boldsymbol{v} \circ \boldsymbol{\varphi}_t$, where $\hat{J}$ and $\hat{\boldsymbol{F}}$ are the transformation of (1) and (2) in $\hat{\Omega}$. Finally, the nabla operator in the fixed coordinate system is indicated with $\hat{\nabla}$.

If we assume that no water is released or consumed during the reactions, mass conservation for the fluid phase reads

$$\frac{\partial(\hat{\phi}\hat{\rho}_w \hat{J})}{\partial t} + \hat{\nabla} \cdot (\hat{\phi}\hat{\rho}_w \hat{\boldsymbol{u}}) = 0 \qquad \text{in } \hat{\Omega} \times (0, T), \tag{3}$$

where $\hat{\rho}_w$ is the density of water, and $\hat{\boldsymbol{u}} = \hat{\boldsymbol{u}}_w - \hat{\boldsymbol{u}}_s$.

The relative velocity $\hat{\boldsymbol{u}}$ of water $\hat{\boldsymbol{u}}_w$ with respect to that of the solid matrix $\hat{\boldsymbol{u}}_s$ is related to the pore pressure $\hat{p}$ by Darcy's law, i.e.

$$\hat{\phi}\hat{\boldsymbol{u}} = -\hat{J}\frac{\widetilde{\boldsymbol{K}}}{\mu_w}\left(\hat{\nabla}\hat{p} - \hat{\rho}_w \hat{\boldsymbol{F}}^T \boldsymbol{g}\right) \qquad \text{in } \hat{\Omega} \times (0, T), \tag{4}$$

with $\boldsymbol{g} = -g\boldsymbol{e}_z$. Here μ_w denotes the water viscosity and $\widetilde{\boldsymbol{K}} := \hat{\boldsymbol{F}}^{-1} \boldsymbol{K}(\hat{\phi}) \hat{\boldsymbol{F}}^{-T}$, where the permeability tensor $\boldsymbol{K}$ is assumed to be isotropic, thus

$$\boldsymbol{K}(\phi) = \mathcal{K}(\phi)\boldsymbol{I} \tag{5}$$

being $\mathcal{K}(\phi)$ given by the following relation, see [3],

$$\mathcal{K}(\phi) = \begin{cases} k_0\phi^3 \text{ if} & \phi \geq 0.1 \\ \dfrac{100\ k_0\phi^5}{(1-\phi)^2} \text{ if} & \phi < 0.1 \end{cases}. \tag{6}$$

Mechanical compaction of porous media is usually modeled by a relation between porosity and effective vertical stress known as Athy's law, [1]. However, in the case of our interest, porosity depends also on the concentration of precipitated mineral. According to [15], we model the coupled effect of dissolution/precipitation and compaction with the following equation,

$$\hat{\phi} = (\phi_0 + (1-\phi_0)(\hat{C}_0 - \hat{C}))\ e^{-\beta\hat{\sigma}}, \tag{7}$$

which is a generalization of Athy's law. Here, $\hat{\sigma}$ denotes the vertical effective stress defined as $\hat{\sigma} = \hat{s} - \hat{p}$, where $\hat{s}$ is the overburden and can be obtained integrating the differential equation

$$\frac{\partial \hat{s}}{\partial \xi} = -[(1-\hat{\phi})\hat{\rho}_s + \hat{\phi}\hat{\rho}_w]\ g\hat{J} \tag{8}$$

with the boundary condition $\hat{s}(x, \xi_{top}, t) = s_{\text{top}}(t)$, where s_{top} is a given function of time and accounts for the weight of the overlying layers. The density ρ_s of the solid matrix is obtained as the average of the density of the inert part of the rock ρ_r and that of the precipitated mineral $\hat{\rho}_p$, weighted with their volume fractions, that is

$$\hat{\rho}_s = \frac{(1-\hat{C}_0)\hat{\rho}_r + \hat{C}\hat{\rho}_p}{1-\hat{C}_0+\hat{C}}. \tag{9}$$

The equations (3), (4), (7), and (8) describe the coupling between water flow and compaction. We now introduce the chemical reactions that model the precipitation and dissolution of the transported mineral specie. Let us introduce the field $\hat{\gamma}$ that represents the dissolved mineral concentration in terms of moles per unit volume of water. The dissolved mineral is allowed to diffuse, to be transported by the fluid flow and to interact with the solid matrix, i.e. to precipitate, behaving as prescribed by the following equation:

$$\frac{\partial}{\partial t}(\hat{\gamma}\hat{\phi}\hat{J}) + \hat{\nabla}\cdot\left(\hat{\phi}\hat{\gamma}\hat{\boldsymbol{u}} - D\hat{\phi}\hat{\boldsymbol{F}}^{-T}\hat{\nabla}\hat{\gamma}\right) = r(\hat{C},\hat{\gamma})\hat{\phi}\hat{J} \qquad \text{in } \hat{\Omega}\times(0,T), \tag{10}$$

where $D > 0$ is the diffusion coefficient and $r(\hat{C},\hat{\gamma})$ is a source/well term that represents the dissolution/precipitation rate of the mineral. We point out that, since we are considering low velocities, we are here neglecting the effect of dispersion.

On the other hand, a source/sink term for equation (10) implies a sink/source term for the equation of the volume fraction of the precipitated mineral $\hat{C}$, as stated by the following equation:

$$\frac{\partial \hat{C}}{\partial t} = -V_m r(\hat{C},\hat{\gamma})\hat{\phi}, \tag{11}$$

where V_m is the molar volume of the mineral.

Following [2], we model the reaction rate r as a discontinuous function of γ and C. Let us introduce the following notation

$$x^+ := \max(0,x), \quad x^- := (-x)^+.$$

We model the source/sink term for the equations (10) and (11) as

$$r(C,\gamma) = \lambda \left(\text{sign}(C)^+ F(\gamma)^- - F(\gamma)^+\right), \tag{12}$$

where

$$F(\gamma) = \frac{\gamma}{\gamma_{eq}} - 1, \quad \gamma_{eq} > 0.$$

Here, γ_{eq} denotes an equilibrium concentration of the dissolved mineral. The rate constant λ is modeled according to Arrhenius law as

$$\lambda = \bar{\lambda} e^{-\frac{E}{RT}} > 0,$$

where E is an activation energy, R is the gas constant and the temperature T is assumed to be a given function of time.

We observe that, if the solute concentration exceeds the equilibrium value, $\gamma > \gamma_{eq}$, then $F(\gamma) > 0$ and $r = -\lambda\left(\frac{\gamma}{\gamma_{eq}} - 1\right) < 0$. In this case, precipitation occurs. On the other hand, if $\gamma < \gamma_{eq}$, then $F(\gamma) < 0$ and $r = \lambda\ \text{sign}(C)^+\left(1 - \frac{\gamma}{\gamma_{eq}}\right) \geq 0$. In this case, if $\text{sign}(C) > 0$ (i.e. if some precipitated is available in the rock), dissolution occurs. Finally, in case $\gamma = \gamma_{eq}$, $F(\gamma) = 0$ and the chemical equilibrium implies $r = 0$.

The solid mass conservation equation allows us to compute, knowing porosity and precipitate concentration, the velocity of the solid matrix and therefore the deformed configuration of the layer. Indeed, in this framework mass conservation of the solid phase in a porous medium implies

$$\frac{\partial}{\partial t}\left((1-\phi)\rho_s\right) + \frac{\partial}{\partial z}\left((1-\phi)\rho_s u_{sz}\right) = Q_s \qquad \text{in } \Omega(t) \times (0,T), \tag{13}$$

where Q_s is a source/sink term that models the growth or consumption of the solid grains, which can be shown to be

$$Q_s = \rho_p \frac{(1-\phi)}{1 - C_0 + C} \frac{DC}{Dt}. \tag{14}$$

To recover the deformed configuration of the layer, one can solve (13) for u_{sz} with a Dirichlet condition on the bottom boundary, and solve then $\frac{\partial z}{\partial t} = u_{sz}$ with a proper initial condition. We point out that, since we are considering the evolution of a single layer, such boundary conditions should be provided by the reconstruction of the history of the whole sedimentary basin.

The described system, complemented with suitable initial and boundary conditions, is a nonlinear system of strongly coupled equations. Indeed, it is clear that the changes in porosity can cause overpressures (i.e. pressures larger than hydrostatic) because permeability is a function of porosity, and moreover porosity plays a role in the storage term of fluid mass conservation. On the other hand, fluid pressure can counteract compaction reducing the effective vertical stress. Finally, chemical reactions are influenced by the fluid flow which transports the solute, and can increase or reduce the porosity if dissolution or precipitation occur. The approximation strategy implemented to tackle this coupling is illustrated in the next section.

3 The numerical approximation

For the solution of the coupled problem described in the previous section one could opt for a fully coupled approach using Newton iterations. However, even if a fully coupled approach is in general more robust, it is computationally very expensive and moreover the Jacobian matrix is likely to be ill-conditioned because of the different scales involved in the equations. For this reason we resort to an iterative splitting where the problems are solved in sequence, performing fixed point iterations until convergence is achieved.

3.1 Time discretization and iterative splitting

In principle all the aforementioned problems, i.e. fluid flow, compaction and solute dissolution/precipitation are coupled. If we assume that the effect of chemistry on porosity is moderate and relatively "slow" we can solve via fixed point iterations only the stronger coupling between pressure and compaction, and solve the advection, diffusion and reaction for the solute once per time step, reducing the computational cost. Therefore, for each time, we perform the following steps:

- integrate the chemical reactions and compute the solute using (10) and (11);
- enter the fixed point loop:
  - compute the sedimentary load and the effective stress with equation (8);
  - update of the porosity with equation (7);
  - solve Darcy's problem to obtain fluid pressure with equations (3) and (4);
  - check for convergence.

Since the reaction term in the ADR equation for γ is discontinuous, its approximation can benefit from a tailored integration scheme. Therefore it is very convenient, though not mandatory, to split the equation into an advection-diffusion part and a reaction part. This way, we can split the coupled problem of (10) and (11) into two sub-problems:

Advection-diffusion equation

$$\frac{\partial}{\partial t}(\hat{\gamma}\hat{\phi}\hat{J}) + \hat{\nabla}\cdot\left(\hat{\phi}\hat{\gamma}\hat{\boldsymbol{u}} - D\hat{\phi}\hat{\boldsymbol{F}}^{-T}\hat{\nabla}\hat{\gamma}\right) = 0 \tag{15}$$

Reaction system

$$\begin{cases} \dfrac{\partial \hat{\gamma}}{\partial t} = r(\hat{C},\hat{\gamma}) \\ \dfrac{\partial \hat{C}}{\partial t} = -V_m r(\hat{C},\hat{\gamma})\hat{\phi} \end{cases} \tag{16}$$

The two problems are solved in sequence according to a second order Strang splitting ([12, 14]). If we denote with $\hat{\gamma}^{n,n+1}$, $\hat{C}^{n,n+1}$ the concentrations at the discrete time t^n and t^{n+1} respectively, with $\hat{\gamma}^{*,**}$ two intermediate values of $\hat{\gamma}$ and with Δt the time step amplitude, the splitting consists in performing the following three steps

$$
\begin{aligned}
&I)\ \frac{2(\hat{\gamma}^*\hat{\phi}\hat{J} - \hat{\gamma}^n\hat{\phi}\hat{J})}{\Delta t} + \hat{\nabla}\cdot\left(\hat{\phi}\hat{\gamma}^*\hat{\boldsymbol{u}} - D\hat{\phi}\hat{\boldsymbol{F}}^{-T}\hat{\nabla}\hat{\gamma}^*\right) = 0\\
&II)\ \begin{cases} \dfrac{\hat{\gamma}^{**} - \hat{\gamma}^*}{\Delta t} = r(\hat{C}^n, \hat{\gamma}^*) \\ \dfrac{\hat{C}^{n+1} - \hat{C}^n}{\Delta t} = -V_m r(\hat{C}^n, \hat{\gamma}^*)\hat{\phi} \end{cases}\\
&III)\ \frac{2(\hat{\gamma}^{n+1}\hat{\phi}\hat{J} - \hat{\gamma}^{**}\hat{\phi}\hat{J})}{\Delta t} + \hat{\nabla}\cdot\left(\hat{\phi}\hat{\gamma}^{n+1}\hat{\boldsymbol{u}} - D\hat{\phi}\hat{\boldsymbol{F}}^{-T}\hat{\nabla}\hat{\gamma}^{n+1}\right) = 0.
\end{aligned}
$$

Note that we have chosen an implicit, thus more stable, discretization for the advection-diffusion part, while as concerns the reaction part an explicit scheme is more suitable for the implementation of the event detection method described in Section 3.3. We also point out that we are employing a higher order splitting to achieve better accuracy. Since for each time step we also need to solve some fixed point iterations for pressure and porosity, the use of an higher order splitting turns out to be more convenient than using a smaller the time step.

3.2 Finite element discretization

We have chosen a mixed finite element method for both the Darcy's problem and the advection-diffusion part of the equation for the solute concentration. This allows us to use the same finite element approximation for the relative velocity $\hat{\boldsymbol{u}}$ in equations (3) and (10). The finite element space chosen for the relative velocity $\hat{\boldsymbol{u}}$ is the lowest order Raviart Thomas $\mathbb{RT}_0(\hat{\Omega}, \mathcal{T}_h) \subset H(\mathrm{div}, \hat{\Omega})$, while the solute concentration γ and the water pressure p are in the space of the piece-wise constant functions $\mathbb{P}_0(\hat{\Omega}, \mathcal{T}_h) \subset L^2(\hat{\Omega})$. In both equations, since we are considering mixed formulations, the Dirichlet boundary conditions on pressure and concentration are naturally included in the weak formulation, while the Neumann boundary conditions on normal velocity and flux are imposed with a Nitsche's penalization technique (see [13]). Finally, the differential equation for the computation of the overburden s is solved with a SUPG stabilized finite element method, using $\mathbb{P}_1$ elements.

3.3 Numerical solution of the discontinuous ODEs for dissolution/precipitation

We now focus on the numerical approximation of the reaction part of the coupled problem (10), (11). Thanks to the splitting we can employ an *ad hoc* method for discontinuous ODEs. Indeed, once the problem has been discretized in space with the finite element method as described in the previous section, equations (16) become a system of ODEs for

each single degree of freedom. If we denote as $\hat{\boldsymbol{\gamma}}$, $\hat{\mathbf{C}}$ the vectors containing the degrees of freedom representing the solute and precipitate concentrations, being $\mathbf{r}$ the corresponding vector of the reaction rates defined as in (12), the two equations (10), (11) make up an ODE system with discontinuous right hand side, to whom Filippov theory can be applied. To this purpose, we define

$$\boldsymbol{y} = \begin{bmatrix} \hat{\boldsymbol{\gamma}} \\ \hat{\mathbf{C}} \end{bmatrix} \qquad \text{and} \qquad \boldsymbol{f} = \begin{bmatrix} \mathbf{r} \\ -V_m \mathbf{r} \hat{\phi} \end{bmatrix},$$

and we observe that

$$\frac{\partial y_i}{\partial t} = \begin{cases} f_i^1 & C = 0 \text{ and } \gamma_i < \gamma_{eq} \\ f_i^2 & \text{elsewhere} \end{cases}, \tag{17}$$

where

$$\boldsymbol{f}^1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \qquad \text{and} \qquad \boldsymbol{f}^2 = \begin{bmatrix} \lambda \left(1 - \frac{\gamma}{\gamma_{eq}}\right) \\ -\lambda \phi V_m \left(1 - \frac{\gamma}{\gamma_{eq}}\right) \end{bmatrix}.$$

Numerical methods for the integration of DRH-systems can deal with the discontinuity with different approaches, such as step adaptation, or smoothing of the right hand side. The computational approach used in this work, proposed by [6], is designed to locate with accuracy the points where the transition occurs and check the transversality/sliding conditions of [6] through the following steps:

- integration of the ODE outside Σ, in particular, we use an explicit order 2 Runge Kutta method;
- location of the point $\mathbf{y} \in \Sigma$ reached by a trajectory, and the corresponding time t^* with an iterative method;
- check whether a crossing of the discontinuity has occurred;
- in case of sliding along the discontinuity, integration on Σ with the proper right hand side.

These steps are performed for each degree of freedom during the integration of (16) to achieve a good accuracy and, most of all, to avoid unphysical solutions such as the occurrence of negative concentrations.

4 Results

We simulate the compaction process of a $200m \times 120m$ sedimentary layer buried at the depth d. At the beginning of the simulation $d = d_0$, then a sedimentation velocity $\frac{\partial d}{\partial t} > 0$ brings the domain at a depth $d(t)$. We do not model the addition of extra layers due

to the progressive burial and the sedimentation acts only as a variation of the boundary conditions. For this reason, the boundary conditions for pressure and overburden are time-dependent. Temperature is a given field and is obtained with a geothermal gradient $\frac{\partial T}{\partial d}$ and a surface temperature T_0. The parameters used for the simulation are summarized in table 1.

	Value	Unit		Value	Unit		Value	Unit
β	10^{-8}	Pa^{-1}	ϕ_0	0.5	-	$\bar{\lambda}$	8.37 10^{-6}	$mol/(m^3\ s)$
d_0	2000	m	$\frac{\partial d}{\partial t}$	100	m/My	E	60.1	kJ/mol
T_0	20	°C	$\frac{\partial T}{\partial d}$	0.035	°C$/m$	γ_{eq}	0.167	mol/m^3
k_0	10^{-6}	$Darcy$	g	9.81	m/s^2	$\bar{\rho}$	2500	kg/m^3
μ_w	0.001	$Pa\ s$	D	1.58 10^{-8}	m^2/s	ρ_d	2500	kg/m^3
ρ_w	1000	kg/m^3	V_m	0.0226	m^3/mol	ρ_m	2660	kg/m^3

Table 1: Physical parameters for the simulation.

The numerical setup of this simulation is shown in figure 1. We have set for pressure hydrostatic Dirichlet boundary conditions at the top of the domain and at the bottom. The domain is considered as a part of a longer thin layer of rock, lying along the x-direction, hence no-flux boundary conditions are imposed on the lateral edges. A Dirichlet condition for the overburden is set at the top and we assume that the bottom of the domain moves downwards with a given, and in our case uniform, velocity. Finally, the dissolved mineral concentration is prescribed both at the top of the domain, where it is greater than that of equilibrium, and at the bottom, where it is equal to zero.

The rock is initially filled with water with no dissolved mineral ($\gamma_0 = 0$). The initial condition for pressure is the hydrostatic pressure and the initial conditions for stress and porosity are computed with some fixed point iterations of the stationary problem. The initial distribution of the precipitated mineral in the rock is sketched in figure 1. Finally, $\hat{\boldsymbol{u}}^0 = \mathbf{0}$.

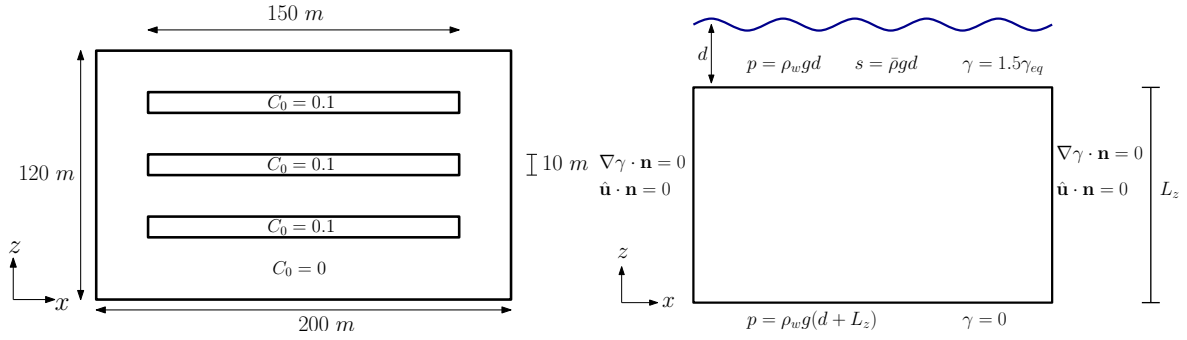


Figure 1: Numerical setup.

We solve the problem on a triangular mesh 50×30 with time step $\Delta t = 0.5\ 10^{11}\ s \approx 1.5\ ky$ to simulate a time span of $T = 25\ My$.

As expected, the mineral in water starts immediately to precipitate at the top of the domain, where its concentration is higher than that of equilibrium. At the same time, the precipitated mineral in rock dissolves in the lower region, where $\gamma < \gamma_{eq}$, behavior which lasts as long as there is mineral in rock that can dissolve, see Figure 2.

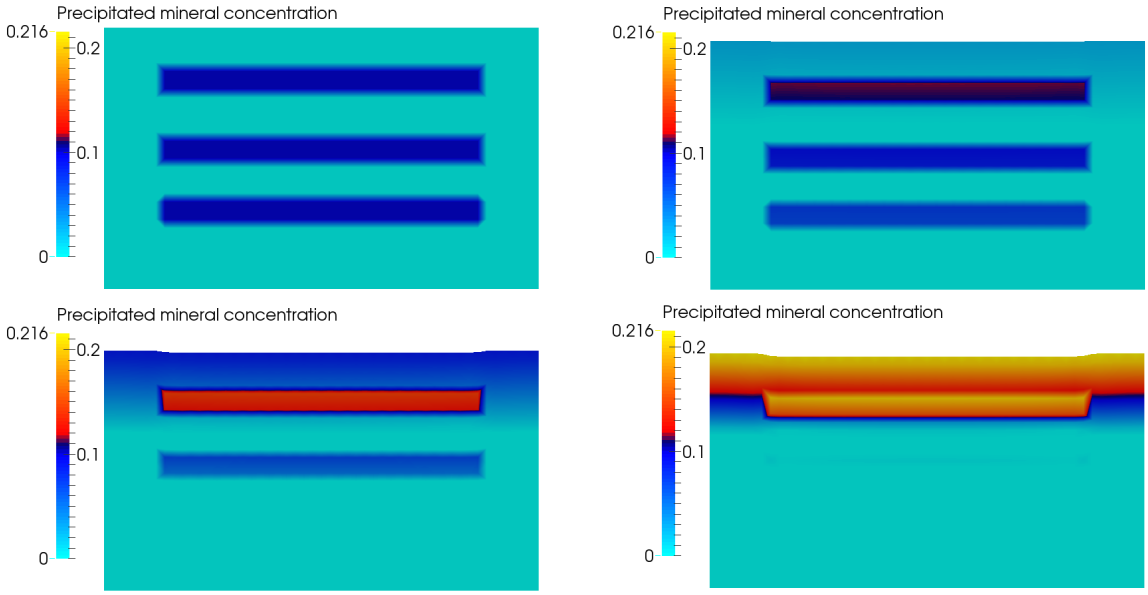


Figure 2: Plot of C at $t = 0$, $t = 8\ My$, $t = 16\ My$, and $t = 24\ My$ on the physical domain $\Omega(t)$.

In Figure 3, the porosity is shown, which decreases during the simulation due to the increase of overburden. An higher porosity is obtained, as expected, where the mineral originally present on the rock dissolves, since the dissolving mineral leaves some void spaces. On the other hand, porosity is lower at the top of the domain, where the mineral precipitates. Due to the non-uniform porosity, the domain compacts in a non-uniform way and we can clearly see at the end of the simulation that the region in which precipitated mineral has dissolved compacts more than its neighbor region.

5 Conclusion

We have proposed and tested a discretization method for the simulation of compaction in porous media with a particular focus on the numerical treatment of the discontinuous reaction terms that may arise in the modeling of geochemical processes. Even if we are considering a simplified model where only one mineral is allowed to precipitate, dissolve and be advected by the water flow, the results are qualitatively correct and the rigorous treatment of the discontinuity avoids the occurrence of negative concentrations and oscillations. The whole approximation strategy has been developed with the aim of finding

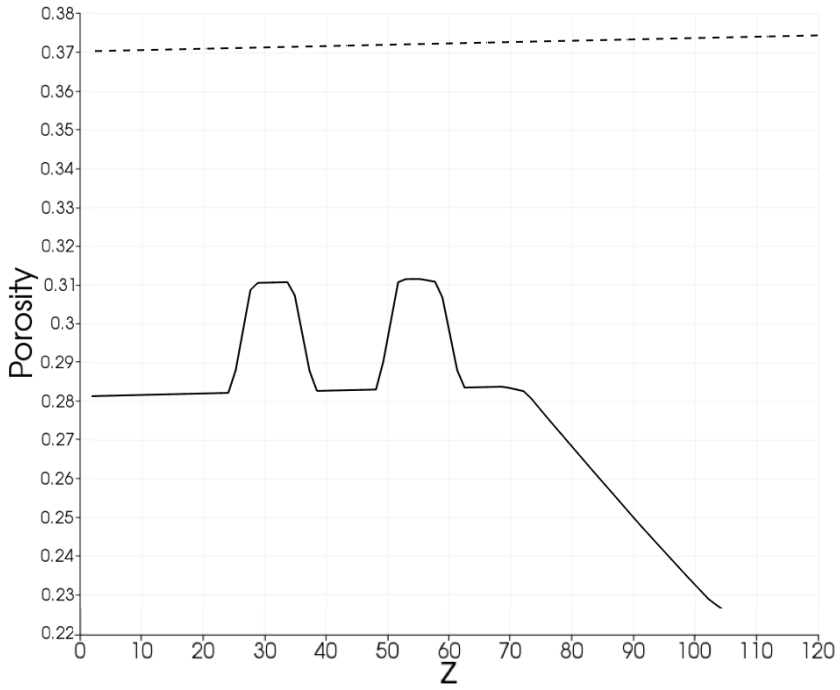


Figure 3: Plot of porosity versus z at $t = 0$, $t = 8\ My$, $t = 16\ My$, and $t = 24\ My$, taken on the vertical line that halves the domain. The dashed and the solid lines represent the initial and the final porosity respectively.

a trade-off between accuracy and computational efficiency. However, there is still room for improvement: in particular a more precise assessment of the error introduced by the splittings will be the subject of future work.

REFERENCES

[1] L. F. Athy. Density, porosity, and compaction of sedimentary rocks. *AAPG Bulletin*, 14(1):1–24, 1930.

[2] N. Bouillard, R. Eymard, R. Herbin, and P. Montarnal. Diffusion with dissolution and precipitation in a porous medium: mathematical analysis and numerical approximation of a simplified model. *ESAIM: Mathematical Modelling and Numerical Analysis*, 41:975–1000, 2007.

[3] Z. Chen, R. E. Ewing, H. Lu, S. L. Lyons, S. Maliassov, M. B. Ray, and T. Sun. Integrated two-dimensional modeling of fluid flow and compaction in a sedimentary basin. *Computational Geosciences*, 6:545–564, 2002.

[4] M.F. Danca. On the uniqueness of solutions to a class of discontinuous dynamical systems. *Nonlinear analysis: real world applications*. in press, doi:10.1016/j.norwa.2009.02.024.

[5] L. Dieci and L. Lopez. A survey of numerical methods for ivps of odes with discontinuous right-hand side. *Journal of Computational and Applied Mathematics*, 236(16):3967–3991, October 2012.

[6] Luca Dieci and Luciano Lopez. Numerical solution of discontinuous differential systems: Approaching the discontinuity surface from one side. *Applied Numerical Mathematics*, 67:98–110, 2013.

[7] A. C. Fowler and X. S. Yang. Dissolution/precipitation mechanisms for diagenesis in sedimentary basins. *Journal of Geophysical Research: Solid Earth (1978–2012)*, 108(B10), 2003.

[8] C. W. Gear and O. Østerby. Solving ordinary differential equations with discontinuities. *ACM transaction on mathematical software*, 10(1):23–24, 1984.

[9] B. Giovanardi, A. Scotti, L. Formaggia, and P. Ruffo. *A general framework for the simulation of geochemical compaction.* Mox Report, 2015.

[10] K. Kumar, I.S. Pop, and F.A. Radu. Convergence analysis of mixed numerical schemes for reactive flow in a porous medium. *SIAM Journal on Numerical Analysis*, 51(4):2283–2308, 2013.

[11] R. H. Lander and O. Walderhaug. Predicting porosity through simulating sandstone compaction and quartz cementation. *AAPG bulletin*, 83(3):433–449, 1999.

[12] D. Lanser and J. G. Verwer. Analysis of operator splitting for advection–diffusion-reaction problems from air pollution modelling. *Journal of Computational and Applied Mathematics*, 111(12):201 – 216, 1999.

[13] J. Nitsche. Über ein Variationsprinzip zur Lösung von Dirichlet-Problemen bei Verwendung von Teilräumen, die keinen Randbedingungen unterworfen sind. *Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg*, 36:9–15, 1971.

[14] D. L. Ropp and J. N. Shadid. Stability of operator splitting methods for systems with indefinite operators: Advection-diffusion-reaction systems. *Journal of Computational Physics*, 228(9):3508 – 3516, 2009.

[15] M. Wangen. Vertical migration of hydrocarbons modelled with fractional flow theory. *Geophysical Journal International*, 115:109–131, 1993.

[16] M. Wangen. Two-phase oil migration in compacting sedimentary basins modelled by the finite element method. *International Journal for Numerical and Analytical Methods in Geomechanics*, 21:91–120, 1997.

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis (Eds)

A MAGNETO-THERMO-METALLURGICAL FINITE ELEMENT MODEL APPLIED TO INDUCTION HARDENING PROCESSES

MATTIA SPEZZAPRIA, MICHELE FORZAN AND FABRIZIO DUGHIERO

Laboratory of Electroheat (LEP) - Department of Industrial Engineering (DII)
University of Padova
Via Gradeniglo 6/a, 35131 Padova, Italy
e-mail: mattia.spezzapria@dii.unipd.it, web page: http://www.dii.unipd.it

Key words: Induction Hardening, Multiphysics, Phase transformation, Quenching

Abstract. Induction hardening has been widely applied for the heat treatment of components mainly in the aeronautical and automotive sectors because of its peculiar advantages like high quality and repeatability of process and its easy automation. A multi-scale multiphysical finite element (FE) analysis is presented in this paper for the prediction of microstructural evolution during induction hardening processes. An ad hoc external routine has been developed in order to calculate the phase changes during heating and cooling process associated with non-isothermal transformations. This routine has been coupled with commercial FEM codes able to solve the coupled electromagnetic and thermal problem that typically describes the induction heating processes. During the heating, the magnetic field generated by the coil induces currents in the workpiece and as consequence the heating of conductive material by Joule effect.
Material properties depend on the temperature distribution but also on the microstructure since the material could be seen as a mixture of different phases, each one with different physical properties. The effect of latent heat of solid-solid phase transformations has been also considered.
From the solution of the coupled steady-state, at a given frequency, electromagnetic and transient thermal problem, temperature distribution as well as heating and cooling rates are used for the evaluation of the existing metallurgical phases at every time step.

1 INTRODUCTION

Heat treatments have been traditionally used in order to improve the mechanical properties of steel parts. In the particular case, superficial heat treatments permit to modify only the external layer of a work-piece, maintaining unchanged the core of the material. Nowadays Contour Induction Hardening (CIH) process is increasingly applied instead of carburizing process, due to its repeatability and easy automation.

During induction hardening, the magnetic field generated by a coil induces current in the work-piece, which is heated by Joule effect. Thermal history induces, during the process, a solid-solid phase transformation in the material.

Due to its multiphysical nature, simulation tools become essential for the prediction of phase transformations.

Many studies have been carried out in the field of heating by induction [1, 2, 3, 4] and in

the specific case of induction hardening [5, 6, 7], but the multiphysical simulation of all the phenomena that occur during the whole process is still an open task.

Several models for the simulation of quenching have been proposed in the past [8, 9, 10, 11, 12, 13], but in the most of cases the body was uniformly heated above the austenitization temperature. In induction hardening, temperature distribution exhibits steep gradients and the actual distribution of austenitization in the piece is not uniform.

In other hand many authors proposed different models for the simulation of laser welding process [14, 15, 16], that can be considered, for some aspects (phase transformations, high heating rate), similar to induction hardening.

2 ELECTROMAGNETIC AND THERMAL FORMULATION

The mathematical analysis of coupled electromagnetic and thermal problem has been carried out by solving Maxwell and Fourier equations. The computation of magnetic flux density and induced current density is obtained by solving a time-harmonic eddy-current problem at a prescribed frequency.

The numerical solution of the EM problem has been carried out by applying the well-known Magnetic Vector Potential formulation where the Coulomb gauge has been imposed in order to guarantee the uniqueness of magnetic vector potential $\boldsymbol{A}$.

In the conductive region Faraday-Neumann law equation implies the existence of an electric scalar potential V, such that [2, 17]

$$\boldsymbol{E} = -j\omega\boldsymbol{A} - \nabla V \tag{1}$$

where $\boldsymbol{E}$ is the electric field and ω the angular frequency. Given that induced current density $\boldsymbol{J}$ can be written as:

$$\boldsymbol{J} = -\sigma(j\omega\boldsymbol{A} + \nabla V) \tag{2}$$

in which σ is the electrical conductivity of the material. The previous relations and Maxwell equations lead to the following system of equations:

$$\begin{cases} \nabla \times \left(\frac{1}{\mu}\nabla \times \boldsymbol{A}\right) + \sigma(j\omega\boldsymbol{A} + \nabla V) = 0 & (3) \\ \nabla \cdot \sigma(j\omega\boldsymbol{A} + \nabla V) = 0 & (4) \end{cases}$$

When the current density distribution is known in the conductive body, the distribution of power densities by Joule effect can be evaluated as follow:

$$w_p = \rho|J|^2 = \rho\omega^2|A|^2 \tag{5}$$

where ρ is the electrical resistivity.

The power densities calculated in the electromagnetic step are used as internal power densities in the transient thermal simulation and the thermal problem has been solved by means of Fourier equation for heat conduction:

$$\nabla(\lambda\nabla T) + w_p = \gamma C_p \frac{dT}{dt} \tag{6}$$

in which γ is the density and C_p the specific heat.

The thermal exchange by convection and radiarion have been taken into account during the heating stage on the surfaces of contact between the trated body and the air:

$$\Phi(T_S, T_\infty) = h(T_S - T_\infty) + k_B \varepsilon \left({T_S}^4 - {T_\infty}^4\right) \tag{7}$$

where h is the coefficient of heat exchange by convection, k_B the Stephan-Boltzmann constant and ε the emissivity of the surface. The temperature distribution calculated has been used for updating the electromagnetic properties in order to calculate accurately the Joule losses distribution for every time step of thermal transient simulation.

3 THERMO-METALLURGICAL FORMULATION

The nodal power densities calculated in the EM step for each time step are the internal heat sources in a coupled thermo-metallurgical model.

In the thermo-metallurgical model the material can be described as a mixture of different phases, each one with different physical properties. The global material properties can be extimated through a linear rule of mixture[8,9]:

$$P(T, \xi_k) = \sum_{k=1}^{N} P_k(T) \xi_k \tag{8}$$

where $P_k(T)$ is the value of the physical property, temperature dependent and ξ_k is the volume fraction of the k-th phase. During phase transformations latent heat is absorbed or released by the body and this effect mustbe included in the calculations for an accurate extimation of temperature distribution. The temperature distribution within the workpiece is determined by the modified heat conduction equation:

$$\nabla(\lambda \nabla T) + w_p + \sum_{k=1}^{N} \Delta H_k \frac{d\xi_k}{dT} = \gamma C_p \frac{dT}{dt} \tag{9}$$

where ΔH_k is the enthalpy change due to phase tranformation.

3.1 Austenitization

During the induction heating the initial microstructure (generally made of pearlite, ferrite and carbides) transforms when austenitization temperatures are reached: below the Ac1 temperature the microstructure is composed by ferrite and pearlite, between Ac1 and Ac3 it is a mixture of ferrite, pearlite and austenite and above Ac3 the microstructure is typically inhomogeneous austenite. These temperature are affected by the chemical composition, heating rate and also by the microstructure. The kinetics of transformation during high heating rate processes can be calculated through the use of Continous Heating Transformation (CHT) Diagrams, which provide an indication of the transformation temperatures at different heating rates. In this study, the CHT diagram has been derived from the literature [18] and the austenite transformation kinetic has been simplified as follow:

$$\xi_A = \begin{cases} 0 & , \quad T < Ac_1(\dot{T}) \\ \dfrac{T - Ac_1(\dot{T})}{Ac_3(\dot{T}) - Ac_1(\dot{T})} & , \quad Ac_1(\dot{T}) \le T \le Ac_3(\dot{T}) \\ 1 & , \quad T > Ac_3(\dot{T}) \end{cases} \tag{10}$$

3.2 Quenching

During quenching several types of microstructures are generated by the decomposition of austenite (pearlite, ferrite, bainite and martensite), depending on the cooling rate.
The microstructure can be evaluated starting from the thermal history using Isothermal (IT) diagrams. These diagrams can be drawn through a numerical method developed by Kirkaldy [19] and reviewed by Victor Li [20] derived from Zener [21] and Hillert [22] formulas, based on kinetic-chemical equations:

$$\tau(X,T) = \frac{F(C,Mn,Si,Ni,Cr,Mo,G)}{\Delta T^n \exp\left(-\frac{Q}{RT}\right)} S(X) \tag{11}$$

in which τ is the time needed to transform, X the chemical composition, T the temperature, ΔT the undercooling, Q the activation energy, R the gas constant, and n is an empirical constant dependent on the diffusion mechanism (n=2 for volume and n=3 for boundary diffusion). S(X) is the reaction term defined by Kirkaldy, which approximates the sigmoidal effect of phase transformation:

$$S(X) = \int_0^X \frac{dX}{X^{0.4(1-X)}(1-X)^{0.4X}} \tag{12}$$

Considering the study carried out by Victor Li in [20], a reasonable value of the activation energy for all diffusional transformation is 27500 Kcal/(mol °C)
The time of transformation for a fixed temperature can be found through the following expressions:

$$\tau_F = \frac{\exp(-4.25 + 4.12C + 4.36Mn + 0.44Si + 1.71Ni + 3.33Cr + 5.19\sqrt{Mo})}{2^{0.41G}(Ae_3 - T)^3 \exp\left(-\frac{27500}{RT}\right)} S(X) \tag{13}$$

$$\tau_P = \frac{\exp(1 + 6.31C + 1.78Mn + 0.31Si + 1.12Ni + 2.70Cr + 4.06Mo)}{2^{0.32G}(Ae_1 - T)^3 \exp\left(-\frac{27500}{RT}\right)} S(X) \tag{14}$$

$$\tau_B = \frac{\exp(-10.23 + 10.18 + 0.85Mn + 0.55Ni + 0.90Cr + 0.36Mo)}{2^{0.29G}(B_S - T)^2 \exp\left(-\frac{27500}{RT}\right)} S(X) \tag{15}$$

Respectively for ferrite, pearlite and bainite transformation.

The transformation temperatures have been calculated through the following models:

$$Ae_3 = 883.49 - 275.89C + 90.91C^2 - 12.26Cr + 16.45CCr - 29.96CMn + 8.49Mo + \\ -10.8CMo - 25.56Ni + 1.45MnNi + 0.76Ni^2 + 13.53Si - 3.47MnSi \tag{16}$$

$$Ae_1 = 739 - 22.8C - 6.38Mn + 18.2Si + 11.7Cr - 15Ni - 6.4Mo - 5V - 20Cu \tag{17}$$

$$B_S = 637 - 58C - 35Mn - 15Ni - 34Cr - 41Mo \tag{18}$$

$$M_S = 539 - 423C - 30.4Mn - 17.7Ni - 12.1Cr - 7.5Mo + 10Co - 7.5Si \tag{19}$$

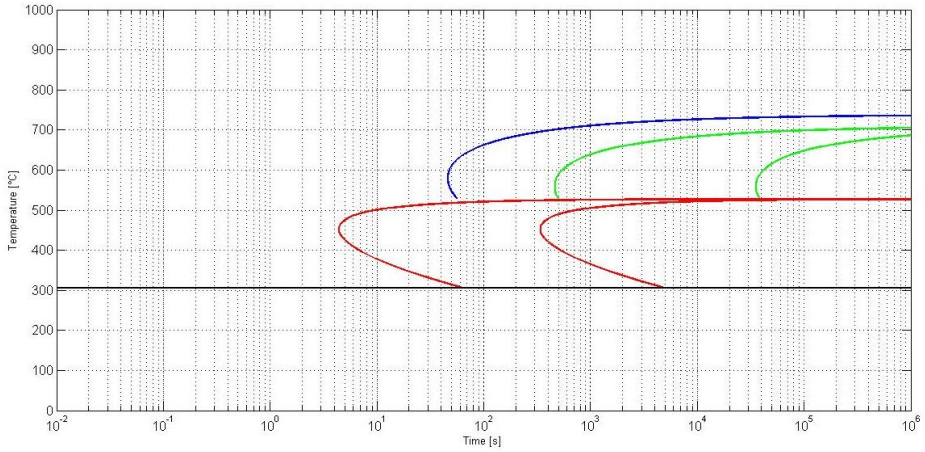


Figure 1: Calculated TTT Diagram for AISI4340

Once the IT diagram have been calculated, an analytical procedure has been developed for the calculation of microstructure evolution during quenching. Considering isothermal conditions, the kinetics of diffusional transformation can be expressed through the Johnson-Mehl-Avrami-Kolmogorov (JMAK) equation:

$$\xi_k = 1 - \exp(-b_k t^{n_k}) \tag{20}$$

where ξ_k is the total ammount of transformed phase, t is the time, b_k and n_k coefficients directly deduced from IT diagram.
Diffusive transformations always require an incubation time before starting and this can be explained through Sheil's additivity rule:

$$S = \sum_{i=1}^{n} \frac{\Delta t_i}{\tau_s(t_i)} = 1 \tag{21}$$

in which $\tau_s(t_i)$ is the incubation time at a current temperature and Δt_i is the time increment. Transformation begins when Scheil's sum is equal to unity.

During quenching the material is never subjected to isothermal conditions and also sometimes the austenite is not the only one metallurgical phase existing in the material. To solve this problem the thermal history during the quenching needs to be discretized into isothermal steps. The effective elapsed time must be corrected, becouse of the different kinetics of transformation that occur during each step and can be calculated as the sum of the time step and a fictitious time evaluated from the IT diagrams[13]:

$$t_k^* = \left[\frac{-ln(-\xi_{k-1})}{b_k}\right]^{\frac{1}{n_k}} \tag{22}$$

where ξ_{k-1} is the transformed phase in the previous step. Once the fictitious time is known a fictitious volume fraction ξ_i^* can be calculated through the JMAK equation:

$$\xi_k^* = 1 - exp[b_k(t_k^* + \Delta t)^{n_k}] \tag{23}$$

Hence the practical transformad volume fraction is:

$$\xi_k = (\xi_{k-1}^A - \xi_{k-1})\xi_{max} \tag{24}$$

in which ξ_{k-1}^{A} and ξ_{k-1} are respectively the austenite volume fraction and the transformed phase in the previous step and ξ_{max} is the maximum possible transformed fraction.
The equation considers that a part of the whole microstructure may not transform becouse already transformed or not austenitized; this is the case of coexistence of different phases due to a partial austenitization that occurs where Ac3 temperature has not been reached.

A different formaulation is used to model the martensitic transformation that describe this process as a diffusionless (or displacive) transformation.

The martensite volume fraction is generally evaluated through the Koistinen-Marburger (KM) model [23], but it tends to underextimate the transformed part in low-alloy steels. In this model, the martensite volume fraction has been calculated through a semi-empirical model proposed by Lee [10]:

$$\xi_M = \xi_A \cdot \{1 - \exp[-K_{LV}(M_S - T)^{n_{LV}}]\} \tag{25}$$

in which ξ_M is the total ammount of martensite, ξ_A the volume of parent phase, K_{LV} and n_{LV} are two coefficients dependent by the chemical composition:

$$K_{LV} = 0.0231 - 0.0105C - 0.0017Ni + 0.0074Cr - 0.0193Mo \tag{26}$$

$$n_{LV} = 1.4304 - 1.1836C + 0.7527C^2 - 0.0258Ni - 0.0739Cr + 0.3108Mo \tag{27}$$

4 FE SIMULATION PROCEDURE AND INPUT DATA

In this study the multiphysical model has been applied to a 2D axy-symmetric geometry. The electromagnetic and thermal simulation for the determination of Joule losses distribution has been solved by means of the commercial FEM software Flux 2D [24].
A circular billett (r=10 mm, h=60 mm) has been heated by a rectangular section coil (15x5x1.5 mm) witha C-shaped flux concentrator. Flux concentrator are often used in induction heating for increasing the efficiency of the process and improving thecontrol of heating pattern.
The billet is made in AISI 4340, a low alloy steel, often used in aeronautical and automotive sectors due to its high mechanical properties and hardenability.

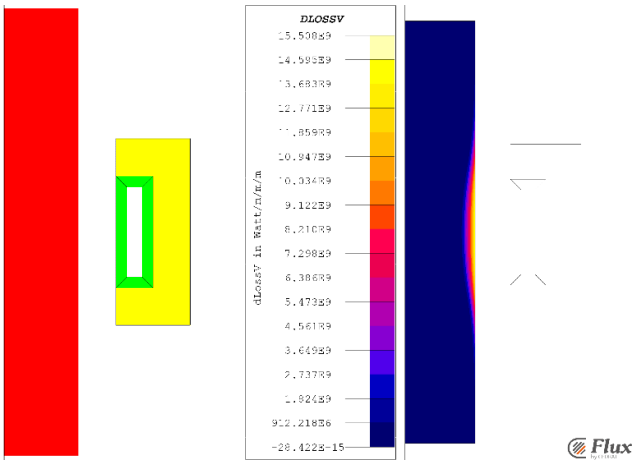


Figure 2: Geometry layout of the electromagnetic-thermal coupled simulation (left) and Joule losses distribution at the beginning of the heating process (right)

In this simulation the material of the billet has been considered as a unique phase with temperature dependent physiscal properties. An approximation of the effect of the latent heat absorbed during austenitic phase transformation have been introduced in the volumetric heat capacity.

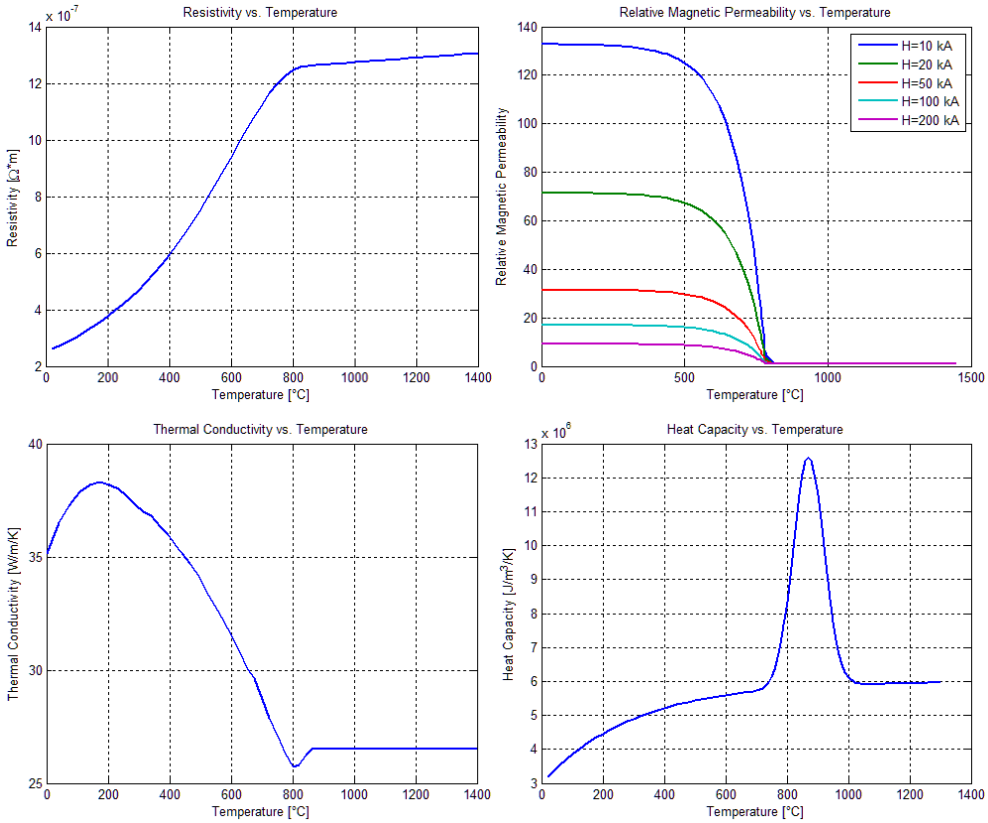


Figure 3: Physical properties of AISI 4340 - Resistivity (top-left), magnetic permeability (top-right), thermal conductivity (bottom-left) and volumetric heat capacity (bottom-right)

The heating process consists on a single shot step in wich the coil is fed by a current of 4300 A at 10 kHz for 1s. During heating the thermal losses by convection and radiation have been considered on the line that describes the external surface of the billet through a constant heat convection coefficient $\varepsilon = 15\ W/(m^2K)$ and a emissivity $\alpha = 0.8\ W/(m^2K^{-4})$.
At the end of electromagnetic-thermal coupled simulation developed in the commercial FEM code Flux2D, nodal power densities are exported at each time step to the thermo-metallurgical simualtion as internal power densities. The thermo-metallurgical simulation has been developed in Comsol with an ad hoc Matlab routine for the calculation of phase transformations.

In this model the material is modeled as a mixture of different phases,each one described by its specific material properties.

At the end of heating the billet is rapidly quenched with a acqueous polymer solution of polyalkilenglicole (PAG) at 12%. In order to simulate the strong temeprature variations of the quenching, a precise description of convective heat exchange parameter is mandatory. The mechanism of quenching is affected by many factors, which significantly influence the

performance of this process. In general, cooling occurs in three different stages: the early vapor phase, governed by convection and conduction through a vapor film around the work-piece; the main boiling phase, ruled by conduction between the hot surface and the quenchant, and the late convection phase into the liquid.

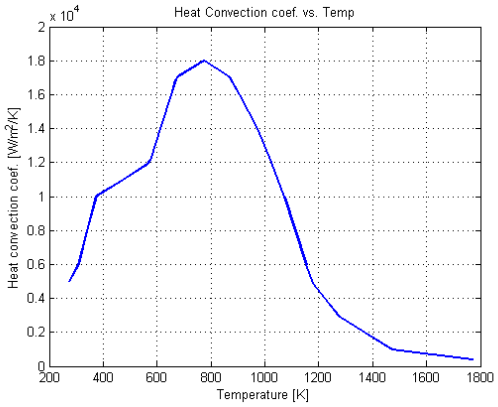


Figure 4: Convective heat exhange coefficient in function of the superficial temperature of the body

In figure 5 the martensite distribution at the end of quenching step is shown. In this type of calculation the effect of overtempering has not been taken into account insofar the tempering kinetics will be a subject of a further work.

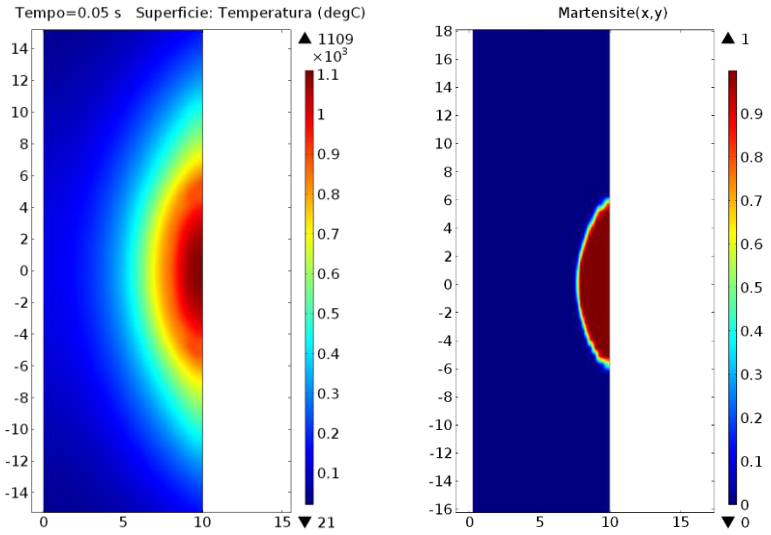


Figure 5: Temperature distriburion on the billet at the end of heating stage in the thermo-metallurgical simulation (left) and martensite distribution at the end of quenching (right)

5 CONCLUSIONS

A FEM based mathematical method has been developed in order to predict the phase transformations kinetics during the whole process of induction hardening. In particular:

- A non-linear magnetic behavior has been considered for the given steel, taking into account both magnetic saturation and Curie temperature.
- Both heating and quenching process have been simulated, calculating the complete thermal history and the microstructure at every time step that discretizes the process time.
- Different physical models have been weakly coupled in order to properly take into account the mutual dependences of material properties.

REFERENCES

[1] Dughiero, F., Forzan, M., Garbin, M., Pozza, C. and Sieni. E. A 3D numerical FEM model for the simualtion of induction welding of tubes, *COMPEL – The international journal for computation and mathematics in electrical and electronic engineering* (2011) **30**:1570-1581

[2] Dughiero, F., Forzan, M., Pozza. C. and Sieni, E., A Traslational Coupled Electromagnetic and Thermal Innovative Model for Induction Welding of Tubes, *IEEE Transaction on Magnetics* (2012) **48**:483-486

[3] Canova, A., Dughiero, F., Fasolo, F., Forzan, M., Freschi, F., Giaccone, L. and Repetto, M., Simplified approach for 3-D nonlinear induction heating problems, *IEEE Transaction on Magnetics* (2009) **45**:1855-1858

[4] Di Barba, P., Dughiero, F., Forzan, M. and Sieni, E., A Paretian Approach to Optimal Design with Uncertainties: Application in Induction Heating, *IEEE Transaction on Magnetics* (2014) **50**

[5] Schwenk, M., Hoffmeister, J. and Schulze, V. Experimental Determination of Process Parameters and Material data for Numerical Modeling of Induction Hardening, *Journal of Materials Engineering and Performance* (2013) **22**:1861-1870

[6] Nacke, B. and Wrona, E. Design of complex induction hardening problems by means of numerical simulation, *Archives of Electrical Engineering* (2005) **214**:461-466

[7] Candeo, A., Ducassy, C., Bocher, P. and Dughiero, F. Multiphysical modeling of induction hardening of ring gears for the aerospace industry, *IEEE Transaction on Magnetics* (2011) **47**:918-921

[8] Simsir, C. and Gür, C. H. A FEM based framework for simualtion of thermal treatments: Application to steel quenching, *Computational Materials Science* (2008) **44**:588-600

[9] Simsir, C. and Gür, C. H. 3D FEM simulation of steel quenching and investigation of the effect of asymmetric geometry on residual stress distribution, *Journal of Materials Processing Technology* (2008) **207**:211-221

[10] Lee, S. J., Matlock, D. K. and Van Tyne, C. J. Comparison of two finite element simulation codes used to model the carburizing of steel, *Computational Materials Science* (2013) **68**:47-54

[11] Leblond, J.B., Mottet, G., Devaux, J. and Devaux, J.C. Mathematical models of anisothermal phase transformations in steels, and predicted plastic behaviour, *Materials Science and Technology* (1985) **1**:815-822
[12] Kang, S.H. and Im, Y.T. Finite Element Investigation of Multi-Phase Transformation within Carburized Carbon Steel, *Journal of Materials Processing Technology* (2007) **183**:241-248
[13] Denis, S. Sjöström, S. and Simon, A. Coupled Temperature, Stress, Phase Transformation Calculation Model Numerical Illustration of the Internal Stresses Evolution during Cooling of a Eutectoid Carbon Steel Cylinder, *Metallurgical Transaction A* (1987) **18A**:1203-1212
[14] Haidemenopoulos, G. N. Coupled thermodynamic/kinetic analysis of diffusional transformations during laser hardening and laser welding, *Journal of Alloy and Compounds* (2001) **320**:302-307
[15] Dai, H. Modelling Residual Stress and Phase Transformations in Steel Welds, *Neutron Diffraction - Intech* (2012)
[16] Ferro, P. The use of Matlab in Advanced Design of Bonded and Welded Joints, *Application of MATLAB in Science and Engineering – Intech* (2011)
[17] Birò, O., Preis, K., Renhart, W., Vrisk, G. and Richter, K. R. Computation of 3D current driven skin effect problems using a current vector potential, *IEEE Transaction an Magnetics* (1993) **29**:1325-1328
[18] Orlich, J., Rose, A. and Wiest, P. *Atlas zur Wärmebehandlung der Stähle*, Vol. 3, Zeit-Temperatur-Austenitisierung-Schaubilder, Verlag Stahleisen, (1973)
[19] Kirkaldy, J. S. and Venugopalan, D. *Phase Transformations in Ferrous Alloys*, D.A.R. Marder and J.I. Goldstein, eds., AIME (1983)
[20] Victor Li, M., Niebuhr, V., Meekisho, L.L. and Atteridge D. G. A Computational Model for the Prediction of Steel Hardenability, *Metallurgical and Materials Transactions B* (1998) **29B**:661-672
[21] Zener, C. *Trans. AIME* (1946) **167**:550-83.
[22] Hillert, M. *Jernkont. Ann.*, (1957) **141**, 557-85.
[23] Koistinen, D. P. and Marburger, R. E. A General Equation Prescribing the Extent of the Austenite-Martensite Transformation in Pure Iron-Carbon Alloys and Plain Carbon Steels, *Acta Metallurgica* (1959) **7**:59-60
[24] Flux User Guide, www.cedrat.com

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis (Eds)

NUMERICAL METHODS IN SIMULATION OF RESISTANCE WELDING

CHRIS V. NIELSEN[1*], PAULO A.F. MARTINS[2], WENQI ZHANG[3] AND NIELS BAY[1†]

[1] Department of Mechanical Engineering, Technical University of Denmark
Produktionstorvet 425, 2800 Kgs. Lyngby, Denmark
[*]e-mail: cvni@mek.dtu.dk [†]e-mail: nbay@mek.dtu.dk

[2] Instituto Superior Tecnico, Universidade de Lisboa
Av. Rovisco Pais, 1049-001 Lisboa, Portugal
e-mail: pmartins@tecnico.ulisboa.pt

[3] SWANTEC Software and Engineering ApS
Diplomvej 373, 2800 Kgs. Lyngby, Denmark
e-mail: wz@swantec.com

Key words: Resistance Welding, Finite Element Method, Electro-Thermo-Mechanical, Phase Changes, Hardness, Damage.

Abstract. Finite element simulation of resistance welding requires coupling between mechanical, thermal and electrical models. This paper presents the numerical models and their couplings that are utilized in the computer program SORPAS. A mechanical model based on the irreducible flow formulation is utilized to simulate plastic deformation and the resulting distribution of stress, a thermal model based on transient heat transfer is used to determine the distribution of temperature, and a steady-state electrical model is employed to calculate the distribution of electrical potential and current density. From a resistance welding point of view, the most essential coupling between the above mentioned models is the heat generation by electrical current due to Joule heating. The interaction between multiple objects is another critical feature of the numerical simulation of resistance welding because it influences the contact area and the distribution of contact pressure. The numerical simulation of resistance welding is illustrated by a spot welding example that includes subsequent tensile shear testing.

1 INTRODUCTION

Resistance welding is a widely applied joining technology. Spot welding, one of its variants, is very important in automotive assembly lines as it is a robust and effective joining process that is under continuous development to meet new demands due to novel materials, complex welding geometries and quality requirements. Projection welding is another widely applied variant of resistance welding which benefits from natural or fabricated projections.

Numerical simulation methods applied to resistance welding can provide a better understanding of the joining process and be utilized in the definition of operating parameters and prediction of the final quality of the welds. The role of numerical simulation in resistance welding is stated well by Singh [1], who points out that simulation cannot replace or substitute ingenuity or creativeness, but it can help in gaining understanding of the process, and hence reduce the amount of time spent during development.

During the 1960s and 1970s, the first models considering resistance welding were exclusively focused on the temperature history and made use of 1D or 2D axisymmetric models without mechanical coupling. Early numerical simulations of resistance welding, starting in the 1980s, include finite element analysis by Nied [2] and application of the finite difference method by Cho and Cho [3]. The finite element method is more suited for the analysis of resistance welding due to the deformations involved in the process. Nied [2], who used the commercial program ANSYS, is recognized as the first to apply finite element analysis to resistance welding. He setup a rather complete model to study resistance spot welding that included electro-thermo-mechanical discretization of the electrodes and workpieces by solid elements and by surface elements at the interfaces between objects to account for elastic mechanical contact (Herzian contact) and for electrical and thermal contact resistances. Although the simulation only considered elastic deformation, it was the first ever to consider the developing contact areas that play a key role in the process. Furthermore, good agreement between presented experiments and simulations provided motivation and inspiration for others to continue developments of numerical methods applied to resistance welding.

Another example of finite elements applied to resistance welding is the numerical modelling of projection welding of an automotive door hinge with two projections welded to a sheet that was performed by Zhu et al. [4] using ANSYS with an electro-thermo-mechanical coupled model that included plastic deformation. Other examples are given by Ma and Murukawa [5].

The objective of this paper is to present an overview of the numerical methods utilized in the commercial finite element program SORPAS for the complete simulation of resistance welding. The program was developed in the 1990s at the Technical University of Denmark to solve 2D industrial applications and later commercialized, maintained and further developed by SWANTEC Software and Engineering [6, 7]. In contrast to general purpose finite element computer programs, SORPAS is a special purpose software dedicated to simulation and optimization of resistance welding processes. The 3D version of SORPAS has been recently developed in collaborative partnership between the Technical University of Denmark, the University of Lisbon and SWANTEC Software and Engineering and has been commercialized since 2012.

The paper is organized such that Section 2 includes basic numerical models and the necessary couplings in the electro-thermo-mechanical model. Section 3 presents the numerical contact implementation as well as the physical contact modeling that are crucial for the simulation of resistance welding. Section 4 describes the prediction of phase changes and resulting hardness related to the weld quality. Section 5 presents a resistance welding example showing the application of the numerical methods to process simulation and subsequent strength test simulation, and Section 6 concludes.

2 ELECTRO-THERMO-MECHNICAL NUMERICAL MODEL

This section presents the mechanical, thermal and electrical formulations that give support to the associated software modules and describes the necessary couplings. Coupled damage modeling is included as part of the mechanical module, while couplings between the three core modules are covered subsequently together with material dependencies.

2.1 Mechanical module

The mechanical module is based on the irreducible flow formulation, which takes the following weak variational form,

$$\int_V \bar{\sigma}\delta\dot{\bar{\varepsilon}}\,dV + K\int_V \dot{\varepsilon}_{ii}\delta\dot{\varepsilon}_{jj}\,dV - \int_{S_t} t_i\delta u_i\,dS + \delta\Pi_C = 0 \tag{1}$$

where the first term covers the energy rate due to plastic deformation in domain volume V, the second term imposes the incompressibility constraint, the third term applies surface tractions over surface S_t and the fourth term is due to the contact contribution to be described in Section 3.1. Variations with respect to velocities u_i are identified by δ while $\bar{\sigma}$ is the effective stress, $\dot{\bar{\varepsilon}}$ is the effective plastic strain rate, the penalty factor K is a large positive number, $\dot{\varepsilon}_{ii}$ is the volumetric strain rate, and t_i are prescribed surface tractions.

The plastic deformation is generally assumed to follow the isotropic von Mises yield criterion,

$$\bar{\sigma}^2 = 3J_2 \tag{2}$$

where J_2 is the second invariant of the deviatoric stress tensor σ'_{ij}. When simulating quasi-static mechanical strength tests the accumulation of damage is accounted for by utilizing constitutive equations of metallic materials with porosity. The formation of porosity is associated with generation and coalescence of voids in average terms over each element. The yield criterion,

$$\bar{\sigma}_R^2 = AJ_2 + BI_1 \tag{3}$$

where I_1 is the first invariant of the stress tensor σ_{ij}, is capable of handling the volumetric changes due to variations in relative density. The effective stress response $\bar{\sigma}_R$ for a given relative density R is given by $\bar{\sigma}_R^2 = C\bar{\sigma}^2$, where $\bar{\sigma}$ is the effective stress response of the fully dense material. The material constants A, B and C are dependent on the relative density and are assumed to follow the porous plasticity theory by Shima and Oyane [8]. The constant $C = \dfrac{R^5}{1+(2.49/3)^2(1-R)^{1.028}} \le 1$ is responsible for the decrease in flow stress when the relative density decreases ($R \le 1$). The accumulated damage D is formulated as,

$$D = 1 - R \tag{4}$$

This expression is zero for fully dense materials and increases linearly with decreasing relative density.

2.2 Thermal module

The thermal module is based on the classical Galerkin treatment of the heat transfer equation giving the following governing equation for the temperature T ,

$$\int_V kT_{,i}\delta T_{,i}\,dV + \int_V \rho_m c_m \dot{T}\delta T\,dV - \int_V \dot{q}_V \delta T\,dV - \int_S \dot{q}_S dS + \delta\Pi_T = 0 \tag{5}$$

where the first term is due to heat conduction, the second term is due to stored energy associated with a temperature rate $\dot{T}$, the third term includes the heat generation rate in the volume $\dot{q}_V$ and the fourth term covers the rate of heat generation (or loses) at the surface $\dot{q}_S$. The last term includes thermal contact between objects as described in Section 3.2.

In equation (5), k is the thermal conductivity, ρ_m is the mass density, c_m is the heat capacity, and δ is used for the arbitrary variations with respect to temperature.

The contributions to $\dot{q}_V$ stem from plastic work and Joule heating, as follows,

$$\begin{aligned} \dot{q}_{plastic} &= \beta\overline{\sigma}\dot{\overline{\varepsilon}} \\ \dot{q}_{Joule} &= \rho J^2 \end{aligned} \tag{6}$$

where β is the fraction of mechanical energy transformed into heat and is usually assumed to be in the range between 0.85 to 0.95, ρ is the electrical resistivity and J is the current density calculated in the electrical module (Section 2.3).

The rate of heat generation $\dot{q}_S$ along surfaces is due to friction, convection and radiation, where the latter two are heat loses during the welding process,

$$\begin{aligned} \dot{q}_{friction} &= \tau_f |v_r| \\ \dot{q}_{convection} &= -h\left(T_s - T_f\right) \\ \dot{q}_{radiation} &= -\varepsilon_{emis}\sigma_{SB}\left(T_s^4 - T_f^4\right) \end{aligned} \tag{7}$$

In the above equations, the heat due to friction is obtained from the product of the friction shear stress τ_f and the relative sliding speed v_r between two surfaces in contact. The convection is associated with the heat transfer coefficient h, the surface temperature T_s and the temperature T_f of the surroundings. The parameters expressing the radiation are the emissivity coefficient ε_{emis} and the Stefan-Boltzmann constant σ_{SB} .

2.3 Electrical module

The electric potential Φ is the major variable in the electrical module. The governing equation is the Laplace equation, which for an arbitrary variation of the electric potential $\delta\Phi$ and application of the divergence theorem can be written as,

$$\int_V \Phi_{,i}\delta\Phi_{,i}dV - \int_S \Phi_{,n}dS + \delta\Pi_\Phi = 0 \tag{8}$$

with the addition of the last term accounting for electrical contact between objects to be presented in Section 3.2. Equation (8) simplifies by the cancellation of the second term

because of the gradient of the potential along free surfaces $\Phi_{,n}$ being zero.

As it was early stated by Greenwood and Williamson [9], the electric field has a much faster reaction rate than the temperature field, and therefore the steady state approximation, $\dot{\Phi} = 0$, behind (8) is generally considered a very good approach. This means that the electric potential is determined solely by geometry.

The current density J is available from the ratio of the potential gradient and the electrical resistivity, $J_i = \Phi_{,i} / \rho$.

2.4 Electro-thermo-mechanical couplings

The three modules are coupled as schematically shown by Figure 1. The mechanical module is run at the beginning of each step to establish the velocity field and geometry change, the contact areas and the overall stress response. Besides the new geometry, the direct influences in other modules are the deformation heat (6) and friction generated heat (7) in the thermal module. Another important influence is on the electrical and thermal contact properties that depend on the contact stresses.

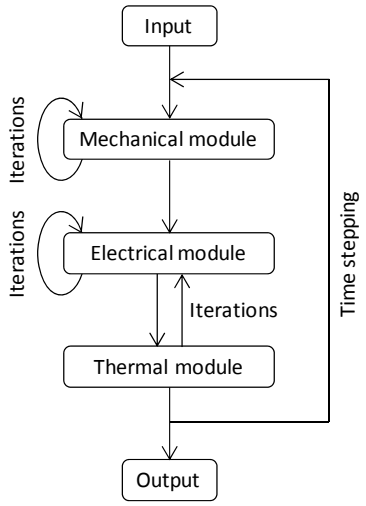


Figure 1: Numerical couplings between the electro-thermo-mechanical simulation of resistance welding.

After convergence of the mechanical module, the electrical and thermal modules are run until individual and mutual convergence. The output from the electrical module is the current density giving rise to Joule heating (6) characterizing the resistance welding process. The resulting temperatures in the thermal module are used to update all temperature dependent material properties. Among the most important influences by the material property updates are the mechanical softening, the increasing electrical resistivity and the changes in thermal properties themselves.

The coupling with the mechanical module is weaker than the coupling between the electrical and thermal modules. This is justified by simulating with very small time steps ensuring that the error in the mechanical module due to temperatures of the previous step is minimal. On the other hand, the time savings by weakening the coupling to the mechanical module is large. The coupling between the electrical and thermal modules is relatively

cheaper due to having only one third (in 3D) of the degrees of freedom for the two scalar fields (potential and temperature) and due to the linear and inexpensive solution of the electrical module. The limitation to using very small time steps is natural in resistance welding in order to capture all physical effects of the welding process. When, for example, simulating 50 Hz alternating current (AC), each half period corresponds to 10 ms and a minimum number of ten simulation points along each half period therefore requires time steps of maximum 1 ms. The time step should preferably be even smaller, especially when noting that alternating current in resistance welding machines typically has a conduction angle ranging from 50-90% (meaning that only 50-90% of the half period is active, and hence the time step should also be scaled to 50-90% to have the same resolution of the active current profile). Further details of the numerical implementation can be found elsewhere [10].

3 CONTACT

The contact implementation is based on the penalty method with identification of contact pairs by a two-pass node-to-face algorithm. The faces in the contact pairs are quadrilateral surface elements of the hexahedral volume meshes of the objects. The normal directions in the contact pairs are determined uniquely by one of the four triangles appearing when introducing a temporary center node in the quadrilateral surface element. The choice of the triangle is determined by the projection point of the contacting node. This method was also adopted by Doghri et al. [11] to overcome symmetry loss that appear with the alternative, and not unique, division of the quadrilateral surface into two triangles by one of the two diagonals.

3.1 Mechanical contact

The last term in Equation (1) due to mechanical contact by the penalty method is given by

$$\delta\Pi_C = P\sum_{c=1}^{N_c} g_n^c \delta g_n^c + P\sum_{c=1}^{N_c} g_t^c \delta g_t^c \tag{9}$$

which selectively penalizes normal gap velocities g_n^c by the first term and tangential gap velocities g_t^c by the second term. The penalty factor P is a large positive number applied to N_c contact pairs. The first term is selective in the sense that it is generally active when otherwise leading to penetration and inactive when the two contacting surfaces are separating. However, in contact pairs identified as already welded, the term is always active. The second term is likewise active in welded contact pairs and also when simulating full sticking conditions. During frictionless or frictional sliding, the second term is inactive.

The frictional stress τ_f during frictional sliding is applied in the contact pairs as surface tractions by the third term in (1) and is modeled by a combination of the Amonton-Coulomb law, $\tau_f = \mu\sigma_n$, and the law of constant friction, $\tau_f = mk$ where μ is the friction coefficient, m is the friction factor, σ_n is the normal pressure and k is the shear flow stress. Using the von Mises yield criterion, the flow stress is $\sqrt{3}k$, and the relation $\mu = m/\left(\sqrt{3}a\right)$ between the two friction laws is therefore valid when the transition between the two models take place at a normal pressure normalized by the flow stress $a = \sigma_n/\left(\sqrt{3}k\right)$, where a is typically 1.5 to 3.

The frictional stress is applied in the direction opposite to the relative sliding velocity $\mathbf{v}_r$ through the following expression by Chen and Kobayashi [12],

$$\tau = -\frac{2}{\pi}\tau_f \cdot \arctan\left(\frac{\mathbf{v}_r}{v_0}\right) \tag{10}$$

where the 'arctan' function is introduced to ensure a continuous derivative for the finite element implementation. It resembles the direction sufficiently fine when v_0 is an arbitrary constant much smaller than the relative speed. The surface integration of (10) over each contact element face is performed by 5x5 Gauss quadrature following Barata Marques and Martins [13] who applied the procedure to contact between finite elements and rigid tools.

3.2 Electrical and thermal contact

The numerical contact ensuring the same potential and temperature on both sides of contact interfaces are described by the following two penalty terms,

$$\begin{aligned} \delta\Pi_\Phi &= P\sum_{c=1}^{N_c}\Phi_d^c\delta\Phi_d^c \\ \delta\Pi_T &= P\sum_{c=1}^{N_c}T_d^c\delta T_d^c \end{aligned} \tag{11}$$

by penalization of potential difference Φ_d^c and temperature difference T_d^c in the N_c contact pairs. The two terms are added in equations (8) and (5), respectively.

Physical contact in the electrical and thermal modules is included through thin layers of elements on the object surfaces. The electrical contact resistivity is higher than the bulk resistivities of the two contacting surfaces due to current restriction and eventual surface contaminants such as oil, dirt and grease. The contact resistivity in SORPAS is modeled as follows [7],

$$\rho_c = \frac{3\sigma_{soft}}{\sigma_n}\left(\frac{\rho_1+\rho_2}{2}+\rho_{contaminant}\right) \tag{12}$$

with the fraction of real contact area to the apparent area in front of the parenthesis being expressed through the flow stress of the softer material σ_{soft} and the contact normal pressure σ_n, as also performed by Bowden and Tabor [14]. The fraction of real contact area describes, in conjunction with the two bulk resistivities of the materials in contact ρ_1 and ρ_2, the overall current restriction. The additional resistance due to contaminants is included through the resistivity $\rho_{contaminant}$. The high electrical contact resistivity expressed by (12) is one of the main contributions to the heat generation by Joule heating as it enters in Equation (6).

Thermal contact resistivity is formulated in a similar way with the exception that the thermal resistance of the surface contaminants is considered negligible. Hence, the thermal contact resistance κ_c is expressed through the ratio of real contact area to the apparent area and the average thermal resistance (inverse conductivity) of the two materials, κ_1 and κ_2,

$$\kappa_c = \frac{3\sigma_{soft}}{\sigma_n}\frac{\kappa_1 + \kappa_2}{2} \tag{13}$$

4 PHASE CHANGES AND HARDNESS PREDICTION

The phase changes during resistance welding involve melting and solidification as well as metallurgical changes of the solid materials due to microstructural changes. Besides the changing material properties of the solid or molten material, the latent heat L also needs to be taken into account when simulating in the temperature range between the solidus temperature T_{sol} and the liquidus temperature T_{liq}. The latent heat is included by replacing the heat capacity in (5) by an effective heat capacity $\widetilde{c}_m$ [15],

$$\widetilde{c}_m = c_m + \frac{L}{T_{liq} - T_{sol}} \tag{14}$$

Another effect of melting that needs to be taken into account is volume mixing of material properties when molten materials of different objects contribute to a molten volume.

As regards microstructural phase changes, the following is based on typical automotive steel grades. Austenitization is considered during heating without consideration of the temperature rate such that zero austenite is formed below the Ac1 temperature, 100% austenitization is assumed above the Ac3 temperature and linear interpolation is assumed between the Ac1 and the Ac3 temperatures. Formation of subsequent phases upon cooling takes the formed austenite as the starting point and is then calculated based on critical cooling rates as defined in the continuous cooling transformation (CCT) diagram of Figure 2. The critical cooling rates v_M for formation of martensite, v_B for formation of bainaite and v_{FP} for formation of ferrite/pearlite are calculated from the chemical compositions by the formulas presented by Blondeau et al. [16].

The hardness of each of the phases are calculated by the formulas given in the work by Maynier et al. [17] based on the chemical compositions and the actual cooling rates calculated in the finite element simulations. Having calculated the fractions of each of the phases together with their hardness, the total hardness of the material is calculated by applying volume mixing.

The change in hardness due to the temperature history is taken into account by scaling the flow stress curves approximated by $\bar{\sigma} = C(\varepsilon_0 + \bar{\varepsilon})^n$, where the pre-strain ε_0 and strain hardening exponent n of the original flow stress curve are kept constant while scaling the factor C. The scaling is performed with the objective of obeying the new tensile strength that can be approximated based on the new hardness. A number of simplified analytical approaches as well as empirical relations between the tensile strength and the hardness are available in literature, e.g. [18-21]. The often applied relation, $\sigma_{TS} = 3H_V$, is utilized in SORPAS, where σ_{TS} is the nominal tensile stress at the instability point corresponding to an effective strain $\bar{\varepsilon}_{inst} = n - \varepsilon_0$ and H_V is the Vickers hardness, which is calculated in the finite element simulation as explained above. Introducing the relation between the flow stress curve

(expressing the true stress) and the nominal tensile stress results in the following relation between the (true) flow stress at instability and the hardness,

$$\sigma_{inst} = 3e^{n-\varepsilon_0} H_V \tag{15}$$

which needs to be obeyed for the corresponding effective strain $n-\varepsilon_0$, and hence the new C-value becomes,

$$C = 3e^{n-\varepsilon_0} H_V / n^n \tag{16}$$

The new flow stress curves due to the calculated hardness distribution influence the simulation of strength testing.

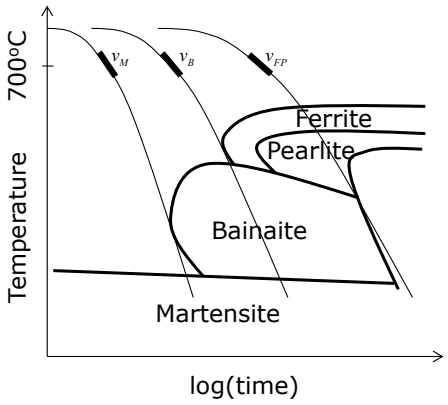


Figure 2: Schematic CTT diagram with indication of critical cooling rates at 700°C for formation of martensite, bainaite and ferrite/pearlite.

5 RESISTANCE SPOT WELDING AND TENSILE SHEAR STRENGTH TESTING

This section presents an example of a resistance spot welding experiment that includes subsequent tensile shear testing up to failure. The simulation is compared to the corresponding experiment to show the capabilities of the finite element implementation in prediction of overall morphology and quality of the weld. Figure 3 includes sheet dimensions, sheet material, electrode type, welding parameters and tensile shear testing speed and direction. The tensile shear testing is in compliance with the ISO standard 14273:2000(E). The initial finite element mesh utilizing one symmetry plane is shown in Figure 3, where the thin layers of elements on both sides of the sheets take into account the coating and interface properties. The electrodes are modeled during the welding simulation but are automatically removed before the tensile shear test simulation.

Experimental and simulated results are included in Figures 4 and 5. The simulated weld nugget diameter is 6.8mm (Figure 4a at the end of the welding time) in comparison with the experimentally observed diameter 6.9mm (Figure 4b). The final shape after tensile shear testing can be compared in Figure 4b and Figure 4c, where the latter includes the simulated damage according to Equation (4). The agreement with the located crack initiation observed in the experiment is good. Both simulation and experiments show full plug failure in the tensile shear tests. Figure 4e shows photographs of the spot weld after failure in one of the

experiments. Figure 4d shows the simulated hardness Vickers that is influencing the tensile shear test simulation by the scaling of the flow stress curves according to Equation (16).

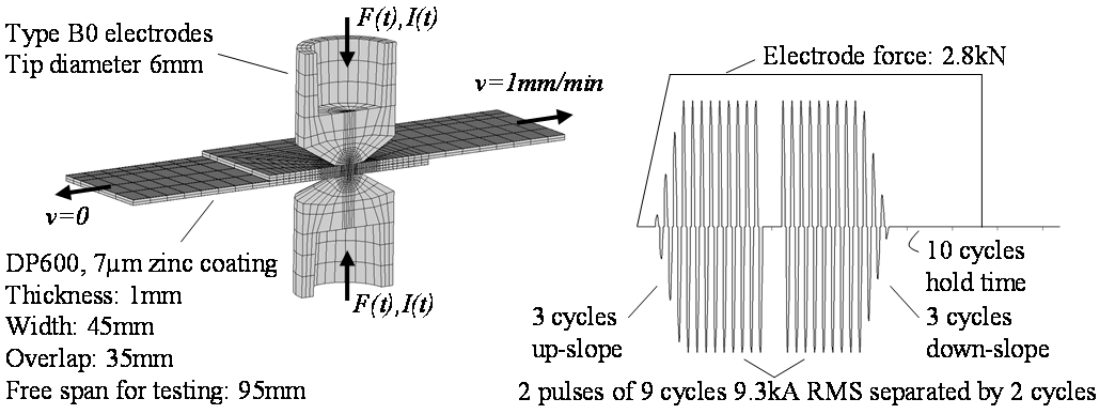


Figure 3: Resistance spot welding case represented by the initial finite element mesh and the welding parameters in terms of electrode force F and welding current I as function of process time t. Subsequent tensile testing velocity is denoted v.

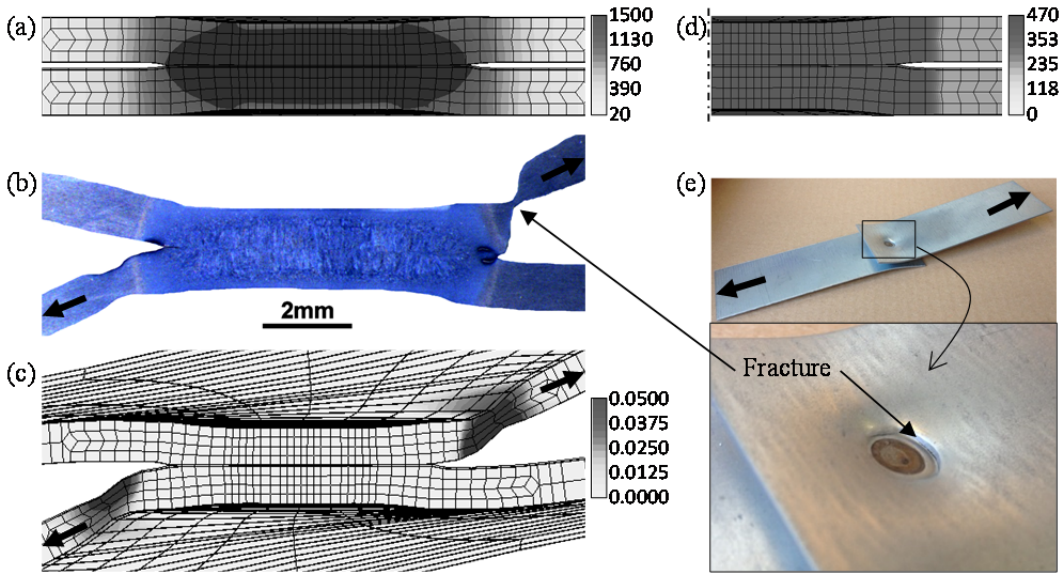


Figure 4: Simulation and experiments. (a) Simulated peak process temperature [°C] during welding and resulting weld nugget. (b) Experimental cross-section after welding and tensile shear testing. (c) Simulated deformation and damage resulting from tensile shear testing. (d) Simulated hardness Vickers distribution. (e) Photographs of spot weld after tensile shear testing until failure. (b,c,e) Arrows identify loaded sheets.

Figure 5 shows the final results of the tensile shear test simulation together with the results of five repetitions of the same experiment. The level of maximum load is captured by the implemented damage model, which is of most industrial relevance, while the sudden drop in load due to fracture does not appear in the simulation due to absence of crack propagation in the finite element model. Good agreement between the simulation and the experiments are observed in the prediction of the tensile shear strength.

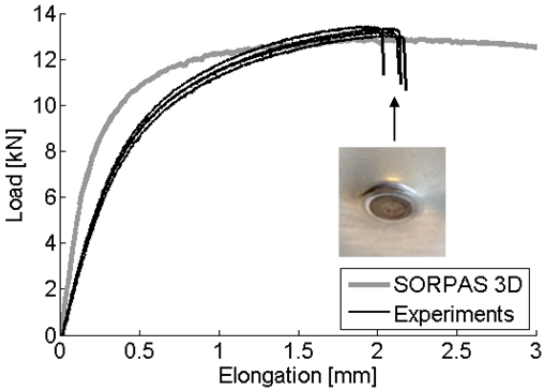


Figure 5: Comparison of simulation and experiments by load-elongation curves for tensile shear testing [22].

6 CONCLUSIONS

The individual modules of the electro-thermo-mechanical finite element implementation in SORPAS 3D and the necessary couplings as well as the level of coupling were comprehensively described. Accurate modeling of contact, phase changes, hardness prediction and resulting new strength of the material in the welded zone allows simulation to be extended into the post-welding destructive tests that are commonly employed to evaluate the overall quality and strength of the resulting weld nuggets.

Comparisons of the weld nugget size, weld strength and failure mode in a test case consisting of a single spot weld of two high strength steels are included to show the overall good agreement between experimental values and observations and numerical predictions.

REFERENCES

[1] Singh, S. Can simulation of the welding process help advance the state of the art in resistance welding? Proceedings of the 3rd International Seminar on: Advances in Resistance Welding, Berlin, Germany, (2012):5-11.

[2] Nied, H.A. The finite element modeling of the resistance spot welding process. Welding Research Supplement (1984) 63(4):123-132.

[3] Cho, H.S. and Cho, Y.J. A study of the thermal behavior in resistance spot welds. Welding Journal (1989) 68(6):236s-244s.

[4] Zhu, W.-F., Lin, Z., Lai X.-M. and Luo, A.-H. Numerical analysis of projection welding on auto-body sheet metal using a coupled finite element method. International Journal of Advanced Manufacturing Technology (2006) 28(1-2):45-52.

[5] Ma, N. and Murukawa, H. Numerical and experimental study on nugget formation in resistance spot welding for three pieces of high strength steel sheets. Journal of Materials Processing Technology (2010) 210:2045-2052.

[6] Zhang, W. and Kristensen, L. Finite element modeling of resistance spot and projection welding processes. The 9th International Conference on Computer Technology in Welding, Detroit, Michigan, (1999):15-23.

[7] Zhang, W. Design and implementation of software for resistance welding process

simulations. Trans. J. Mater. Manuf. (2003) 112(5):556–564.
[8] Shima, S. and Oyane, M. Plasticity theory for porous metals. International Journal of Mechanical Sciences (1976) 18(6):285-291.
[9] Greenwood, J.A. and Williamson, J.B.P. Electrical conduction in solids II. Theory of temperature-dependent conductors. In: Proceedings of the Royal Society of London, Series A, Mathematical and Physical Sciences (1958) 246:13-31.
[10] Nielsen, C.V., Zhang, W., Alves, L.M., Bay, N. and Martins, P.A.F. Modeling of Thermo-Electro-Mechanical Manufacturing Processes with Applications in Metal Forming and Resistance Welding, Springer, (2012).
[11] Doghri, I., Muller, A. and Taylor, R.L. A general three-dimensional contact procedure for implicit finite element codes. Engineering Computations (1998) 15(2):233-259.
[12] Chen, C. and Kobayashi, S. Rigid plastic finite element analysis of ring compression. Application of numerical methods in forming processes, AMD 28, ASME, New York, (1978):163-174.
[13] Barata Marques, M.J.M. and Martins, P.A.F. Three-dimensional finite element contact algorithm for metal forming. International Journal for Numerical Methods in Engineering (1990) 30:1341-1354.
[14] Bowden, F.P. and Tabor, D. The Fabrication and Lubrication of Solids, Oxford University Press, (1950).
[15] Zhang, W., Jensen, H.H. and Bay, N. Finite element modeling of spot welding similar and dissimilar metals. The 7th International Conference on Computer Technology in Welding, San Francisco, USA, (1997):364-373.
[16] Blondeau, R., Maynier, P., Dollet, J. and Vieillard-Baron, B. Prévision de la dureté, de la résistance et de la limite d'elasticité des aciers au carbone et faiblement alliés d'après leur composition et leur traitement thermique. Mémoires Scientifiques Revue Métallurgie (1975):759-769.
[17] Maynier, P., Jungmann, B. and Dollet, J. Creusot-Loire system for the prediction of the mechanical properties of low alloy steel products. Hardenability Concepts with Applications to Steels. The Metallurgical Society of AIME Heat Treatment Committee / American Society for Metals Activity on Phase Transformations (1978):518-545.
[18] Tabor, D. The physical meaning of indentation and scratch hardness. British Journal of Applied Physics (1956) 7:159-166.
[19] Cahoon J.R. An improved equation relating hardness to ultimate strength. Metallurgical Transactions (1972) 3:3040.
[20] Busby, J.T., Hash, M.C. and Was, G.S. The relationship between hardness and yield stress in irradiated austenitic and ferritic steels. Journal of Nuclear Materials (2005) 336:267-278.
[21] Pavlina, E.J. and Van Tyne, C.J. Correlation of yield strength and tensile strength with hardness for steels. Journal of Materials Engineering and Performance (2008) 17(6):888-893.
[22] Nielsen, C.V., Bennedbæk, R.A.K., Larsen, M.B., Bay, N., Chergui, A., Zhang, W. and Martins, P.A.F. Experimental and simulated strength of spot welds. The 8th International Seminar on Advances in Resistance Welding, Baveno, Italy, (2014): 161-172.

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis(Eds)

A COMPUTATIONAL MULTISCALE APPROACH TO COUPLE HYGRO-MECHANICAL RESPONSES OF LARGE-SCALE MASONRY WALLS

G. CASTELLAZZI*, S. de MIRANDA*, G. FORMICA†, L. MOLARI* AND F. UBERTINI*

*DICAM Department
University of Bologna
Viale del Risorgimento 2, 40136 Bologna, Italy
Web page: http://www.eng.dicam.unibo.it

†Dipartimento di Architettura
Università degli Studi "Roma Tre"
via della Madonna dei Monti 40, 00184 Roma, Italy
e-mail: formica@uniroma3.it, web page: http:/http://host.uniroma3.it/docenti/formica

Key words: Computational Multiscale approach, mechanical and environmental coupled degradations.

Abstract. We present a computational multiscale approach to the nonlinear problems of humidity diffusion and mechanical damage of large-scale masonry walls, and their coupling in terms of the effects of the humidity diffusion on the mechanical response and the effects of the mechanical degradation on the diffusion process. Such an approach allows us to recover, both efficiently and accurately, the complex nonlinear response of large-scale walls, which are in general hard to be solved by means of standard numerical tools.
Two representative tests of two- and three-storey walls are here analyzed, and the corresponding results reported and commented, aiming to show how samples like these can potentially serve as reference solutions for more applicative purposes.

1 INTRODUCTION

The models of the two nonlinear problems, humidity diffusion and structural equilibrium, consider the wall as a lattice-like discrete system [1, 2], where the field variables are associated to each block composing the wall, i.e. sandstone units, while the nonlinear response is assumed to be concentrated at interface elements, i.e. cement mortar joints. Although such models drastically simplify the corresponding physical phenomena, or more sophisticated models [3], they present all the relevant nonlinearities of the problems, such as the water uptake process under evaporation conditions, and the structural frictional

toughness along with the softening material response. In order to address to more realistic structural contexts, large-scale walls are hard to be solved numerically, even working with physical models as simplified as the discrete models here adopted [1]. Regardless of the computational costs which increase with the number of blocks in the masonry texture, the nonlinearities of the physical problems can lead to difficulties in convergence [4]. Such nonlinearities and consequent failures in convergence are essentially related to localization phenomena of the physical responses: roughly speaking, the water uptake process tends to concentrate the humidity distribution in sharp fronts, in reason of the simultaneous evaporation process; on the other hand, the strain localizes and produces mechanical damage, so that the wall stress response becomes softening.

Following the approach proposed in [5] for pure mechanical contexts and then extended to the diffusive context in [2], the discrete systems are solved by an iterative multilevel strategy mimicking algebraic multigrid methods, where local solutions, those of balancing between two linked blocks, and global solutions, those of balancing two contiguous patterns of blocks, work as nested sequences able to correct error distributions with high and low frequency oscillation, respectively.

We conduct a numerical testing, whose meaningful results are here reported. We essentially show how the diffusion properties are affected by existing damage patterns in equilibrium with certain loading conditions, and conversely, how strength and elastic properties are degraded by existing distributions of humidity through water uptake processes in steady-state. In particular, the results obtained on a two- and a three-storey wall are discussed and compared.

2 COMPUTATIONAL MULTISCALE APPROACH

2.1 A sketch of the hygro-mechanical model

Here we summarize modelling equations of both diffusive and mechanical problems, first proposed and validated in [2] and [1], respectively. The two physical problems are both modelled in a 2D context basing on a Lagrangean formulation which leads to describe the wall as a lattice-like discrete system.

Each brick is ideally divided in two blocks, where both humidity and displacement variable fields are described by discrete parameters which refer to the block centroid: one parameter generalizes the humidity scalar field over the block, while the kinematics vector field is generalized in terms of three parameters, namely two displacement components and the rotation of the block. Each block interacts with the n surrounding blocks through the corresponding (two bed plus two head) mortar joints (n is equal to 3 or 4, depending on whether the block is at the boundary or not). Such interactions are governed by constitutive prescriptions which allow us to model each mortar joint as an interface exhibiting the nonlinear response of both physical problems; in particular, the interface corresponding to one of the two head joints accounts for the nonlinear response that develops within the single real brick. Figure 1 sketches the key-points building up the discrete model:

(a) initial 3D model to be discretized; (b) 2D discrete model; (c) the block as physical Representative Elementary Volume (REV); kinematical (d) and humidity (g) descriptors; interface nonlinear responses in terms of stresses (e) and moisture flux (h), and corresponding tangent mechanical stiffness (f) and diffusive permeability (i) of the interfaces depicted as springs.

Referring to Figure 1 for definitions of geometrical and physical symbols, and denoting with h_i and $h_{(k)}$ the values of the humidity h at the centroids of the i-th block and of the k-th surrounding block, the discrete moisture-balance equation of each block reads as:

$$\overline{c}_{w,h}[h_i]\frac{\partial h_i}{\partial t} + \frac{1}{\Omega_b}\sum_{k=1}^{n}\frac{b_{(k)}}{l_{(k)}}\overline{C}_h[h_{(k)}]\Big(h_{(k)} - h_i\Big) - \overline{\gamma}_w[h_i]\Big(h_i - h_{\text{env}}\Big) = 0 \tag{1}$$

where barred quantities mean that the corresponding continuum quantities are homogenized over the REV, according to an averaging procedure of material and geometric properties of bricks and joints involved in the same REV. In particular, $\overline{c}_{w,h}$ is the equivalent ratio of moisture content c_w, $\overline{C}_{h(k)}$ is the equivalent interface permeability between two blocks, and $\overline{\gamma}_w$ is the humidity convective term accounting for the evaporation conditions across the wall thickness. Finally, h_{env} is the prescribed environmental humidity. Further details can be found in [3, 2].

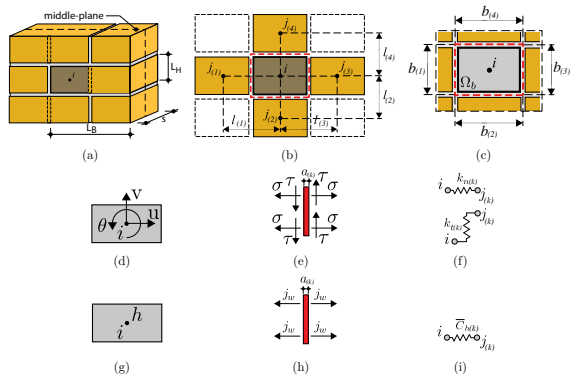


Figure 1: Representation of the generation of the discrete model.

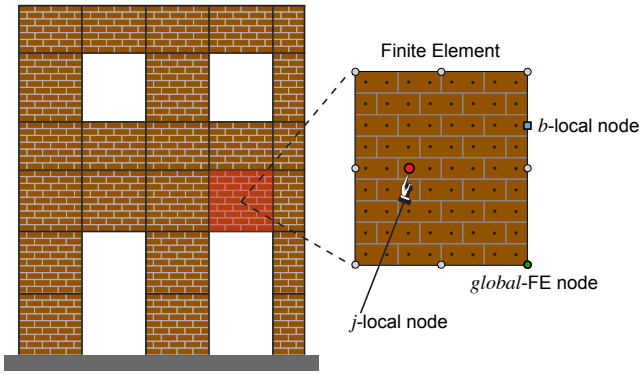


Figure 2: Representation of the different scales/levels.

The discrete equilibrium equations of each block are generated similarly: the second term in (1) assembles in the same way the stress contributions through the interfaces surrounding the block, while the last source term in (1) corresponds in the mechanical problem to possible loads referred to the block centroid; the first ratio-dependent term is instead not present. The equivalent, normal and shear, stress contributions (σ, τ) across the interfaces depend nonlinearly on the differential displacements (ε, γ) of the surrounding blocks with respect to each considered block [1]:

$$\sigma = \frac{k_{n(k)}\,\varepsilon}{1 + \alpha_{(k)}H(\varepsilon)}\,, \qquad \tau = \frac{k_{t(k)}\gamma\left\{1 + \alpha_{(k)}[1 - H(\varepsilon)]\right\} - \alpha_{(k)}\overline{\tau}_{(k)}H(-\varepsilon)}{1 + \alpha_{(k)}}. \tag{2}$$

In (2), $k_{n(k)}$ and $k_{t(k)}$ are, respectively, the normal and tangential stiffness associated to the k-th joint, $\alpha_{(k)}$ is the damage parameter of the k-th joint, and $\overline{\tau}_{(k)}$ is the frictional toughness equivalent to a fully-damaged joint. $\overline{\tau}_{(k)}$ obeys to the Coulomb criterion:

$$\overline{\tau}_{(k)} = k_{t(k)}\gamma - \max\left\{-\mu_{(k)}\sigma,\, \min[\mu_{(k)}\sigma,\, \tau_{0(k)} + k_{t(k)}(\gamma - \gamma_0)]\right\}, \tag{3}$$

being $\mu_{(k)}$ the friction coefficient of the mortar/brick interface, and $\tau_{0(k)}$ and γ_0 the current values of $\tau_{(k)}$ and γ. Moreover, H is the Heaviside function, so as to switch compressive to tensile conditions of the joint. Therefore, in compression the joint behaves as linearly elastic, except for a Coulomb frictional toughness $\overline{\tau}_{(k)}$.

2.2 Numerical strategy

The two analyses of humidity diffusion and mechanical response run separately, the former requiring a time-integration solution, the latter a parametric solution in the shear-load amplification factor. In particular, for the humidity diffusion problem we use the generalized-alpha method, as defined in [6]; a path-following technique recovers the pseudo-static equilibrium paths of the mechanical problem, according to the scheme first proposed in [1] and then developed in [5].

The two physics are coupled just in the initialization phase of the processes, where the material diffusion parameters are modified in dependence on a known damage pattern, and the material mechanical coefficients in dependence on a given humidity distribution. From a computational point of view, both analyses require iterative solvers which actually reveal the same complexity, sharing several parts of the algorithmic frameworks.

Indeed, both the iterative solution schemes solve the linearized problem serving as corrector step of an initial guess. Let $\boldsymbol{u}_l$ be the vector of initial guess: $\boldsymbol{u}_l$ collects the discrete parameters that correspond to the values of humidity of all the sandstone blocks, in a case, and the displacements of all the sandstone blocks, in the other case. The corrector step within the iterative strategies solves the following residual problem:

$$\boldsymbol{r}_l\left[\delta\boldsymbol{u}_l\right] = \boldsymbol{f}_l - \boldsymbol{s}\left[\delta\boldsymbol{u}_l\right] = \boldsymbol{0}\,, \quad \text{with } \boldsymbol{s}\left[\delta\boldsymbol{u}_l\right] = \boldsymbol{K}\left[\boldsymbol{u}_l\right]\delta\boldsymbol{u}_l \tag{4}$$

where $\boldsymbol{r}_l$ is the residue, $\boldsymbol{f}_l$ is the known term, and $\boldsymbol{s}\left[\delta\boldsymbol{u}_l\right]$ is the linearized physical response: $\boldsymbol{K}\left[\boldsymbol{u}_l\right]$ is the tangent matrix of the problem computed at the current solution $\boldsymbol{u}_l$, and $\delta\boldsymbol{u}_l$ is the correction variable to find that updates $\boldsymbol{u}_l$; $\boldsymbol{K}\left[\boldsymbol{u}_l\right]$ serves as iteration matrix.

As proved in [5] and then in [2], computational multiscale approaches can successfully avoid problems in convergence occurring when analyzing nonlinear responses of large-scale walls. We distinguish two scales of representation (see Figure 2), i.e. the (reference) *local* scale, defined at each block by the discrete models, and the *global* scale, defined as Finite Elements which are a representation of the same local problems for a certain pattern of blocks. Within an iterative scheme that reaches convergence at the reference scale, the two scales serve both as partial solutions of the whole problem: the global scale is used as a pre-conditioner of the local scale and corrects smooth distributions of the iterative

error, while the local scale is used to solve the interaction between two blocks at a time and then corrects high-frequency oscillations of distribution of the error.

Introducing for vectors and matrices subscripts l and g to label local and global quantities, respectively, the global-level sequence of algorithmic operations starts from a trial solution $\boldsymbol{u}_l$ affected by the error $\boldsymbol{r}_l$, and consists of the following solutions:

$$\boldsymbol{r}_g = \boldsymbol{A}^\top \boldsymbol{r}_l \quad \Rightarrow \quad \boldsymbol{d}_g = \boldsymbol{K}_g^{-1} \boldsymbol{r}_g \quad \Rightarrow \quad \boldsymbol{u}_l^{\text{new}} = \boldsymbol{u}_l + \boldsymbol{A}\boldsymbol{d}_g \tag{5}$$

In such a sequence, the first and latter items use the so-called scale bridging operator $\boldsymbol{A}$, to perform local-to-global and global-to-local transformations, respectively: in particular, $\boldsymbol{A}^\top$ works as restriction of the local error to the global (coarser) error, while $\boldsymbol{A}$ interpolate the global correction into a (finer) local correction which update the initial local solution. Since this updated solution is usually affected by a new, high frequency, error $\boldsymbol{r}_l$, a smoothing process works at local level, according to the following equations:

$$\boldsymbol{d}_l = \boldsymbol{K}_l^{-1} \boldsymbol{r}_l \quad \Rightarrow \quad \boldsymbol{u}_l^{\text{new}} = \boldsymbol{u}_l + \boldsymbol{d}_l \tag{6}$$

Note that the local-level iteration (6) is enabled on the basis of a threshold criterion, which is imposed on the global error e_g left by the global-level solutions (5); on the other hand, the global level can run when the local error e_l, resulting from local-level corrections, reduce by a factor $c < 1$ the global error e_g, ie. $e_l < c\, e_g$.

The key-role is actually ruled by the scale bridging operator $\boldsymbol{A}$, which is derived numerically for each Finite Element simultaneously. The algorithm, implementing a sort of algebraic homogenization procedure, results from a linearized solution computed at the level of the single element: in one stage the linearized balance equations are solved in terms of the discrete variables of the blocks surrounding the element boundary; these boundary variables are then imposed to be perfectly conform to the shape functions describing the same variables in terms of the element at the global scale. This way to proceed minimizes the numerical error introduced while smoothing the local evaluation of the variables. Besides, it allows us to derive the global tangent matrix $\boldsymbol{K}_g$, which turns out to be the best coarse representation of the (finer) tangent matrix of the local problem $\boldsymbol{K}[\boldsymbol{u}_l]$ when the local problem is linear [4].

Consider the local linearized balance, which represents a general form valid for both the mechanical and the diffusive solution scheme, as the following system,

$$\begin{bmatrix} \boldsymbol{K}_l^{jj} & \boldsymbol{K}_l^{jb} \\ \left(\boldsymbol{K}_l^{jb}\right)^\top & \boldsymbol{K}_l^{bb} \end{bmatrix} \begin{bmatrix} \boldsymbol{d}_l^j \\ \boldsymbol{d}_l^b \end{bmatrix} = \begin{bmatrix} \boldsymbol{r}_l^j \\ \boldsymbol{r}_l^b \end{bmatrix}. \tag{7}$$

Here we assume a local linearized balance for each Finite Element, i.e. the global scale, containing a certain pattern of bricks, which, in turn, contains the local scale, separating brick local variables internal to the Element (labeled with superscript j) from those at the boundary of the same Element (labeled with superscript b), see Figure 2. The Finite Element size is defined by the user and, of course, must be a multiple of the brick size.

Therefore, solving by condensation in the local problem (7) the internal variables $\boldsymbol{d}_l^j$ in terms of the boundary variables $\boldsymbol{d}_l^b$, and enforcing $\boldsymbol{d}_l^b$ to be exactly conform to the interpolation used to define the Finite Element representing the global scale, i.e. $\boldsymbol{d}_l^b = \boldsymbol{N}\boldsymbol{d}_g$, we finally obtain the scale bridging matrix for the generic Element as

$$\boldsymbol{A} = \left(\boldsymbol{K}_l^{jj}\right)^{-1}\boldsymbol{K}_l^{jb}\boldsymbol{N}, \tag{8}$$

where $\boldsymbol{N}$ is the matrix collecting the shape functions of each Element, evaluated at the centroids of the boundary bricks. In particular, a 8-node Serendipity interpolation is used as shape functions.
The global tangent matrix is then picked up by the remaining internal variables of the same local problem (7): by imposing a mutual work identity between the two scales, i.e. $\boldsymbol{r}_l^\top \boldsymbol{d}_l = \boldsymbol{r}_g^\top \boldsymbol{d}_g$, for each Element we have

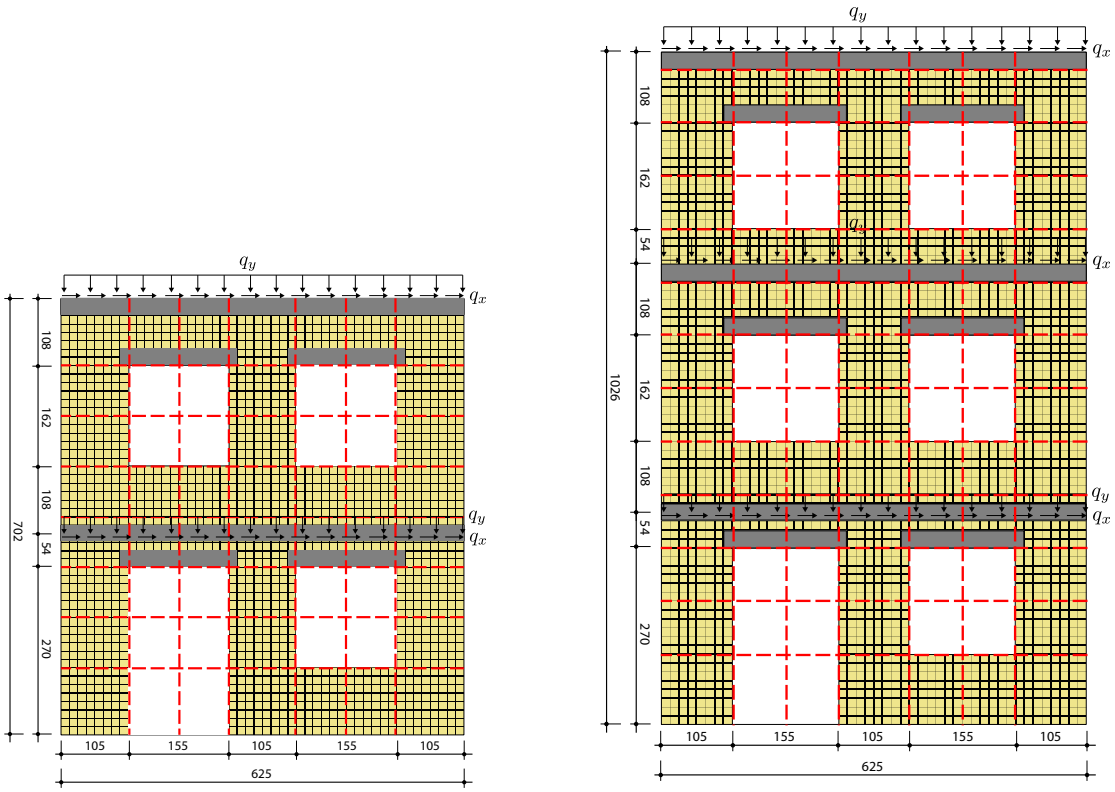
$$\boldsymbol{K}_g = \boldsymbol{N}^\top \left[\left(\boldsymbol{K}_l^{bb}\right)^{-1} - \left(\boldsymbol{K}_l^{jb}\right)^\top \left(\boldsymbol{K}_l^{jj}\right)^{-1}\boldsymbol{K}_l^{jb}\right]\boldsymbol{N}. \tag{9}$$

Each Element is therefore balanced internally, i.e. the local linearized balance problem is satisfied at each element separately, while it is compatible along its boundary with a local solution. It means that if the local solution is exactly the one predicted by the interpolating Serendipity functions at each Element boundary, then the global level reach the exact linearized solution everywhere. As concerns $\boldsymbol{K}_l$, since it has to be able to correct local fluctuations of the error, it corresponds to the tangent matrix of the problem defined at a single block. Thus, the correction of the local variables is made block by block, regardless of the influence of the blocks surrounding it, as well as a Gauss-Seidel iterative scheme.

3 NUMERICAL RESULTS

The validation of the discrete model in terms of pure mechanical response has been carried out in [5], considering applications at the structural scale as well. The validation of the discrete diffusive model has been made in [2] by comparing the two-dimensional discrete model with the three-dimensional continuum diffusive model, already proposed and tuned on experimental evidences in [3].

In this section the proposed numerical strategy is applied to two significant large-scale walls, characterized by different aspect ratios (height-over-width) of the wall size: the former about 1, the latter about 3/2. Two masonry walls of two and three storeys respectively were tested, whose geometries are depicted in Figure 3.
For both the walls we report: i) equilibrium curves providing shear-load carrying capacity, see Figure 4; ii) uptake curves, measuring the cumulative volumetric water inflow up to steady-state conditions, see Figure 5; iii) damage maps depicting in white-to-red color scale joints with different values of damage (red color to 100% damaged joints), overlapped with humidity maps depicting in white-to-aqua scale blocks with different values of water content (aqua color to 100% wet blocks), see Figures 6 and 7 .



(a) Two-storey wall (b) Three-storey wall

Figure 3: Geometry of the tested walls (units in centimetres). FE mesh in red dashed lines.

3.1 Material properties and coupling relationships

Isothermal conditions are assumed, with the temperature $T = 293K$. The material properties of the sandstone and cement mortar come from various experimental evidences (see [1]–[5] for further details and references), and they are listed in Table 1. Note that we adopt the following empirical expressions, as proposed in [7], for adsorption isotherm and for liquid conductivity:

$$c_w = \phi\rho_w \frac{\psi - 1}{\psi - h} h\,, \quad K_l = 3.8 \left(\frac{A}{\phi\rho_w}\right)^2 10^{3(S_w - 1)} \frac{\partial c_w}{\partial h} \frac{h}{\rho_w R_v T}, \tag{10}$$

where A is the water adsorption coefficient, R_v the gas constant of water vapour, ρ_w the water density, and the parameter $\psi = 0.8 \dfrac{c_{w80} - \phi\rho_w^l}{c_{w80} - \phi\rho_w^l 0.8}$ fits the experimental data by enforcing the measured value of water content at relative humidity $h = 0.8$.

The coupling phenomena are here modelled by constitutive relationships investing key material quantities: i) the sandstone elastic (Young) modulus and the mortar tensile strength, both decreasing as the water saturation degree increases; ii) the mortar perme-

Table 1: Material property parameters.

meule sandstone	parameter	cement mortar
$A = 0.225$	water adsorption [kg/m^2/s$^{1/2}$]	$A = 0.015$
$\phi = 0.235$	porosity	$\phi = 0.23$
$\rho = 2077$	density [kg/m^3]	$\rho = 2053$
$c_{w80} = 8.31$	water content at $h = 0.8$ [kg/m^3]	$c_{w80} = 59.3$
$\frac{D_v}{R_v T} = 4.401\text{e-}11$	vapour permeability [kg/s/Pa/m]	$\frac{D_v}{R_v T} = 1.965\text{e-}11$
$E = 19$	Young's modulus [GPa]	$E = 0.5$
$G = 3.8$	tangential modulus [GPa]	$G = 0.1$
$\sigma_t = 40$	tensile strength [MPa]	$\sigma_t = 21$
$\mu = 0.467$	friction coefficient	$\mu = 0.467$

ability which increases as the mechanical damage increases.
The former relationships rule how mechanical characteristics are affected by the presence of humidity in the wall, and they are expressed in terms of the water saturation degree $S_w = c_w \big/ \rho_w \phi$. The sandstone elastic modulus obeys to the following law

$$E = 3.121 + 19.06 \Big/ S_w 100 \text{ [GPa]} \tag{11}$$

according to experimental results reported and interpreted in [8]. On the basis of the same work, the decay law of the cement tensile strength was derived and proposed in [2]:

$$\sigma_t = -7 \left((S_w - 10) \big/ 100 \right)^{1/5} + 21 \text{ [MPa]} \tag{12}$$

Such a law is valid for values of S_w greater than 10%, while the constant value $\sigma_t = 21$ [MPa] is considered for values lower than 10%.
The influence of the mechanical damage on the mortar permeability is supported by more recent experiments [9, 10]. In particular, the water vapour diffusion C_h^v and the liquid water diffusion C_h^l, which define the mortar permeability $C_h = C_h^v + C_h^l$, are scaled by two factors g_d^v and g_d^l, respectively, which both depend on the percentage of damaged mortar area $\beta = \alpha \Big/ (1 + \alpha)$, according to the following laws:

$$g_d^v = -1.875\beta^2 + 1.5\beta + 1 \,, \quad g_d^l = 1 + (b_1\beta)^{b_2} + 1/2(b_1\beta)^{2b_2} + 1/6(b_1\beta)^{3b_2} \tag{13}$$

where $b_1 = 11.3$ and $b_2 = 1.64$ are the coefficient values tuned on concrete materials.

3.2 Effects of moisture rising on the structural response

The structural response and carrying capacity are determined by performing push-over like analysis: once humidity diffusion gets the steady-state and preliminary gravity loading are applied, shear loading are monotonically increased according to an amplification

factor, up to the ultimate condition of the structure. The equilibrium path during this varying loading condition is recovered in terms of such an amplification factor and the average lateral displacement of the wall.
Figure 4 reports the equilibrium paths followed by the two tested walls, so as to compare structural performance in dry conditions with respect to those in wet conditions, for both the walls. Such wet conditions are attained by conducting a diffusive analysis as we will describe in the next Section. Both structural responses highlight how the presence of humidity rising on the wall significantly affect the pre- and post-critical mechanical behaviours. Less evident is the influence of different diffusive conditions when the ultimate shear load is attained. Such conditions indeed affect only the softening response of the stress according to how the damage can diversely evolve and distribute; the limit load is instead characterized almost by the friction toughness which is predominant on the structural response as the damage becomes greater and greater.

By comparing the equilibrium paths in Figure 4, the two-storey wall is more sensitive than the three-store wall to the different damage distributions passing from dry to wet case. This result can be explained by considering that the shear limit conditions are expected to be more evident in lower than higher walls.
In order to clarify this point two-by-two snapshots of damage maps over the structure are reported in Figures 6 and 7. Dry and wet cases have quite similar spatial distribution over the two walls. However, damage can grow along the left side more significantly in the three-storey wall compared to that of the two-storey wall, since the former wall is subject to a mechanical behaviour more flexural than shear, with respect to the latter wall.

3.3 Effects of damage level on moisture rising

The humidity diffusion analysis is conducted by simulating moisture rising from ground, with evaporating conditions either applied as boundary Neumann conditions on all the external surfaces, and applied in the entire wall domain as source term.
As the curves of volumetric water uptake show (see Figure 5), the presence of a damage pattern increases noticeably the water content evolving in time. For both the walls, such curves describe a qualitatively similar increment at steady-state conditions, as expected by the fact that the moisture rising is imposed from the ground of both the walls, and the damage pattern is mostly concentrated in the lower parts. However, the increment of the volumetric water content in the lowest wall doesn't exceed a value of 8%, while in the highest wall it attains about a 20%. This result can be explained in reason of a wider distribution of both damage and humidity in the three-storey wall, for the same inflow surface area at the ground.

The differences in humidity distribution over the walls are also emphasized in Figures 6 and 7. Due to the different mechanical conditions (undamaged versus damaged wall), the humidity significantly rises up to the first floor in the initially damaged wall for both the analyzed cases. In particular, both the walls are characterized by a very similar moisture rising in the undamaged cases (see figures on the left), owing to the identical boundary

conditions imposed in the diffusion problems. However, the damage distribution in the three-storey wall, which has a more pronounced flexural response with respect to the two-storey wall, leads the moisture rising to assume a more evident straight front, while the steady-state front attained in the two-storey wall follows more evidently diagonal, shear-like, lines.

REFERENCES

[1] Formica G, Sansalone V, Casciaro R. A mixed solution strategy for the non linear analysis of brick masonry walls. *Comput Method Appl M* (2002) **191**:5847–5876.

[2] Castellazzi G, de Miranda S, Formica G, Molari L, Ubertini F. Coupled hygro-mechanical multiscale analysis of masonry walls. *Eng Struct* (2015) **84**:266–278.

[3] Castellazzi G, Colla C, de Miranda S, Formica G, Gabrielli E, Molari L, Ubertini F. A coupled multiphase model for hygrothermal analysis of masonry structures and prediction of stress induced by salt crystalization. *Constr Build Mater* (2013) **41**:717–731.

[4] Brasile S, Casciaro R, Formica G. Multilevel approach for brick masonry walls - Part II: On the use of equivalent continua. *Comput Method Appl M* (2007) **196**:4801–4810.

[5] Brasile S, Casciaro R, Formica G. Multilevel approach for brick masonry walls - Part I: A numerical strategy for the nonlinear analysis. *Comput Method Appl M* (2007) **196**:4934–4951.

[6] Jansen KE, Whiting CH, Hulbert GM. A generalized-α method for integrating the filtered Navier-Stokes equations with a stabilized finite element method. *Comput Method Appl M* (2000) **190**:305–319.

[7] Sykora J, Krejci T, Sejnoha M. Computational homogenization of non-stationary transport processes in masonry structures. *J Comput Appl Math* (2012) **236**: 4745–4755.

[8] Glucklich J, Korin U. Effect of Moisture Content on the strength and strain energy release rate of cement mortar. *J Am Ceram Soc* (1975) **58**:517–521.

[9] Pijaudier-Cabot G, Dufour F, Choinska M. Permeability due to the increase of damage in concrete: from diffuse to localized damage distribution. *J Eng Mech-ASCE* (2009) **135**:1022–1028.

[10] Rouchier S, Foray G, Woloszyn M, Roux JJ. Influence of diffuse damage on the water vapour permeability of fibre reinforced mortar. *Transport Porous Med* (2012) **93**:543-559.

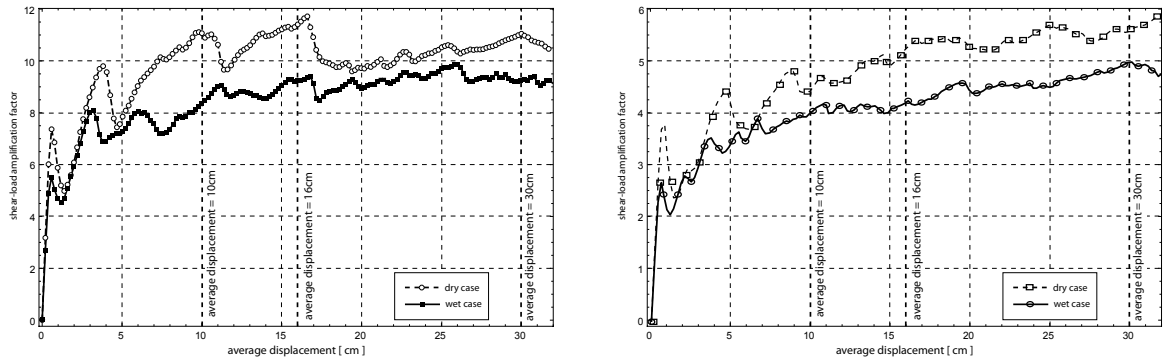


Figure 4: Structural responses in the presence of moisture (wet case) compared with the dry case: two-storey wall on the left, three-storey wall on the right.

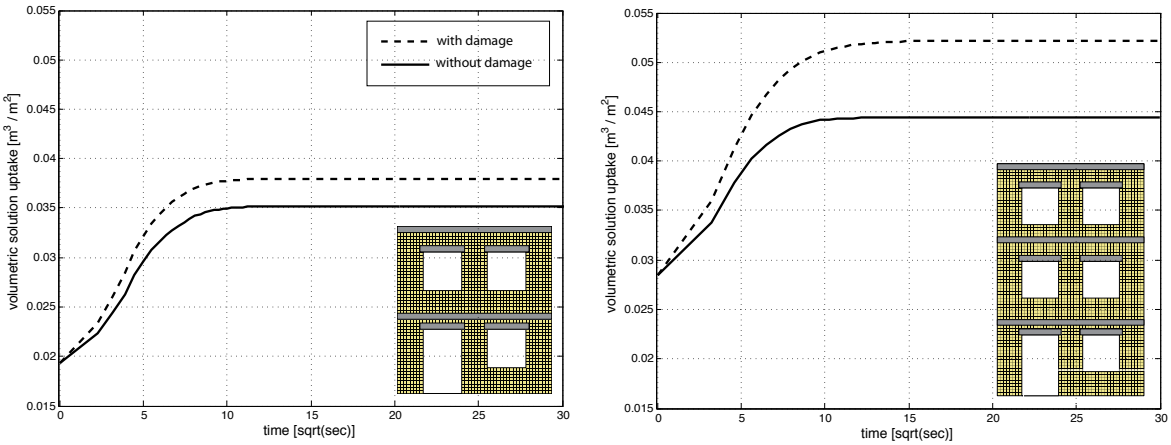


Figure 5: Volumetric water uptake with and without initial damage: two-storey wall on the left, three-storey wall on the right.

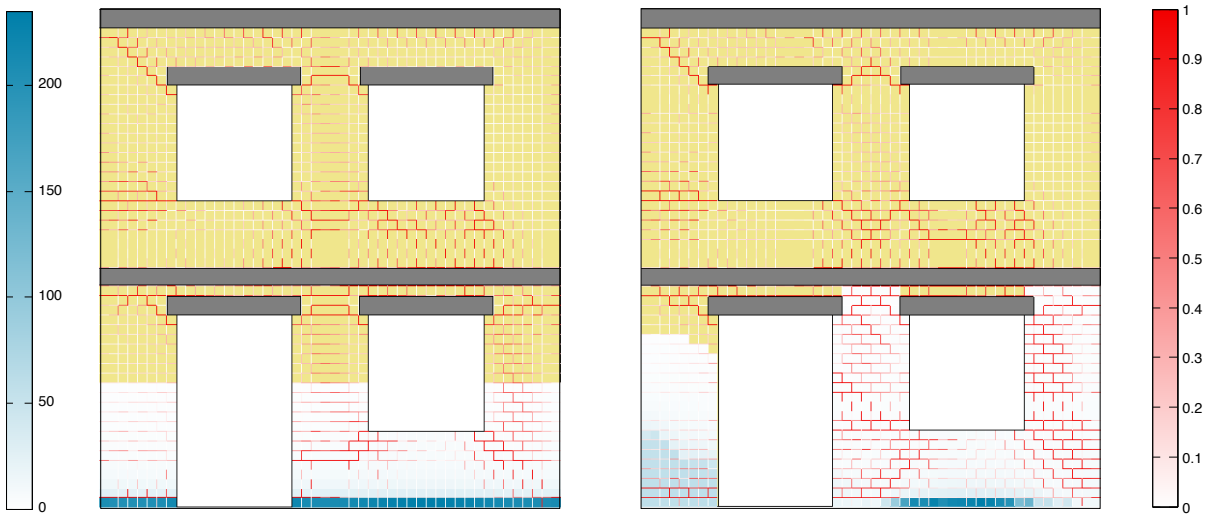


Figure 6: Two-storey wall: damage and water content distributions. Damage in dry conditions on the left, in wet conditions on the right; water content in undamaged conditions on the left, in damaged conditions on the right.

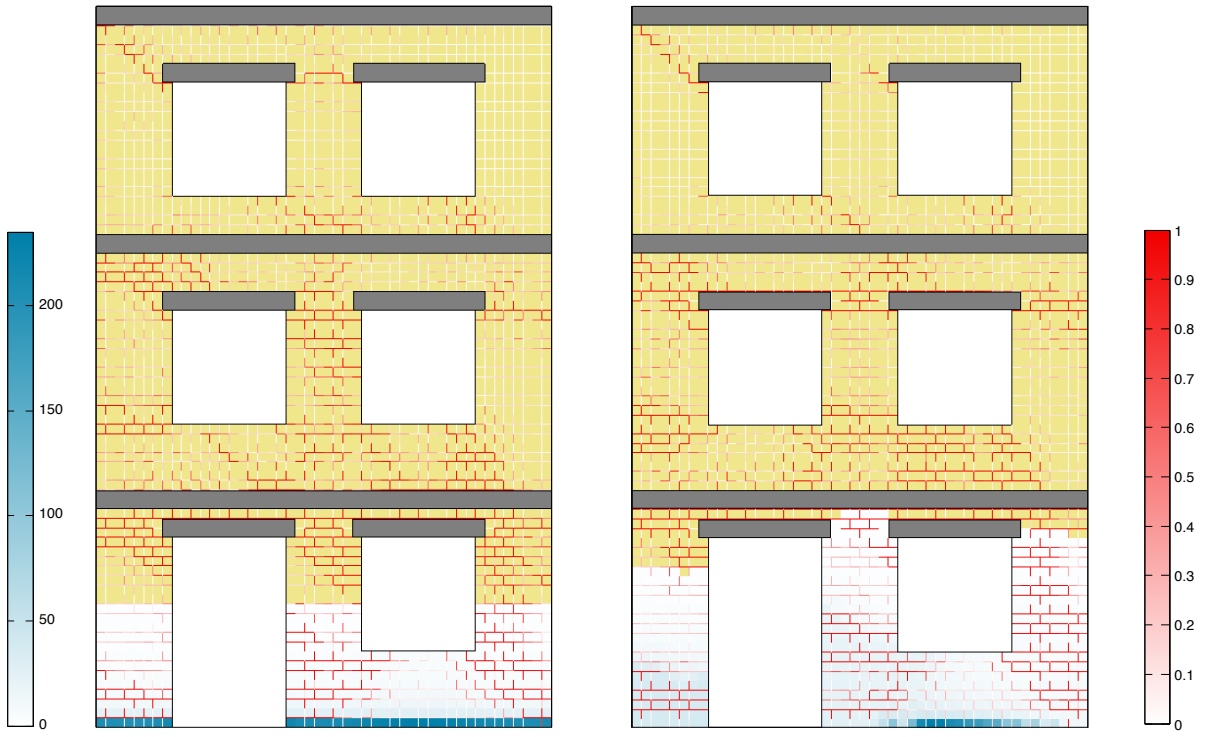


Figure 7: Three-storey wall: damage and water content distributions. Damage in dry conditions on the left, in wet conditions on the right; water content in undamaged conditions on the left, in damaged conditions on the right.

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis(Eds)

MODELLING OF AIR CHAMBER SUPPORTED FLOATING PLATFORMS – COUPLING FREE SURFACE FLOW, COMPRESSIBLE AIR, AND FLEXIBLE STRUCTURES

Florian TOTH*, Manfred KALTENBACHER† and Franz G. RAMMERSTORFER*

*Institute of Lightweight Design and Structural Biomechanics
Vienna University of Technology
Gusshausstrasse 27-29, 1040 Vienna, Austria
e-mail: ftoth@ilsb.tuwien.ac.at – Web page: http://www.ilsb.tuwien.ac.at

†Institute of Mechanics and Mechatronics
Vienna University of Technology
Getreidemarkt 9, 1060 Vienna, Austria
e-mail: manfred.kaltenbacher@tuwien.ac.at – Web page: http://www.mec.tuwien.ac.at

Key words: Finite Element Method, Surface Waves, Flexible Structure, Acoustics

Abstract. Air chamber supported floating platforms can significantly decrease wave induced structural responses. Novel applications, like floating arrays of solar collectors, with low payload requirements allow the design of floating platforms supported by large, cylindrical air chambers made of highly flexible membranes. In order to predict the dynamics of such systems a modelling strategy capturing all important phenomena: incompressible free surface flow, compressible air and flexible structures is presented. The governing partial differential equations and boundary conditions are given in their linearised form, and subsequently solved by the finite element method. A frequency domain formulation is chosen to compute the steady state response to harmonic excitation. In order to handle problems in unbounded domains a perfectly matched layer formulation is used. Thereby, radiating waves are efficiently damped at the edge of the computational domain. For the sake of simplicity we present two-dimensional, test problems used for the validation of the developed modelling strategy. Finally, we present a fully coupled simulation of wave interactions with a flexible, air chamber supported floating platform.

1 INTRODUCTION

Air chamber supported platforms typically consist of a plate-like structure which is carried by a number of air chambers. The chambers are formed by vertical walls extending from the plate-like structure into the water. The lower end of the air chambers is formed

by the free water surface. Thus, in a static situation, the equilibrium air pressure in the chamber is proportional to the weight of the platform. Travelling waves will change the volume of the chamber which leads to pressure variations and hence to a dynamic excitation of the structure.

The concept of air chamber supported floating platforms was pioneered by Pinkster and Meevers Scholte, who showed that air chambers can significantly reduce wave induced forces [14]. Early works assumed spatially constant pressure in the air chambers [13, 15], which is valid if the fundamental frequency of the first acoustic mode is well above the highest expected wave frequency. Using a semi-analytic approach for the acoustic modes in the air chamber, Lee and Newman [11] derived a model which takes spatially varying air pressure in a rigid chamber into account. The governing Laplace equation in the semi-infinite water domain is solved by the boundary element method (BEM). Small deformations of the platform structure can be taken into account by the use of generalised modal coordinates [7, 19]. The air chambers of all platform concepts investigated to date have relatively stiff bounding walls [6, 7, 14, 18, 19], allowing no significant volume change due to chamber deformations. These strong bounding walls are required because of the high payload requirements (in terms of mass) of the platforms.

For novel target applications, e.g. offshore solar collectors, these payload requirement can be as low as 25 kg/m^2. Therefore, the internal pressure in the air chamber is less than 1 MPa. This allows for air chambers being formed of cylindrical, highly flexible polymer membranes. When modelling such systems the deformation of the chamber walls due to pressure oscillations must be taken into account, because they lead to significant changes in the volume, which in turn influences the pressure: the system is strongly coupled.

2 PROBLEM FORMULATION

As an example a two-dimensional problem is depicted in Fig. 1. An arc-shaped flexible structure is partially submerged in water. It is entrapping compressible air, thereby forming an air chamber. All formulations are developed for the general three-dimensional case.

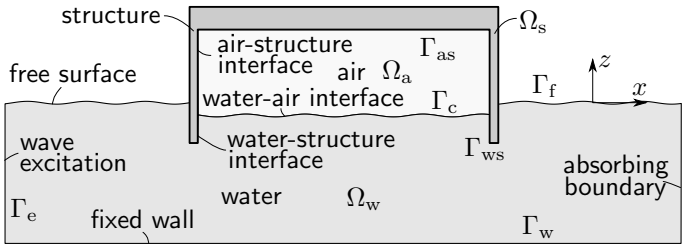


Figure 1: Sketch of the model domains.

2.1 Field Equations

Wave interaction problems require the modelling of unsteady free surface flow in the water domain denoted by Ω_w. A common modelling assumption in marine engineering is to idealise the water as an inviscid, incompressible fluid [3]. The governing equation then is the Laplace equation written as

$$-\nabla^2 p_\mathrm{w} = 0, \tag{1}$$

where ∇^2 denotes the Laplace operator, and p_w denotes the dynamic pressure in the water domain.

For the air in the chamber, i.e. the air domain Ω_a, compressibility must be taken into account. The air in the chamber is modelled as an ideal gas undergoing reversible, adiabatic changes. Furthermore, the dynamic changes in pressure, p_a, are assumed small in comparison with the static pressure and the gas is considered as inviscid. The governing equation then is the linear acoustic wave equation

$$\frac{1}{c_\mathrm{a}^2}\frac{\partial^2 p_\mathrm{a}}{\partial t^2} - \nabla^2 p_\mathrm{a} = 0\,, \tag{2}$$

where c_a is the speed of sound in the medium.

The wall of the air chamber, i.e. the structure domain Ω_s, will be considered as a linear elastic, isotropic solid. The governing equations then are the balance of momentum, the constitutive relation (Hooke's law), and the linearised strain displacement relationship

$$\rho\frac{\partial^2 \boldsymbol{u}}{\partial t^2} = \boldsymbol{\nabla}\cdot\boldsymbol{\sigma}\,, \tag{3a}$$

$$\boldsymbol{\sigma} = \boldsymbol{C}\boldsymbol{\epsilon}, \tag{3b}$$

$$\boldsymbol{\epsilon} = \frac{1}{2}\left(\boldsymbol{\nabla}\boldsymbol{u} + (\boldsymbol{\nabla}\boldsymbol{u})^T\right)\,, \tag{3c}$$

where $\boldsymbol{u}$ denotes the displacement vector, $\boldsymbol{\sigma}$ the (Cauchy-)stress tensor, $\boldsymbol{C}$ the stiffness tensor, and $\boldsymbol{\epsilon}$ the strain tensor. The juxtaposition of two elements, i.e. tensors of rank 0, 1, 2, or 4, denotes their dyadic product. Body forces were not included in the balance of momentum.

2.2 Boundary and Coupling Conditions

At the water surface both the dynamic boundary condition, i.e. the equality of the total pressure in water and air, and the kinematic boundary condition, i.e. interface particles must stay on the interface, have to be fulfilled. Introducing a coordinate system with the x-y-plane coinciding with the undisturbed water surface and the z axis pointing upwards against the vector of gravity, the hydrostatic pressure in the water domain is $p_{\mathrm{w},0} = p_{\mathrm{a},0} - \rho_\mathrm{w} g z$, where ρ_w denotes the mass density of water. The hydrostatic pressure

in the air, $p_{\mathrm{a},0}$, is assumed constant. The interface between air and water can be described by the function $z = \eta(x, y, t)$. Requiring equality of the total pressure, composed of hydrostatic and dynamic pressure, in water and air yields

$$p_{\mathrm{w}} - p_{\mathrm{a}} = \rho_{\mathrm{w}} g \eta \quad \text{on } \Gamma_{\mathrm{c}} \,. \tag{4}$$

The kinematic boundary condition, linearised and combined with the balance of momentum, yields

$$-\rho_i \frac{\partial^2 \eta}{\partial t^2} = \frac{\partial p_i}{\partial \boldsymbol{n}_\eta} \quad \text{on } \Gamma_{\mathrm{c}} \,, \tag{5}$$

where the index i is used for the water and air domain, respectively, and $\boldsymbol{n}_\eta$ denotes the normal vector of the water surface. In case of the free surface, i.e. water surface not coupled to compressible air, the dynamic boundary condition reduces to $p_{\mathrm{w}} = \rho_{\mathrm{w}} g \eta$ which can be used to eliminate η form the above equation giving

$$-\frac{1}{g} \frac{\partial^2 p_{\mathrm{w}}}{\partial t^2} = \frac{\partial p_{\mathrm{w}}}{\partial \boldsymbol{n}_\eta} \quad \text{on } \Gamma_{\mathrm{f}} \,. \tag{6}$$

At the solid-fluid interfaces between air/water and chamber wall the velocities perpendicular to the wall must be equal. In combination with the linearised momentum equation for the fluid one obtains

$$-\frac{1}{\rho_i} \nabla p_i \cdot \boldsymbol{n}_{\mathrm{s}} = \frac{\partial^2 \boldsymbol{u}}{\partial t^2} \cdot \boldsymbol{n}_{\mathrm{s}} \quad \text{on } \Gamma_{\mathrm{is}} \,, \tag{7}$$

where $\boldsymbol{n}_{\mathrm{s}}$ is the normal vector of the interface. The second coupling condition at the interface is the equality of surface stress and fluid pressure written as

$$\boldsymbol{\sigma} \cdot \boldsymbol{n}_{\mathrm{s}} = -p_i \boldsymbol{n}_{\mathrm{s}} \quad \text{on } \Gamma_{\mathrm{is}}. \tag{8}$$

The case of a stationary rigid wall can be considered as a special case of fluid structure coupling with immovable structure. The boundary condition then becomes

$$-\frac{1}{\rho_i} \nabla p_i \cdot \boldsymbol{n}_{\mathrm{s}} = 0 \quad \text{on } \Gamma_{\mathrm{w}} \,. \tag{9}$$

Waves in the fluid domain are generated by specifying a Neumann boundary condition for the pressure in the form

$$\frac{\partial p_i}{\partial \boldsymbol{n}} = f(x, y, z, t) \quad \text{on } \Gamma_{\mathrm{e}} \,, \tag{10}$$

where f is a known forcing function. For propagating surface gravity waves with low amplitude an analytical solution for the pressure field is known [9, 12], and can be used to compute the forcing function.

2.3 Perfectly Matched Layer Formulation

When studying wave propagation in semi-infinite domains the waves must be allowed to leave the finite computational domain. Therefore, a perfectly matched layer (PML) formulation is used to absorb outgoing waves without reflection. The water domain of interest, Ω_{w}, is surrounded by a PML of thickness L_k in each coordinate direction k. In this PML region, Ω_{PML}, we introduce to the complex change of variables

$$\tilde{x}_k(x_k) = x_k + \frac{1}{j\omega} \int_0^{x_k} \sigma_k(x) \, \mathrm{d}x \,; \quad x_k \in \{x, y, z\} \,, \tag{11}$$

according to Teixeira and Chew [16]. The damping function σ_k is positive inside Ω_{PML} and vanishes in Ω_{w}. Hence, we obtain the following relations

$$\frac{\partial \tilde{x}_k}{\partial x_k} = 1 + \frac{\sigma_k}{j\omega} = \zeta_k \quad \text{and} \quad \frac{\partial}{\partial \tilde{x}_k} = \frac{1}{\zeta_k} \frac{\partial}{\partial x_k} \,. \tag{12}$$

Performing a Fourier transform of the governing equation, Eq. (1), and the free surface boundary condition, Eq. (6), and inserting above relations we obtain the modified equations, which have to be solved in Ω_{PML}

$$-\frac{\partial}{\partial x}\left(\frac{1}{\zeta_x}\frac{\partial \hat{p}_{\mathrm{w}}}{\partial x}\right) - \frac{\partial}{\partial y}\left(\frac{1}{\zeta_y}\frac{\partial \hat{p}_{\mathrm{w}}}{\partial y}\right) - \frac{\partial}{\partial z}\left(\frac{1}{\zeta_z}\frac{\partial \hat{p}_{\mathrm{w}}}{\partial z}\right) = 0 \,; \qquad \frac{\partial \hat{p}_{\mathrm{w}}}{\partial \boldsymbol{n}} = \frac{w^2}{g}\hat{p}_{\mathrm{w}} \,, \tag{13}$$

with

$$\zeta_x = 1 + \frac{\sigma_x}{j\omega} \,; \quad \zeta_y = 1 + \frac{\sigma_y}{j\omega} \,; \quad \zeta_z = 1 + \frac{\sigma_z}{j\omega} \,. \tag{14}$$

The proper choice of the damping functions is of great importance, especially in order to obtain a very robust and efficient PML-technique [4]. A wave travelling through the PML region with a speed of c_p will be damped, but totally reflected at the outer boundary of the PML region. Assuming a layer thickness of L, e.g. in x-direction, the reflected wave reaches the interface between propagation and PML region with an amplitude of

$$\hat{p}_{\mathrm{w}} = p_0 e^{-(2/c_p)\cos\alpha \int_0^L \sigma_x(x)dx} = p_0 R \,. \tag{15}$$

Here, α models the angle with respect to the x-axis. In the choice of the reflection factor R the trade-off between the necessity of sufficient reduction of reflected waves according to Eq. (15) and possible disturbances of the numerical solution by a too rapid damping in a narrow PML region has to be taken into account. In our computations we use a value of $R = 10^{-3}$. Using Eq. (15) one obtains the relations:

- constant damping

$$\sigma_x^{\mathrm{const}} = \frac{-c_p \ln R}{2L\cos\varphi} \,,$$

- quadratically increasing damping

$$\sigma_x^{\text{quad}} = \frac{3c_p \ln R}{2L\cos\varphi}\,\frac{x^2}{L^2}\,,$$

- inverse distance damping

$$\sigma_x^{\text{inverse}} = \frac{c_p}{L-x}\,,$$

for different types of damping functions (for details see [8]). All of them were chosen directly proportional to the speed of the surface wave $c_p = \frac{\omega}{k} = \frac{g}{\omega}$ to remove the dependence on the wave frequency. As shown by Bermúdez et al. [2], the inverse distance damping function leads to an optimal damping behaviour for the acoustic wave equation in the frequency domain, and compared to other damping functions does not need specification of the reflection coefficient R. Within our numerical investigations, we will demonstrate that this is also true for our equation.

3 FINITE ELEMENT FORMULATION

To derive the FE formulation of the coupled problem, we assume a computational domain as displayed in Fig. 1 and just extend it by Ω_{PML} surrounding the water region Ω_{w}. Since we are interested in the time harmonic solution, we first perform a Fourier transform of the PDEs. The hat symbol, e.g. $\hat{p}$, is used to denote the Fourier transform of a quantity, e.g. p, and ω denotes the angular frequency. In order to arrive at the FE formulation, we introduce appropriate test functions φ, ψ, $\boldsymbol{\delta}$ and μ, multiply our coupled system of PDEs by these test functions and integrate over the whole computational domain. Furthermore, integration by parts* and incorporation of the boundary conditions, yields the weak (variational) formulation: Find $\hat{\boldsymbol{u}} \in (H_0^1)^3$, $\hat{p}_{\text{w}} \in H_0^1$, $\hat{p}_{\text{a}} \in H_0^1$, and $\hat{\eta} \in H_0^1$ such that†

$$\int_{\Omega_{\text{w}}} \boldsymbol{\nabla}\varphi \cdot \boldsymbol{\nabla}\hat{p}_{\text{w}}\,\mathrm{d}x - \int_{\Gamma_{\text{f}}} \frac{\omega^2}{g}\varphi\hat{p}_{\text{w}}\,\mathrm{d}s - \int_{\Gamma_{\text{c}}} \rho_{\text{w}}\omega^2\varphi\hat{\eta}\,\mathrm{d}s - \int_{\Gamma_{\text{ws}}} \rho_{\text{w}}\omega^2\varphi\hat{\boldsymbol{u}}\cdot\boldsymbol{n}_{\text{s}}\,\mathrm{d}s = \int_{\Gamma_{\text{w}}} \varphi\hat{f}\,\mathrm{d}s\,, \quad (16a)$$

$$\int_{\Omega_{\text{PML}}} \tilde{\boldsymbol{\nabla}}\varphi \cdot \tilde{\boldsymbol{\nabla}}\hat{p}_{\text{w}}\,\mathrm{d}x - \int_{\Gamma_{\text{f}}} \frac{\omega^2}{g}\varphi\hat{p}_{\text{w}}\,\mathrm{d}s = 0\,, \quad (16b)$$

*Using the relation $\boldsymbol{\nabla}\cdot(\boldsymbol{\delta}\cdot\boldsymbol{\sigma}) = (\boldsymbol{\nabla}\cdot\boldsymbol{\sigma})\cdot\boldsymbol{\delta} + \boldsymbol{\sigma} : (\boldsymbol{\nabla}\boldsymbol{\delta})$, where ':' denotes the double dot product, i.e. the sum of the products of conjugated tensor elements.

†H_0^1 is the space of functions, which are square integrable along with their first derivatives in a weak sense [1].

$$-\int_{\Omega_\mathrm{a}} \frac{\omega^2}{c_\mathrm{a}^2}\psi\hat{p}_\mathrm{a}\,\mathrm{d}x + \int_{\Omega_\mathrm{a}} \boldsymbol{\nabla}\psi\cdot\boldsymbol{\nabla}\hat{p}_\mathrm{a}\,\mathrm{d}x + \int_{\Gamma_\mathrm{c}} \rho_\mathrm{a}\omega^2\psi\hat{\eta}\,\mathrm{d}s - \int_{\Gamma_\mathrm{as}} \rho_\mathrm{a}\omega^2\psi\hat{\boldsymbol{u}}\cdot\boldsymbol{n}_\mathrm{s}\,\mathrm{d}s = 0\,, \tag{16c}$$

$$-\int_{\Omega_\mathrm{s}} \rho_\mathrm{s}\omega^2\boldsymbol{\delta}\cdot\hat{\boldsymbol{u}}\,\mathrm{d}x + \frac{1}{2}\int_{\Omega_\mathrm{s}} \boldsymbol{\nabla}\boldsymbol{\delta} : \boldsymbol{C}\left(\boldsymbol{\nabla}\hat{\boldsymbol{u}} + (\boldsymbol{\nabla}\hat{\boldsymbol{u}})^T\right)\mathrm{d}x$$
$$-\int_{\Gamma_\mathrm{ws}} \hat{p}_\mathrm{w}\boldsymbol{n}_\mathrm{s}\cdot\boldsymbol{\delta}\,\mathrm{d}s - \int_{\Gamma_\mathrm{as}} \hat{p}_\mathrm{a}\boldsymbol{n}_\mathrm{s}\cdot\boldsymbol{\delta}\,\mathrm{d}s = 0\,, \tag{16d}$$

$$\int_{\Gamma_\mathrm{c}} \mu\hat{p}_\mathrm{w}\,\mathrm{d}s - \int_{\Gamma_\mathrm{c}} \mu\hat{p}_\mathrm{a}\,\mathrm{d}s - \int_{\Gamma_\mathrm{c}} \mu\rho_\mathrm{w}g\hat{\eta}\,\mathrm{d}s = 0\,, \tag{16e}$$

for all test functions $\varphi \in H_0^1$, $\psi \in H_0^1$, $\boldsymbol{\delta} \in (H_0^1)^3$, and $\mu \in H_0^1$. It has to be noted that the surface elevation η acts as a Lagrange-multiplier to enforce the interface condition between water and air along Γ_c. The scaled Nabla-operator $\tilde{\boldsymbol{\nabla}}$ in Eq. (16b) computes according to Eqs. (13) and (14)

$$\tilde{\boldsymbol{\nabla}} = \left(\frac{1}{\zeta_x}\frac{\partial}{\partial x}\,,\ \frac{1}{\zeta_y}\frac{\partial}{\partial y}\,,\ \frac{1}{\zeta_z}\frac{\partial}{\partial z}\right)^T\,. \tag{17}$$

Furthermore, we want to note that the complex change of variables according to Eq. (11) has also to be performed, when computing the Jacobian within the FE procedure for Eq. (16b).

For the spatial discretisation we apply standard quadrilateral finite elements of 2nd order. The whole scheme has been implemented in the in-house research software CFS++ (Coupled Field Simulation) and used to compute the results presented in the following section. For some comparison, we have also applied the commercial software Abaqus [5].

4 RESULTS OF TEST PROBLEMS

Various numerical tests were carried out in order to validate the implementation of the described equations. The assumed material properties are given in Table 1(a). For the acceleration of gravity, g, a value of 9.81 m/s^2 was assumed. For the sake of simplicity two-dimensional problems are presented in the following.

Table 1: Assumed material properties for the fluids (a), and linear elastic, isotropic solid (b).

(a) Fluid properties

Property	Air	Water
c_i in m/s	343.05	1484.58
ρ_i in kg/m^3	1.2041	998.2

(b) Solid properties

E in MPa	800.0
ν	0.3
ρ in kg/m^3	1000.0

Table 2: Sloshing frequencies for the anticipated geometry.

Mode number n	0	1	2	3	4	5	6
Frequency in Hz	0.0306	0.0565	0.0764	0.0910	0.1044	0.1152	0.1247

4.1 Sloshing

In order to test the free surface boundary condition alone, the natural modes of a two-dimensional, rectangular tank with length $l = 350$ m and a water depth of $h = 50$ m were computed numerically. The domain was discretised by finite elements and the appropriate boundary conditions were applied (stationary wall at the left, bottom and right side; free surface condition at the top). The natural frequencies and corresponding mode shapes were computed by solving the standard eigenvalue problem which is obtained from the discretised ordinary differential equation system by making an harmonic Ansatz for the time dependant pressure. The obtained natural frequencies are given in Table 2. Comparison with the known analytical solution [9, 10] showed excellent agreement.

4.2 Travelling Waves

In order to test the efficiency of the PML formulation in a further test example a rectangular domain of water of uniform depth h=50 m was considered. The length of the domain was 350 m. Steady state solution of travelling waves were computed in a harmonic analyses at a range of wave frequencies. Waves were generated on the left wall, and a PML was used to absorb them at the right boundary.

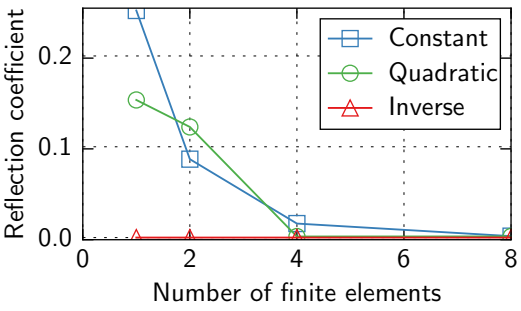
If the absorption is not perfect, a part of the wave is reflected, and a superposition of waves travelling in positive and negative direction is obtained. The combined wave system shows a spatially varying amplitude. The reflection coefficient, i.e. the ratio between the amplitude of the reflected wave and the amplitude of the incident wave, can be computed by

$$\frac{a_{\text{ref}}}{a_{\text{inc}}} = \frac{a_{\max} - a_{\min}}{a_{\max} + a_{\min}}, \tag{18}$$

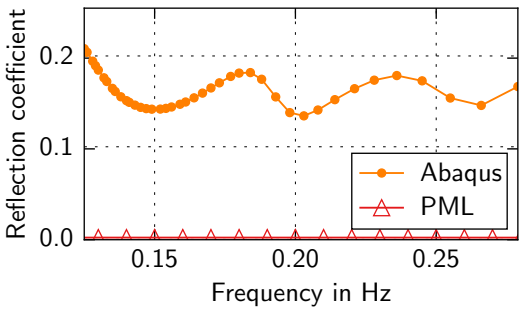
where $a_{\max}$ is the maximum amplitude of the combined wave system, and $a_{\min}$ is the minimum amplitude of the combined wave system.

The impact of the thickness of the PML was investigated by comparing results from computations with PMLs consisting of 1, 2, 4 and 8 finite elements, respectively. For all computations elements with 2nd order basis functions were used. The functions for the quadratically increasing as well as inverse distance damping were accordingly adjusted in each integration point. The results obtained for a wave frequency of 0.1 Hz, displayed in Fig. 2(a), demonstrate the extreme performance of the inverse distance damping function, which already shows an optimal performance with a layer thickness of just one finite element.

In Fig. 2(b) a comparison of the PML formulation with inverse damping function and acoustic infinite elements implemented in Abaqus [5] is shown for a range of wave



(a) Comparison of different damping functions at a wave frequency of 0.1 Hz.



(b) Comparison of Abaqus infinite elements and PML with inverse damping function.

Figure 2: Reflection coefficient in dependence of PML thickness (a) and wave frequency (b).

frequencies. The acoustic infinite elements in Abaqus are designed for the absorption of sound waves. In order to achieve absorption of surface gravity waves, the compressibility of the acoustic medium of the acoustic infinite elements was modified such that the speed of sound equals the wave speed of surface gravity waves. Therefore, the compressibility was set to $K_{\text{inf}} = \rho_{\text{w}} \left(\frac{g}{\omega}\right)^2$. The PML formulation with inverse damping function shows the best performance (with almost vanishing reflection coefficients) over all frequencies, due to the scaling of the damping function with speed c_p of the surface wave.

5 FLOATING STRUCTURE WITH SINGLE AIR CHAMBER

As an example of a fully coupled, two-dimensional problem a single floating air chamber as sketched in Fig. 1 was considered. A single chamber will likely become statically unstable, therefore, in a more realistic system at least three chambers must be used to provide sufficient safety against overturning [17]. The single chamber system is nevertheless chosen, because it is the simplest system to illustrate wave interaction effects with a flexible air chamber. Figure 3 shows the computed pressure in the water and the air domain. The flexible structure reflects a part of the incoming waves, visible in the spatially varying

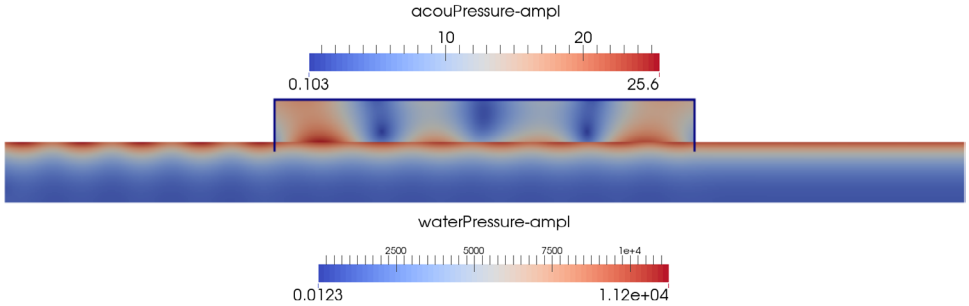


Figure 3: Computed pressure in the water and in the entrapped air. Values in Pa.

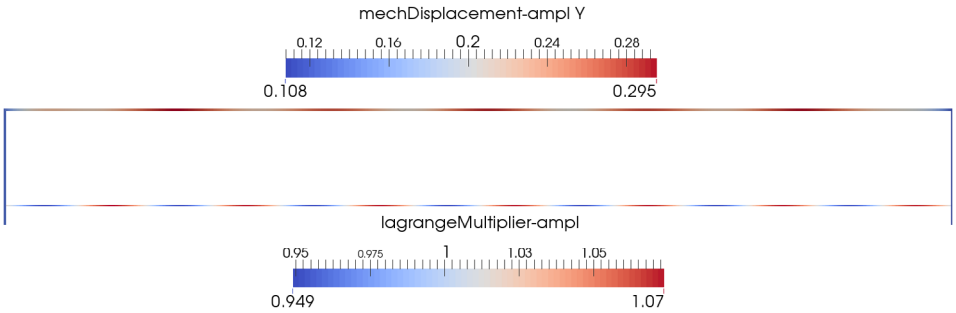


Figure 4: Computed mechanical displacement of the structure (in y-direction) and surface elevation. Values in m.

pressure amplitudes at the water surface on the left side. On the right side of the water domain, the pressure amplitude is constant, corresponding to waves propagating towards the right edge where they are completely absorbed by the PML. The wave amplitude is reduced to about 88% of the excitation amplitude behind the structure.

The structural displacements are displayed in Fig. 4, together with the surface elevation of the air-water interface, i.e. the Lagrange multiplier in the coupling equation. A combination of rigid body heave displacement, superimposed by a bending type deformation arises. The maximum displacement amplitudes are about one third of the wave amplitude.

6 CONCLUSIONS

A monolithic finite element formulation for the fully coupled analysis of incompressible free surface flow, deformable structures and compressible fluids was developed. As the formulation is based on linearised theory it is applicable to cases with sufficiently small disturbances from the equilibrium position. An harmonic formulation was chosen as it is most suitable for the analysis of wave interaction problems. To allow for the analysis of wave propagation in semi-infinite domains a highly efficient PML formulation was developed. Selected test cases were presented, to demonstrate the validity and performance of the modelling strategy.

Wave interactions with an air chamber supported floating platform were treated in a fully coupled manner. Here, the effect of a deformable chamber filled by a compressible fluid could be directly included. Typical target applications for the modelling approach are, hence, very large floating platforms, wave energy converters, or air chamber supported floating platforms. Furthermore, the presented modelling strategy can be applied to investigate sloshing effects in liquid filled tanks.

References

[1] Adams, R.A. *Sobolev Spaces*. Pure and Applied Mathematics. Academic Press, (1975).

[2] Bermúdez, A., Hervella-Nieto, L., Prieto, A., and Rodríguez, R. An optimal perfectly matched layer with unbounded absorbing function for time-harmonic acoustic scattering problems. *Journal of Computational Physics* (2007) **223**(2):469–488.

[3] Chakrabarti, S.K. *Numerical Models in Fluid-structure Interaction*. Advances in fluid mechanics. WIT, (2005).

[4] Collino, F. and Monk, P.B. Optimizing the perfectly matched layer. *Computer Methods in Applied Mechanics and Engineering* (1998) **164**:157–171.

[5] Dassault Systmes, . *Abaqus Analysis User's Manual*, 6.13-4 edition, (2013).

[6] Guret, R.A.M. *Interaction of free surface waves with elastic and air-cushion platforms*. PhD thesis, Delft University of Technology, (2003).

[7] Ikoma, T., Rheem, C.K., Maeda, H., and Masuda, K. Hydroelastic Behaviors of VLFS Supported by Many Aircushions With the Three-Dimensional Linear Theory. *Journal of Offshore Mechanics and Arctic Engineering* October (2011) **134**(1): 011104(8).

[8] Kaltenbacher, M. *Numerical Simulation of Mechatronic Sensors and Actuators*. Springer, Berlin, 2. edition, (2007). ISBN: 978-3-540-71359-3.

[9] Kundu, P.K., Cohen, I.M., and Dowling, D.R. *Fluid Mechanics*. Academic Press. Academic Press, (2012).

[10] Lamb, H. *Hydrodynamics*. Cambridge University Press, (1895).

[11] Lee, C.H. and Newman, J.N. Wave effects on large floating structures with air cushions. *Marine Structures* (2000) **13**(4-5):315–330.

[12] Lighthill, J. *Waves in Fluids (Cambridge Mathematical Library)*. Cambridge University Press, (2001).

[13] Malenica, S. and Zalar, M. An alternative method for linear hydrodynamics of air cushion supported floating bodies. In *Proceedings of the 15th IWWWFB*, (2000).

[14] Pinkster, J.A. and Meevers Scholte, E.J.A. The behaviour of a large air-supported MOB at sea. *Marine Structures* (2001) **14**(1-2):163–179.

[15] Pinkster, J. The effect of air cushions under floating offshore structures. In Vugts, J., editor, *Proceedings of the International Conference on Behaviour of Offshore Structures '97*. Pergamon, (1997).

[16] Teixeira, F. and Chew, W. Complex space approach to perfectly matched layers: a review and some developments. *International Journal of Numerical Modelling: Electronic Networks, Devices and Fields* (2000) **13**:441–455.

[17] Toth, F. *Static and Dynamic Modelling of Lightweight Floating Platforms Supported by Flexible Air Chambers*. PhD thesis, Vienna University of Technology, (2014).

[18] Tsubogo, T. and Okada, H. Hydroelastic Behavior of an Aircushion-Type Floating Structure. In Chung, J.S., Sayed, M., Kashiwagi, M., Setoguchi, T., and Hong, S.W., editors, *The Proceedings of The Twelfth (2002) International Offshore and Polar Engineering Conference*, (2002).

[19] van Kessel, J.L.F. *Aircushion Supported Mega-Floaters*. PhD thesis, Delft University of Technology, (2010).

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis(Eds)

COUPLING OF BOUNDARY ELEMENT REGIONS WITH THE BOUNDARY ELEMENT TEARING AND INTERCONNECTING METHOD (BETI)

CHRISTAN DUENSER*, BERNHARD LINDNER* AND GERNOT BEER†

*Institute for Structural Analysis
Graz University of Technology
Lessingstrasse 25, 8010 Graz, Austria
e-mail: duenser@tugraz.at, lindner@tugraz.at, www.ifb.tugraz.at

†Centre for Geotechnical and Materials Modelling
University of Newcastle
NSW 2308, Australia
e-mail: beer@tugraz.at

Key words: sequential excavation, BEM, BETI

Abstract. The boundary integral equation for elasticity is valid for a single domain consisting of homogeneous material properties. In the case of heterogeneity the consideration of different material properties is possible with a coupling of boundary element regions. Of course each region is again homogeneous. Another simulation application of multiple regions is the simulation of an industrial process, where different subdomains of a homogenous domain are treated differently due to a mechanical process. For instance, this is the case in tunnelling, where excavation is performed in a staged procedure. In the simulation of such an excavation process regions are deactivated step by step. As the material behaviour can be nonlinear an accurate simulation of such a staged process is a necessary requirement. Thus, the domain is decomposed into subregions which are coupled to neighbouring regions. There are different coupling strategies existing. In some of them stiffness matrices of subdomains are worked out which are the basis for the coupling and solution of the problem. A traditional method is the coupling of interface surfaces only [1]. In this method the stiffness matrix of a region is computed on the basis of the coupling surfaces (interfaces), whereas the coupling surface may be not identical to the complete surface of a subdomain and the size of the stiffness matrix is determined by the degrees of freedom of the coupling surface. In an application where the boundary conditions change (e.g. from interface to Neumann condition) from one calculation step to the other, the stiffness matrix has to be calculated new. A modern coupling technique is the Boundary Element Tearing and Interconnecting (BETI) method [2], similar to the method of Finite Element Tearing and Interconnecting (FETI) [3]. In

this method the region stiffness matrix is worked out for the entire boundary of the region. The stiffness matrices of all regions remain the same during the whole analysis, even if the boundary conditions change during the simulation process. In setting up the equation system each subdomain is treated completely separated and independent from the others. Thus, a parallelisation of the computational work is ideally suited and implemented in the present computer code. In this work the theory of both mentioned coupling techniques are introduced briefly. The differences of both methods are worked out and advantages/disadvantages are shown and will be demonstrated. The accuracy of the results as well as the computational performance will be shown and compared based on a realistic simulation example.

1 INTRODUCTION

Within a tunnel excavation according the New Austrian Tunneling Method (NATM) parts of the tunnel volume are excavated in a staged procedure. Due to this process the tunnel construction is dependent on spatial and temporal development. In order to provide certain predictions or in the case of verification of ongoing tunnel constructions a numerical simulation has to consider those requirements. A typical tunnel construction is shown in Figure 1. The tunnel cross section is divided into a top heading and a bench part. In longitudinal direction the excavation of the top heading is more advanced then the bench. The excavation is done in a way where volumes of rock with specified thickness are removed.

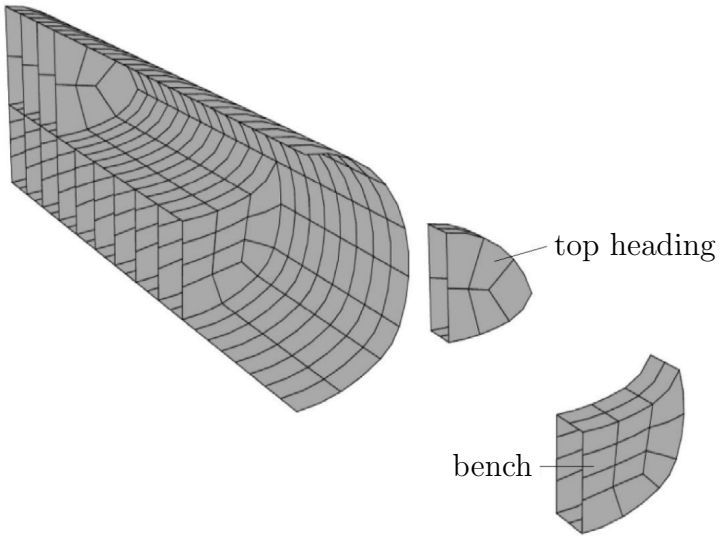


Figure 1: Sequential tunnel excavation

As a tunnel construction typicaly occurs in an infinite or semi-infinite domain the Boundary Element Method (BEM) is ideally suited for the numerical simulation of this type of problems [4]. The fundamental solution used in BEM [5] are special solutions for infinite domains, they allready fullfill the radiation conditons. There is no need to truncate a mesh and therefore no artificial boundary conditions have to be applied. With the BEM only the surface of the regions has to be discretised, thus the effort of mesh generation is drastically reduced for such simulation problems. As shown in Figure 1 the domain is subdivided into several regions. For the excavation volumes top heading and bench finite regions are used. These regions are embedded in an infinite region which represents the infinite extend of the domain. For the numerical solution of such problems there is a need to couple these regions. In chapter 2 the displacement boundary integral equation, its discretisation and solution is shown. In chapter 3 the development of two multiple region BEM (MRBEM) coupling methods will be addressed. In chapter 4 the coupling methods are demonstrated on a practical tunnel excavation example in 3D.

2 BOUNDARY INTEGRAL EQUATION

The displacement boundary integral equation [5] shown in Equation (1) is the basis of the BEM:

$$\mathbf{C}(\mathbf{y})\mathbf{u}(\mathbf{y}) + \cancel{\int_{\Gamma}} \mathbf{T}(\mathbf{y},\mathbf{x})\mathbf{u}(\mathbf{x})\,\mathrm{d}\Gamma = \int_{\Gamma} \mathbf{U}(\mathbf{y},\mathbf{x})\mathbf{t}(\mathbf{x})\,\mathrm{d}\Gamma \tag{1}$$

$\mathbf{T}(\mathbf{y},\mathbf{x})$ and $\mathbf{U}(\mathbf{y},\mathbf{x})$ are the fundamental solutions and $\mathbf{u}(\mathbf{x})$ and $\mathbf{t}(\mathbf{x})$ are the boundary displacements and tractions, respectively. The boundary integral equation (1) is valid for a single region whose boundary is discretised by boundary elements. Due to the discretisation of the integral equation the boundary Γ is divided into a sum of elements E and nodes N. Introducing the discretisation the integral equation (1) is transformed into the following form:

$$\mathbf{C}\mathbf{u}_i + \sum_{e=1}^{E}\sum_{n=1}^{N} \boldsymbol{\Delta}\mathbf{T}^e_{ni}\mathbf{u}^e_n = \sum_{e=1}^{E}\sum_{n=1}^{N} \boldsymbol{\Delta}\mathbf{U}^e_{ni}\mathbf{t}^e_n \tag{2}$$

$\boldsymbol{\Delta}\mathbf{T}^e_{ni}$ and $\boldsymbol{\Delta}\mathbf{U}^e_{ni}$ are integrated kernel coefficients with respect to the collocation node i and element n. $\mathbf{C}$ is the integral free term which depends on the geometrical conditions at node i. Equation (2) is evaluated for all collocation points i and the coefficients $\boldsymbol{\Delta}\mathbf{T}^e_{ni}$ and $\boldsymbol{\Delta}\mathbf{U}^e_{ni}$ are assembled into matrices $[\Delta T]$ and $[\Delta U]$, whereas the following equation arise:

$$[\Delta T]\{u\} = [\Delta U]\{t\} \tag{3}$$

With equation (3) a single boundary element region can be solved. At the nodes of the boundary either displacements or tractions are known. The unknown boundary conditions

(BC's) are solved by rearranging equation (3). The unknown BC's with its corresponding columns of matrices $[\Delta T]$ and $[\Delta U]$ are shifted to the left side and the known BC's are multiplied with the columns of the matrices $[\Delta T]$ or $[\Delta U]$ and form the right hand side vector $\{f\}$ of the following equation:

$$[A]\{x\} = \{f\} \tag{4}$$

In case of a mixed boundary value problem the content of the solution vector $\{x\}$ are either displacements or tractions, matrix $[A]$ is filled up either with columns of matrix $[\Delta T]$ or $[\Delta U]$.

3 MULTIPLE REGION BEM (MRBEM)

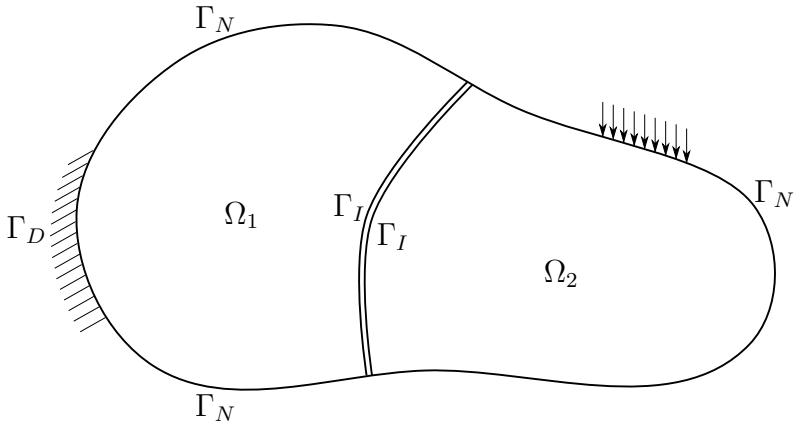


Figure 2: Multiple region system

If there is a system with more then one region present (as shown in Figure 2) a multiple region boundary element method (MRBEM) is a possibility to solve such a problem. In comparison to the single region BEM the boundary of a region is extended by the interface boundary. The whole boundary of a region may be devided into a Neumann (Γ_N), Dirichlet (Γ_D) and Interface (Γ_I) part. At the interface displacements and tractions are unknown. Thus additional conditions are necessary to solve this problem. These conditions are equilibrium and compatability at the interface of adjacent regions. For the solution of such a coupled system of regions different method formulations are available. Two of them are discussed next. For both methods stiffness matrices are worked out. In the first method (Interface Coupling (IC) method) the stiffness matrix of each region is based on the degrees of freedom (DOFs) at the interface, in the second (BETI method) the stiffness matrix is based on the DOFs of the whole region.

3.1 IC method

For the derivation of the equation system of a multiple region system Equation (3) is shown again as following:

$$[\Delta U]\{t\} = [\Delta T]\{u\} \tag{5}$$

Due to the three different boundary types (Neumann, Dirichlet and Interface) Equation (5) is expanded to:

$$\begin{bmatrix} [U_I] & [U_N] & [U_D] \end{bmatrix} \begin{Bmatrix} \{t_I\} \\ \{t_N\} \\ \{t_D\} \end{Bmatrix} = \begin{bmatrix} [T_I] & [T_N] & [T_D] \end{bmatrix} \begin{Bmatrix} \{u_I\} \\ \{u_N\} \\ \{u_D\} \end{Bmatrix} \tag{6}$$

The tractions $\{t_N\}$ are known BC's and they will be shifted to the right side of the equation system together with the associated kernel matrix $[U_N]$. $\{u_N\}$ are unknown displacements at the Neumann boundary and they will be moved to the left side together with matrix $[T_N]$ as shown next:

$$\begin{bmatrix} [U_I] & -[T_N] & [U_D] \end{bmatrix} \begin{Bmatrix} \{t_I\} \\ \{u_N\} \\ \{t_D\} \end{Bmatrix} = \begin{bmatrix} [T_I] & -[U_N] & [T_D] \end{bmatrix} \begin{Bmatrix} \{u_I\} \\ \{t_N\} \\ \{u_D\} \end{Bmatrix} \tag{7}$$

The matrix on the left side of Equation (7) is renamed by $[A]$ and on the right hand side the interface displacements $\{u_I\}$ are seperated as following:

$$\begin{bmatrix} \quad A \quad \end{bmatrix} \begin{Bmatrix} \{t_I\} \\ \{u_N\} \\ \{t_D\} \end{Bmatrix} = [T_I]\ \{u_I\} + \begin{bmatrix} -[U_N] & [T_D] \end{bmatrix} \begin{Bmatrix} \{t_N\} \\ \{u_D\} \end{Bmatrix} \tag{8}$$

Solving Equation (8) for the vector on the left will result in:

$$\begin{Bmatrix} \{t_I\} \\ \{u_N\} \\ \{t_D\} \end{Bmatrix} = [A]^{-1}[T_I]\ \{u_I\} + [A]^{-1}\big[-[U_N]\ [T_D]\big] \begin{Bmatrix} \{t_N\} \\ \{u_D\} \end{Bmatrix} \tag{9}$$

This equation can be simplified to:

$$\begin{Bmatrix} \{t_I\} \\ \{u_N\} \\ \{t_D\} \end{Bmatrix} = \begin{bmatrix} [K]^* \\ [D] \end{bmatrix} \{u_I\} + \begin{Bmatrix} \{t_{I0}\} \\ \{u_{N0}\} \\ \{t_{D0}\} \end{Bmatrix} \tag{10}$$

Taking the first equation from (10), which is:

$$\{t_I\} = [K]^*\{u_I\} + \{t_{I0}\} \tag{11}$$

and multiplying it with the mass matrix $[M]$ the tractions of Equation (11) are transformed to work equivalent nodal point forces. Because of this a coupling to the Finite Element Method (FEM) is possible. As a result Equation (11) is transformed to:

$$\{f_I\}_r = [K]_r\{u_I\}_r + \{f_{I0}\}_r \tag{12}$$

whereas $\{f_I\}_r = [M]_r\{t_I\}_r$, $[K]_r = [M]_r[K]_r^*$ and $\{f_{I0}\}_r = [M]_r\{t_{I0}\}_r$. The final forces at the interface of region r are the forces due to the interface displacements plus the forces at the interface due to the loading (given tractions $\{t_N\}_r$ and applied displacements $\{u_D\}_r$). The final interface forces $\{f_I\}_r$ and the interface displacements $\{u_I\}_r$ are unknown at the present state. Thus Equation (12) has to be applied to every region of the coupled system and the final system of equation can be assembled under the following conditions:

- Equilibrium of forces at the interface:

$$\{f_I\}_1 + \{f_I\}_2 = 0 \tag{13}$$

  Equation (13) states that the forces at the interface of region 1 are in equilibrium with those of the neighbouring region 2.

- Compatability of displacements at the interface:

$$\{u_I\}_1 = \{u_I\}_2 \tag{14}$$

  Equation (14) states that the displacements at the interface of region 1 are equal with those of the neighbouring region 2.

Considering these conditions Equation (12) for every region is assembled into a global equation system which is shown as following:

$$\{f_I\}_r = [K]^{sys}\{u_I\} + \{f_{I0}\} = 0 \tag{15}$$

where $[K]^{sys}$ is the assembled stiffness matrix related to all coupling interfaces of the system. $\{f_{I0}\}$ is the right hand side vector related to the loading of the system and $\{u_I\}$ is the vector of interface displacements. This equation is solved for the interface displacements. Once $\{u_I\}$ is known all remaining unknowns (tractions at the interface $\{t_I\}$, displacements at the Neumann boundary $\{u_N\}$ and tractions at the Dirichlet boundary $\{t_D\}$) can be evaluated using Equation (10).

3.2 BETI method

The Boundary Element Tearing and Interconnecting Method (BETI) is a domain decomposition method for the Symmetric Galerkin BEM [2] similar to the Finite Element Tearing and Interconnecting Method (FETI) for the FEM introduced by [3]. Using similar concepts the BETI method will be applied in this work to the collocation BEM. For each region this method works out a stiffness matrix which is based on the DOFs of the whole region surface in contrast to the method applied in chapter 3.1 where the stiffness matrix is based on the coupled DOFs only. From Equation (3) the boundary tractions of a region can be calculated as following:

$$[\Delta U]^{-1}[\Delta T]\{u\} = \{t\} \tag{16}$$

To make possible a coupling to the FEM, Equation (16) is multiplied with the mass matrix $[M]$. Using this procedure the boundary tractions are transformed to work equivalent nodal point forces acting at the nodes of the boundary:

$$[M][\Delta U]^{-1}[\Delta T]\{u\} = [M]\{t\} = \{f\} \tag{17}$$

where the stiffness matrix $[K]$ is:

$$[K] = [M][\Delta U]^{-1}[\Delta T] \tag{18}$$

Inserting this into Equation (17) will result in the well known relation between displacements and forces:

$$[K]\{u\} = \{f\} \tag{19}$$

In order to formulate a coupled system of R boundary element regions two conditions have to be satisfied:

- Equilibrium
- Compatability

3.2.1 Equilibrium of a boundary element region

The equilibrium state of a region can be described by using Equation (19):

$$[K]\{u\} = \{f_N\} + [B]^T\{\lambda\} \tag{20}$$

whereas the force vector on the right hand side of Equation (19) is splitted into:

$$\{f\} = \{f_N\} + [B]^T\{\lambda\} \tag{21}$$

and inserted into Equation (20). $[K]\{u\}$ are the forces at the boundary of the region due to deformation, $\{f_N\}$ is the force vector of the given loading (Neumann boundary conditions) and $[B]^T\{\lambda\}$ are the coupling forces (Lagrange multipliers) to the neighbouring regions.

3.2.2 Compatability of interface displacements

The compatability of a system of R regions can be written in following form:

$$[B]_1\{u\}_1 + [B]_2\{u\}_2 + \cdots + [B]_R\{u\}_R = \{b\} \tag{22}$$

Equation (22) either guaranties that the displacements at the interface of adjacent regions are equal or that the displacements at the Dirichlet boundary are equal to the applied Dirichlet boundary conditions which are entries of vector $\{b\}$.

3.2.3 System of equation

The final system of equation of a coupled system of R boundary element regions is shown as following:

$$\begin{bmatrix} [K]_1 & & 0 & & -[B]_1^T \\ & [K]_2 & & & -[B]_2^T \\ 0 & & \ddots & & \vdots \\ & & & [K]_R & -[B]_R^T \\ [B]_1 & [B]_2 & \cdots & [B]_R & 0 \end{bmatrix} \cdot \begin{Bmatrix} \{u\}_1 \\ \{u\}_2 \\ \vdots \\ \{u\}_R \\ \{\lambda\} \end{Bmatrix} = \begin{Bmatrix} \{f_N\}_1 \\ \{f_N\}_2 \\ \vdots \\ \{f_N\}_R \\ \{b\} \end{Bmatrix} \tag{23}$$

Equations 1 to R of Equation (23) are representing the equilibrium of each region and the last equation of (23) guaranties compatability of displacements at every node at the interface of adjacent regions and at the nodes of the Dirichlet boundary.

In the implementation of the BETI method Equation (23) is not assembled to an equation system. The equation system (23) is condensed to the solution of the coupling forces λ (Lagrange multipliers). This is done by inserting equations 1 to R into the last equation of (23). From this equation λ is solved either directly or iteratively with a BiCGSTAB iterative solver. As the stiffness matrix $[K]_r$ of a finite region (floating region) is singular special treatment of rigid body motions have to be considered. The whole solution formulation is shown in detail by [2].

3.3 Comparison of coupling methods

The main advantage of the BETI method is that the stiffness matrix of each region has to be calculated only once and in the case of a sequential tunnel excavation these matrices can be used for each load step of excavation. Using the BETI method for this application type the stiffness matrices are independent on the changing boundary conditions. Changing boundary conditions due to sequential excavation are considered by the coupling matrix $[B]_r$ of Equation (23). The coupling matrices have to be computed again for each calculation step. As those matrices are sparsely populated they are implemented as sparse matrices. The effort to set up those matrices is small and it is insignificant compared to the overall computing time. The way how the equation system is formulated

makes the treatment of operations at the region independently from the other regions. Thus, the BETI method is ideally suited for parallelisation.

The advantage of the IC method is that the size of the stiffness matrix is related to the number of DOFs at the interface of the coupled system of regions. In the case of a sequential excavation the coupling surfaces are reduced from one excavation step to the other. In each load step one or more regions are deactivated from the simulation model. Due to the deactivation the boundary condition of surfaces adjacent to the deactivated regions change from Interface condition to Neumann condition. Thus, the size of the assembled system stiffness matrix reduces from one load step to the other and the solution of the equation system gets faster. For regions for which a change of boundary conditions happens the stiffness matrix has to be calculated again. Compared to the BETI method stiffness matrices do not remain constant throughout the entire analysis of such an excavation simulation.

4 EXAMPLE - 3D TUNNEL EXCAVATION

In this example a 3D tunnel excavation is investigated. In Figure 3 the boundary discretisation is shown. The tunnel geometry is subdivided into top heading and bench regions (finite regions). These regions are embedded in an infinite region which represent the infinite extend of the domain. The mesh is discretised with quadratic boundary elements. At the starting of the tunnel quadratic plane strain infinite boundary elements are used to simulate the infinite extend of the tunnel behind the tunnel face. Two different meshes are investigated, a coarse mesh and a refined mesh as shown in Figure (3).

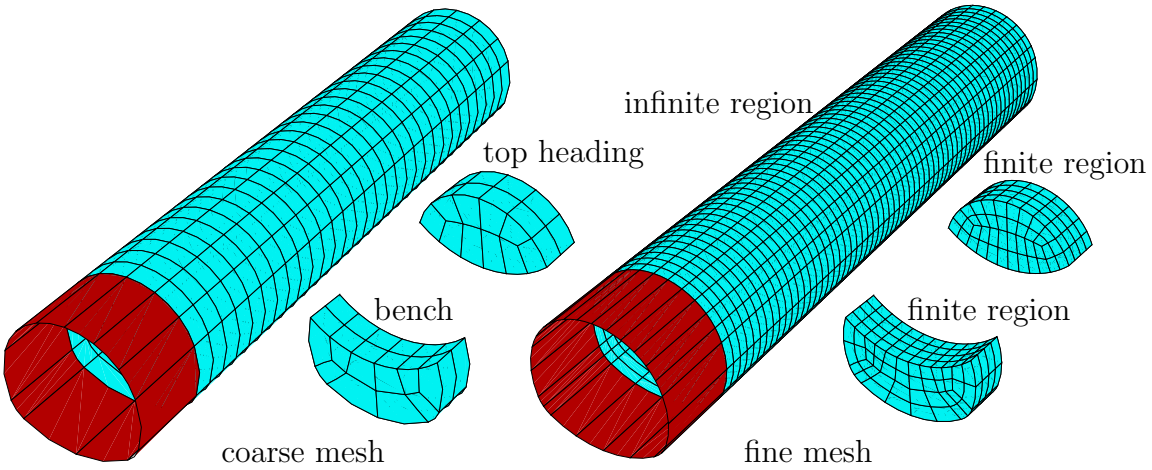


Figure 3: Discretisation

A constant primary stress field with $\sigma_{xx} = -1.375$, $\sigma_{yy} = -1.375$, $\sigma_{zz} = -2.75$ and $\sigma_{xy} = \sigma_{yz} = \sigma_{zx} = 0.0$ is assumed. From this stress field the excavation tractions are computed by $t = \sigma \cdot n$, where n is the unit vector normal to the surface of the tunnel.

This traction is applied at the free surface, this is the surface of the tunnel just excavated in load step 1. The loading for every subsequent step is the traction applied at the parts of the tunnel surface which changed from coupled interface conditions to free surfaces, these are surfaces at adjacent regions to the deactivated regions. The tractions which are applied in the current load step are taken from the result tractions at the coupled interfaces from the previous laod step. From step 2 to 15 top heading and bench regions are excavated alternately. The sequence of excavation is shown in Figure 4, where regions are marked by the number of load step they are excavated. The excavation of the top heading part takes place further ahead then the bench, which can be observed in Figure 4, too. Within an excavation step one or several regions might be excavated. In the first step from the beginning top heading and bench regions are excavated to reach a typical staged excavation configuration in the longitudinal direction.

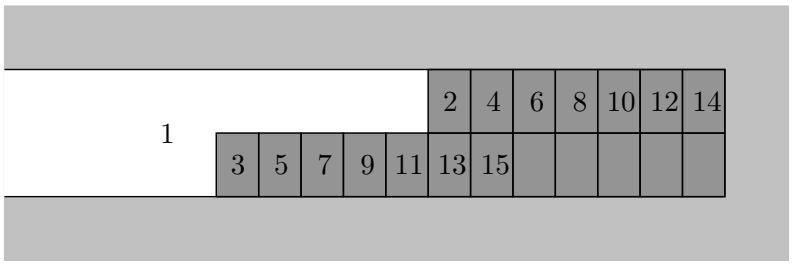


Figure 4: Sequence of excavation load steps

In Figure 5 results for the displacements u_z are shown at the infinite region and at finite regions (embedded in the infinite region) for the refined discretisation.

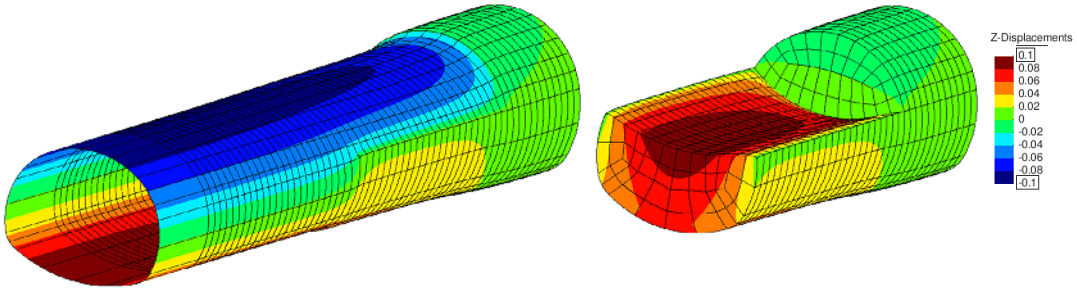


Figure 5: Displacements u_z for excavation load step 8

In the diagram of Figure 6 the deformation curves along a line (parallel to the longitudinal tunnel axis) at the crown of the tunnel for the displacements u_z are shown . The curves are shown for the coarse and fine discretisation and for all load cases. Results are shown only for the interpolation type Discontinuous. As can be seen the solutions for the coarse mesh are almost of the same quality as the solution of the fine mesh.

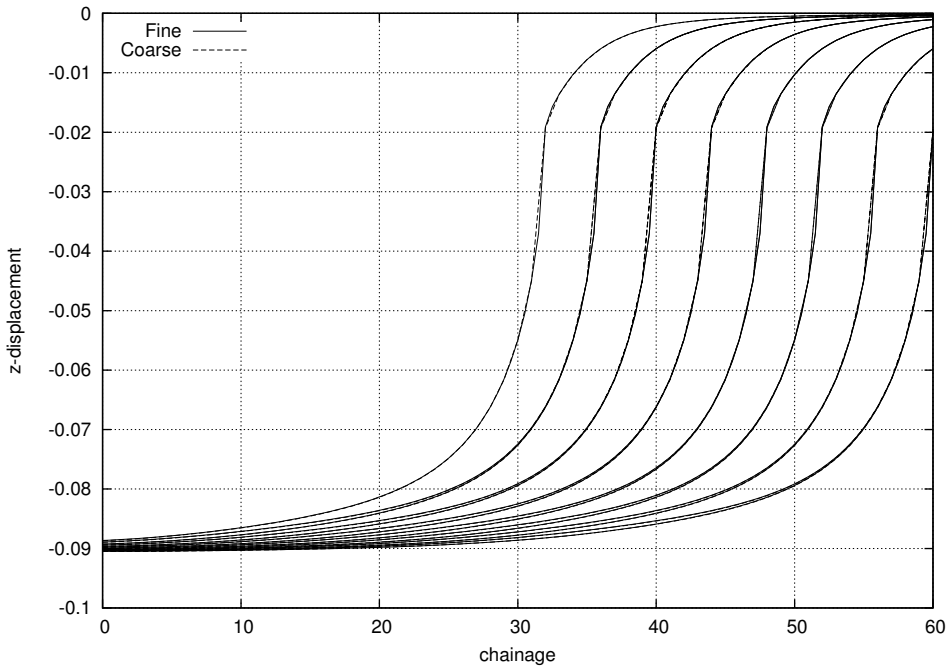


Figure 6: Displacements u_z at the crown of the tunnel

In Table 1 the number of DOFs for the infinite region and for the size of the global system of equation are shown for the coarse and fine discretisation. Three types of interpolation are choosen for the evaluation of this example - Continuous, Mixed and Discontinuous. For the Continuous type continuous interpolation is applied for the displacement and traction field, the Mixed type uses continuous interpolation for the displacement field and discontinuos for the traction field. Within the Discontinuous type both fields are interpolated discontinuous. The order of interpolation is choosen quadratic for all types of interpolation.

Table 1: Degrees of Freedom

	infinite Region				Global	
	Dirichlet DOFs		Neumann DOFs		DOFs	
	Coarse	Fine	Coarse	Fine	Coarse	Fine
Continous	3459	13755	3459	13755	5403	19785
Mixed	3459	13755	9192	36624	5403	19785
Discontinous	9192	36624	9192	36624	12048	48192

In relation to the numbers of DOFs of Table 1 in Table 2 the calculation times are shown for the methods IC and BETI. The calculations are done with the parallised version of

the code using 24 CPUs. As can be seen for all discretisations and interpolation types the BETI method is faster then the IC method. According to our experiences the difference in computing time of both methods is getting greater if the number of load cases is increased. The reason for this is the repeated calculation of region stiffness matrices due to changing boundary conditions for the method IC.

Table 2: Calculation Time

	Calculation Time			
	IC		BETI	
	Coarse	Fine	Coarse	Fine
Continous	140	2027	59	524
Mixed	241	3613	136	1116
Discontinous	394	8706	215	3034

5 CONCLUSIONS

In the present work two coupling methods are investigated, the IC method and the BETI method. Due to the formulation of stiffness matrices based on work equivalent nodal point forces both methods are able to couple BE regions (BEM/BEM coupling) and BE regions to FE regions (BEM/FEM coupling). The main differences of both methods are worked out and on a practical 3D tunnel excavation example results are shown. Particular attention has been paid to the computational performance of both coupling techniques.

REFERENCES

[1] Duenser, C. and Beer, G. *Simulation of sequential excavation with the Boundary Element Method. Computers and Geotechnics*, 44(0):157 – 166, (2012).

[2] Langer, U. and Steinbach, O. *Boundary element tearing and interconnecting methods. Computing*, 71(3):205–228, (2003).

[3] Farhat, C. and Roux, F.-X. *A method of finite element tearing and interconnecting and its parallel solution algorithm. International Journal for Numerical Methods in Engineering*, 32(6):1205–1227, (1991).

[4] Duenser, C. *Simulation of Sequential Tunnel Excavation with the Boundary Element Method*. Monographic Series TU Graz: Structural Analysis. Verlag der Technischen Universität Graz, (2007).

[5] Beer, G., Smith, I. and Duenser, C. *The Boundary Element Method with Programming*. SpringerWienNewYork, (2008).

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis(Eds)

IMPLEMENTATION OF AN EFFICIENT COUPLED FEM-SBFEM APPROACH FOR SOIL-STRUCTURE-INTERACTION ANALYSIS

Marco Schauer*, Sabine C. Langer†

*Technische Universität Braunschweig
Institut für Statik
Beethovenstrae 51, 38106 Braunschweig, Germany
e-mail: m.schauer@tu-braunschweig.de, web page: https://www.tu-braunschweig.de/statik

†Technische Universität Braunschweig
Institut für Konstruktionstechnik
Langer Kamp 8, 38106 Braunschweig, Germany
e-mail: s.langer@tu-braunschweig.de, web page: https://www.tu-braunschweig.de/ik

Key words: scaled boundary finite element method, model reduction, soil-structure-interaktion

Abstract. Buildings are grounded in the surrounding soil, so that soil and structure interact with each other. Consequently in the soil induced vibrations are transmitted to the structures. Neighbouring buildings and structures interact with each other, as they are connected by the soil. Nowadays numerical simulation of soil structure interaction is of great interest and is applied to very different problems. These include for example the construction of reliable earthquake-resistant structures in seismic active areas, and also the increase of comfort of buildings by decouple them form surrounding emissions like vibrations induced by traffic of machine foundations.

This work shows that the simulation of soil-structure-interaction taking unbounded domains into account, which fulfils the Sommerfeld radiation condition exactly, is not only possible for academic examples, but for large scale real life problems as well. Therefore two numerical methods where coupled to create an efficient coupled method, which can be used to simulate soil-structure-interaction in time domain. The numerical implementation of this coupled approach bases on a combination of finite element method [1] and scaled boundary finite element method [2]. The finite element method is used to discretise the near-field, containing structures and its surrounding soil. The coupled infinite half-space, the far-field is realised by the scaled boundary finite element method.

A contemporary parallel implementation of the coupling algorithms is done, since the simulation of soil structure interaction in time domain is very time and memory consuming [3]. Subsequent the numerical performance of the implemented software is discussed

in terms of speed-up and efficiency. Different geotechnical applications are illustrated and the applicability of the coupled method is shown and discussed on chosen examples.

1 INTRODUCTION

Accurate simulations of wave-propagation is essential for soil-structure interaction (SSI) analysis. Unbounded domains or infinite half-spaces, require careful analysis and call for efficient methods in order to model wave-propagation to infinity. Whenever vibrations or impulses are emitted to soil, they induce waves traveling through the ground that can provoke structures to vibrate and even to fail. It is essential not only to analyze the structure itself but also to take the surrounding soil into account [4, 5, 6].

Different methods considering the surrounding unbounded domain by a transmitting boundary, have been developed during the last years. One of the simplest transmitting boundary conditions is a viscous boundary condition, acting like a dashpot [7]. Other local, arbitrary order absorbing boundary conditions [8, 8, 9] and several other types of transmitting boundaries (e.g., infinite elements [10, 11]) have been proposed, but none of them is able to fulfill the radiation condition exactly.

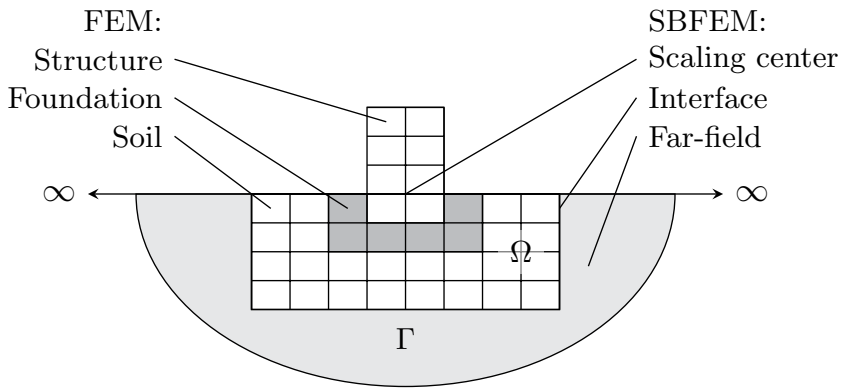


Figure 1: Problem definition.

Here a substructuring method is used (see Fig. 1) therefore the soil-structure system is spit up into a near-field and a far-field. Structure and foundation as well as parts of the surrounding soil are represent by the near-field Ω. The unbounded domain is represented by the far-field. The near-field can be easily discretized by finite elements the far-field is discretized by scaled boundary finite elements. Both methods are coupled at their common interface Γ. This discretisation assures that the radiation condition to infinity is fulfilled [2].

2 COUPLED FEM-SBFEM APPROACH

The equation of motion for the displacement-based FEM is given by

$$\mathbf{M}\ddot{\mathbf{u}} + \mathbf{C}\dot{\mathbf{u}} + \mathbf{K}\mathbf{u} = \mathbf{p}, \tag{1}$$

where $\mathbf{M}$ is the mass matrix, $\mathbf{C}$ is the damping matrix and $\mathbf{K}$ the stiffness matrix. The vectors $\mathbf{u}$, its first and second derivative represents displacement, velocity and acceleration. Supposing that period T can be divided into n time steps of constant size yields to a time step length $\Delta t = \frac{T}{n}$. Applying the Newmark scheme [12] to equation (1) leads to

$$\mathbf{M}\ddot{\mathbf{u}}_{n+1} + \mathbf{C}\dot{\mathbf{u}}_{n+1} + \mathbf{K}\mathbf{u}_n = \mathbf{p}_{n+1}\,, \tag{2}$$

with the update rules for displacement

$$\mathbf{u}_{n+1} = \mathbf{u}_n + \Delta t\dot{\mathbf{u}}_n + \left(\frac{1}{2} - \beta\right)\Delta t^2\ddot{\mathbf{u}}_n + \beta\Delta t^2\ddot{\mathbf{u}}_{n+1}\,, \tag{3}$$

and velocity

$$\dot{\mathbf{u}}_{n+1} = \dot{\mathbf{u}}_n + (1-\gamma)\,\Delta t\ddot{\mathbf{u}}_n + \gamma\Delta t\ddot{\mathbf{u}}_{n+1}\,. \tag{4}$$

The parameters β and γ of the time step integration scheme should be set as follows

$$\gamma \geq \frac{1}{2} \quad \text{and} \quad \beta \geq \frac{1}{4}\left(\gamma + \frac{1}{2}\right)^2\,. \tag{5}$$

In order to couple FEM and SBFEM the entries of the matrices in equation (1) have to be reordered

$$\begin{bmatrix} \mathbf{M}_{\Omega\Omega} & \mathbf{M}_{\Omega\Gamma} \\ \mathbf{M}_{\Gamma\Omega} & \mathbf{M}_{\Gamma\Gamma} \end{bmatrix}\ddot{\mathbf{u}} + \begin{bmatrix} \mathbf{C}_{\Omega\Omega} & \mathbf{C}_{\Omega\Gamma} \\ \mathbf{C}_{\Gamma\Omega} & \mathbf{C}_{\Gamma\Gamma} \end{bmatrix}\dot{\mathbf{u}} + \begin{bmatrix} \mathbf{K}_{\Omega\Omega} & \mathbf{K}_{\Omega\Gamma} \\ \mathbf{K}_{\Gamma\Omega} & \mathbf{K}_{\Gamma\Gamma} \end{bmatrix}\mathbf{u} = \begin{bmatrix} \mathbf{p}_{\Omega\Omega} \\ \mathbf{p}_{\Gamma\Gamma} \end{bmatrix} - \begin{bmatrix} \mathbf{0} \\ \mathbf{p}_b \end{bmatrix} \tag{6}$$

so that the block with the subscript "ΩΩ" contains all nodes located in the near-field while the block with subscript "ΓΓ" contains all nodes at the far-field-interface. The blocks with the subscripts "ΩΓ" and "ΓΩ" include the coupling information of near-field and far-field nodes. The vector $\mathbf{p}_b$ acts on the boundare Γ only. This additional force describes the response of the infinite half-space and can be applied to the near-field as a load.

The near-field is described by the FEM the far-field by the SBFEM. The forces acting at the interface of near-field and far-field are given by the convolution integral

$$\mathbf{p}_b(t) = \int_0^t \mathbf{M}^\infty(t-\tau)\ddot{\mathbf{u}}(\tau)d\tau\,, \tag{7}$$

where $\mathbf{M}^{\infty}(t)$ is the unit-impulse matrix. To solve the convolution integral (7) the unit-impulse matrices $\mathbf{M}_i^{\infty}$ are assumed to be constant within the time step Δt,

$$\mathbf{M}^{\infty}(t) = \begin{cases} \mathbf{M}_0^{\infty} & t \in [0; \Delta t]\,, \\ \mathbf{M}_1^{\infty} & t \in [\Delta t; 2\Delta t]\,, \\ \vdots & \vdots \\ \mathbf{M}_n^{\infty} & t \in [(n-1)\Delta t; n\Delta t]\ . \end{cases} \tag{8}$$

Due to this assumption and applying the time step integration scheme, equation (7) can be rewritten as

$$\mathbf{p}_b(t_n) = \gamma \Delta t \mathbf{M}_0^{\infty} \ddot{\mathbf{u}}_n + \sum_{j=1}^{n-1} \mathbf{M}_{n-j}^{\infty} \left(\dot{\mathbf{u}}_j - \dot{\mathbf{u}}_{j-1} \right). \tag{9}$$

The coupling of FEM and SBFEM is realized by simply adding equation (9) to the presorted FEM (6)

$$\begin{aligned} &\begin{bmatrix} \mathbf{M}_{\Omega\Omega} & \mathbf{M}_{\Omega\Gamma} \\ \mathbf{M}_{\Gamma\Omega} & \mathbf{M}_{\Gamma\Gamma} + \gamma \Delta t \mathbf{M}_0^{\infty} \end{bmatrix} \ddot{\mathbf{u}} + \begin{bmatrix} \mathbf{C}_{\Omega\Omega} & \mathbf{C}_{\Omega\Gamma} \\ \mathbf{C}_{\Gamma\Omega} & \mathbf{C}_{\Gamma\Gamma} \end{bmatrix} \dot{\mathbf{u}} + \\ &\begin{bmatrix} \mathbf{K}_{\Omega\Omega} & \mathbf{K}_{\Omega\Gamma} \\ \mathbf{K}_{\Gamma\Omega} & \mathbf{K}_{\Gamma\Gamma} \end{bmatrix} \mathbf{u} = \begin{bmatrix} \mathbf{p}_{\Omega\Omega} \\ \mathbf{p}_{\Gamma\Gamma} - \sum_{j=1}^{n-1} \mathbf{M}_{n-j}^{\infty} \left(\dot{\mathbf{u}}_j - \dot{\mathbf{u}}_{j-1} \right) \end{bmatrix}, \end{aligned} \tag{10}$$

so that the coupling is described in equation (10) completely.

3 IMPLEMENTATION

For the SSI analysis two programs are developed. The far-field is computed by ScaBo the coupled FEM-SBFEM analysis is computed by elPaSo [13, 14]. A more detailed description about the concept of ScaBo is published in [3]. Different third party libraries are used to achieve reasonable performance. ScaBo has to handle dense and sparse matrices as well, that is why ScaLAPACK builds the main core of this program [15]. Additional libraries like LAPACK, BLAS, BLACS, SLICOT, PLICOC (see Fig. 2(left)) are used to gain an optimal performance [16, 17, 18, 19].

The coupling of near-field and far-field is performed using elPaSo, here PETSc [20] builts the main core, hence sparse matrices are used. The libraries linkt to elPaSo are shown in Fig. 2 (right). To get access to direct solvers for linear equation systems MUMPS and SuperLU dist are also liked to elPaSo [21, 22, 23].

The communication is done by using the message passing interface (MPI) [24]. This allows to run both programs on distributed memory systems like clusters of computers and shared memory systems like multi core compute servers as well. This implementations guarantee a high flexibility regarding the utilized compute resources.

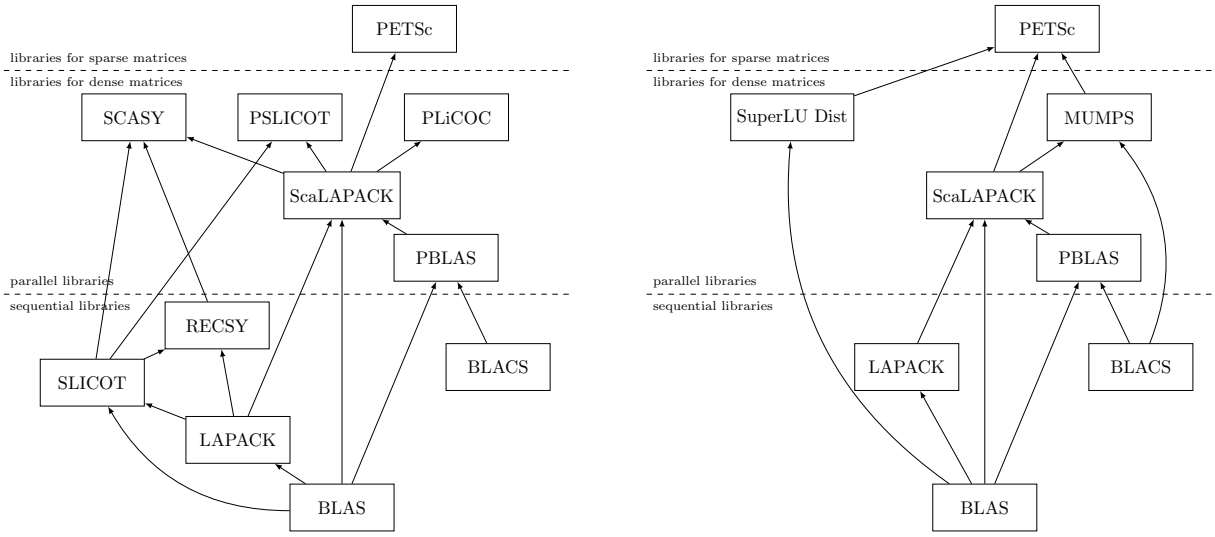


Figure 2: External libraries linked to ScaBo (left) elPaSo (right).

4 SETTLEMENT SIMULATION

In order to validate the implemented algorithms as illustrated in section 2 a settlement problem is carried out. Therefore two different discretisations of the near-field far-field interface are analysed. As already discussed in section 1, the problem has to be split into two separate domains. The near-field is represented by linear finite elements and the far-field by linear scaled boundary finite elements.

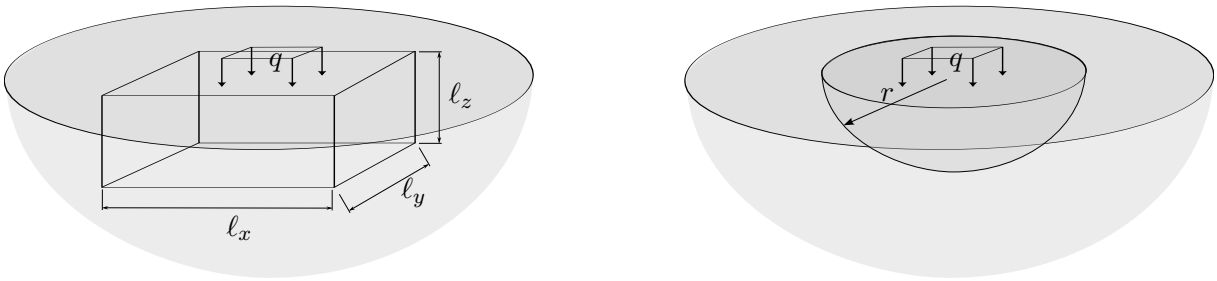


Figure 3: Infinite half space under constant area load q. (left) The computational domain is realized using a hemisphere. (right) The computational domain is realized using a cuboid.

Isotropic, homogeneous and fully linear elastic material is assumed for near-field and far-field as well. The material parameter are chosen as follows: Young's modulus $E = 21000$ [kNm^{-2}], Poisson ratio $\nu = 0.15$ and density $\rho = 2100$ [kgm^{-3}]. In this case a semi-analytical solution is available [25].

The numerical model is implemented for time domain analysis so that the displacements are time-dependent. As the coupled FEM-SBFEM approach fulfills the radiation condition, the nodal displacements become constant after a certain time, whenever con-

stant loads are applied. In those cases the numerical results approximate the static semi-analytical solution. To show the accuracy of the method we use two different meshing strategies, which are discussed in detail in the following sections.

One interface is discretised by a cubuid shaped mesh (CSM) see Fig. 3(left). The dimensions of the mesh are set to $\ell_x = \ell_y = 457.2$ [m] and $\ell_z = 190.5$ [m]. The second interface is discretised by a hemisphere shaped mesh (HSM) see Fig. 3(right). The distance between this centre and all interface nodes at the boundary Γ is exactly $r = 190.5$ [m]. The scaling centre is located in both cases in the middle of the loaded area, so that the smallest distance between scaling centre and boundary Γ is 190.5 [m]. On a square region of 152.4×152.4 [m^2] an area load $q = 70$ [kNm^{-2}] is applied. Table 1 summarizes the degrees of freedom (DoF) for the two discretised meshes.

Table 1: Discretisations with different number of DoF.

	$\mathrm{DoF}_{\mathrm{FEM}}$	$\mathrm{DoF}_{\mathrm{SBFEM}}$	$\Theta \frac{\mathrm{DoF}_{\mathrm{FEM}}}{\mathrm{DoF}_{\mathrm{SBFEM}}}$
CSM	14520	3720	25.6%
HSM	49155	3603	7.3%

The chosen material parameters yield to a longitudinal wave speed $c_p = 102.001$ ms^{-1}. So that the critical time step length is given by $\Delta t = \frac{r}{30 c_p} \approx 0.06$ s [26]. The parameters of Newmarks time step integration scheme are set to $\beta = 0.3025$ and $\gamma = 0.6$.

Evaluating the mentioned semi-analytical solution lead to a constant displacement $d(z = 0) = 0.560443138$ [m] in the centre of the loaded area directly at the near-fields surface. This solution is used to normalize the numerical time dependent solution as shown in Fig. 4.

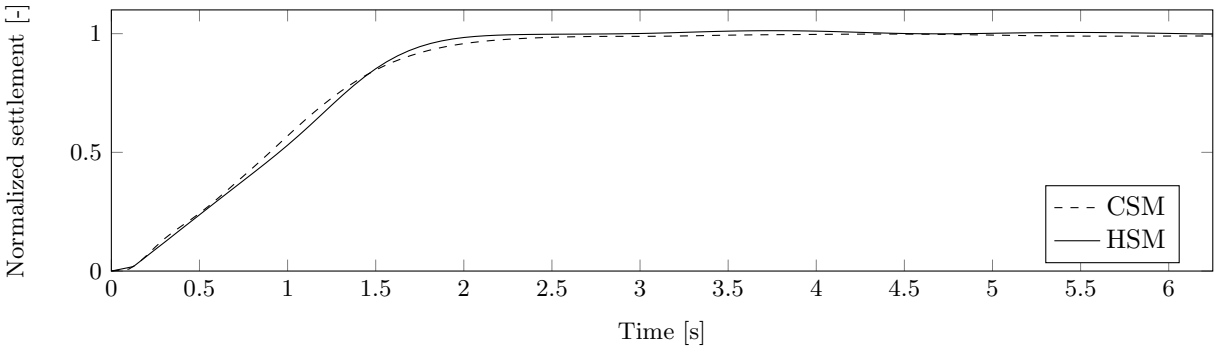


Figure 4: Time dependent normalized settlement.

5 FAR-FIELD MODEL REDUCTION

Computation of the unit-impulse response matrices $\mathbf{M}^{\infty}$ and the convolution integral eq. (7) is very expensive and memory consuming. Hence different model reductions are used to reduce the computational effort and memory consumption.

Since $\mathbf{M}^{\infty}$ increases after a certain time step t_m ($t_0 \leq t_m \leq t_n$), with a constant increment, not all n matrices have to be provided for the numerical simulation. For a simulation with n time steps, it is necessary to compute at least m matrices. All other matrices can be computed by a recursive algorithm [27].

Applying this model reduction to the meshes discussed in section 4 lead to the results as shown is Fig. 5. It is obvious that the used memory as well as the needed computation time decreases with the reduction of provided $\mathbf{M}^{\infty}$ matrices. As long as the number of $\mathbf{M}^{\infty}$ matrices is big enough the simulations relative error stays constant. If the number of precomputed matrices is too small the relative error increases. With $t_m = 50$ the relative error and also the solution is the same, compared to the reference solution without matrix extrapolation, but the memory usage is reduced to 32% and the run time is reduced to 37%.

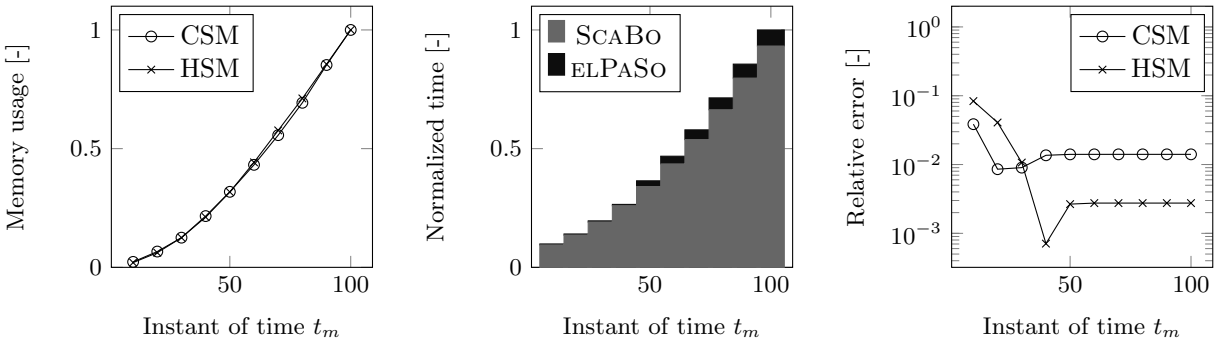


Figure 5: Memory usage (left), normalized compute time (mid) and relative error (right) when matrices $\mathbf{M}_t^{\infty}$, $t \leq m$ are extrapolated.

To make use of matrix extrapolation for model reduction, a recursive algorithm has to be implemented. Another approach without modifying the calculus is the geometrical decoupling. Therefore a threshold ϵ_z is introduced. This value is used to compare the entities in the unit-impulse response matrices. If the matrix entry is smaller than ϵ_z the influence of the corresponding nodes is supposed to be very small and so the value is not used for further computation. The influence of ϵ_z is studied within a range $10^{-8} \leq \epsilon_z \leq 10^{-2}$. The influence regarding the time used for computation is rather small, but the memory consumption is significant (see Fig.6). The fully populated matrices are computed and only at very end of the far-field computation the model is reduced. This reduced model is then used for later near-field far-field coupling, regarding the solutions relative error this model reduction is very robust.

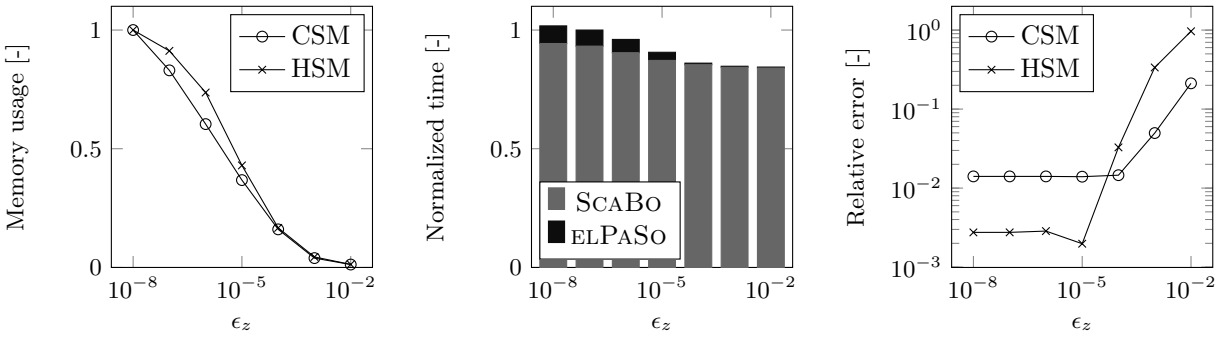


Figure 6: Memory usage (left), normalized compute time (mid) and relative error (right) when geometrical threshold ϵ_z is used.

The next step is consequently to decouple the far-field nodes not at the end of the far-field computation, but in the beginning. Hence the far-field is separated into an arbitrary number of substructures. Each substructure is solved separately. When the coupling is carried out, the substructures are combined with the near-field, so that the the complete far-field is considered. Memory usage, needed computational time and relative error are shown in Fig. 7. This reduction is very efficient in terms of memory and time consumption, but the relative error increases imminently.

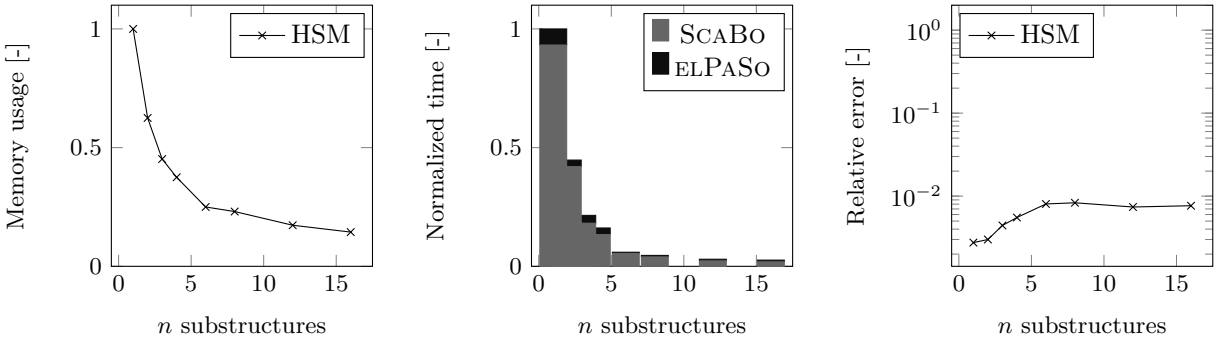


Figure 7: Memory usage (left), normalized compute time (mid) and relative error (right) when far-field is split up into substructures.

All three model reductions can be combined. The largest relative error of one method is dominating the other two. If substructuring is done carefully and the parameters t_m and ϵ_z are chosen in a valid range a major reduction of memory usage and computational time can be achieved. Here for instance if HSM is analysed with $t_m = 50$, $\epsilon_z = 10^{-5}$ and the far-field is divided into four substructures only 11% of memory is required and only 16.2% of the computational time. The solution is the same as the one which uses 4 substructures only.

6 PARALLEL PERFORMANCE

Desktop computer and even compute server are limited in memory and performance, hence the usage of compute cluster is a way to provide lots of memory and handle large numbers of unknowns in order to simulate realistic problems. As already mentioned both presented programs are designed to perform on distributed memory systems. In order to show the parallel performance a weak scaling is done. According to Amdahl [28] the time T needed to solve a certain problem can be split into one part which is going to be performed sequentially αT and another part which can be done in parallel $(1-\alpha)\,T$:

$$T = \alpha T + (1-\alpha)\,T\,. \tag{11}$$

Applying this to p processes lead to

$$T(p) = \alpha T(1) + \frac{(1-\alpha)\,T(1)}{p}\,. \tag{12}$$

Here the parallel performance of the coupled simulation is analysed. A detailed study of the parallel performance of ScaBo can be found in [3]. The weak scaling is conducted on a compute cluster of 15 nodes, each of them with 2 AMD Opteron 240 processors at 1.4 GHz and 3 GB of RAM, linked with a Myrinet-2000 interconnect. In all runs, one MPI rank per CPU is used (for instance, runs with 2 processes use only one node).

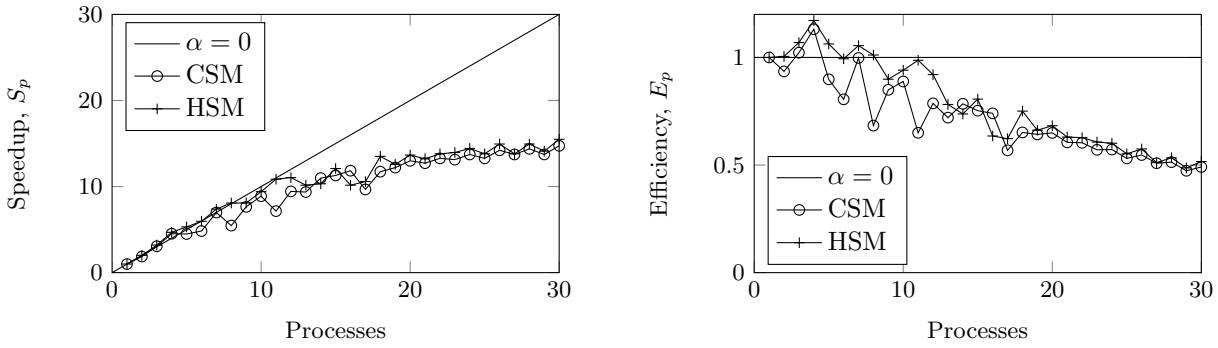


Figure 8: Parallel performence of the coupled simulation. Speedup (left), efficiency (right).

Fig. 8 displays the parallel speedup, which is defined as follows:

$$S_p = \frac{T(1)}{T(p)} \tag{13}$$

The parallel speedup is equal to the gain factor when the code runs with p processors. The plot shows an increasing of the speedup, which does not decay until 30 processes are employed. The speedup is not very close to the theoretical maximum $\alpha = 0$, but it shows a significant gain of about 15 with 30 processes. In terms of efficiency

$$E_p = \frac{S_p}{p} = \frac{T(1)}{p \cdot T(p)}\,, \tag{14}$$

which gives an indication of how well the code utilizes the available processors, a level of 50% is achieved in the worst case.

7 CONCLUSIONS

It is shown, that different model reductions techniques applied to the far-field can reduce the required memory and computational time without a leak of accuracy. Both, model reduction techniques and parallel computing can provide an answer to the high computational cost resulting from the application of SBFEM to very large scaled engineering problems. It is demonstrated that, in practice, when run on a moderately-sized cluster, the coupled FEM-SBFEM approach can yield the solution not only to academic benchmarks, but also to real problems. More complex, large-scale applications using this coupling approach will follow.

REFERENCES

[1] Zienkiewicz, O.C and Taylor, R.C, *The finite element method*. 4th Edition, Vol. 1, McGraw Hill, (1989).

[2] Wolf, J. and Song, C. *Finite-Element Modelling of Unbounded Media*, John Wiley & Sons, Chichester, (1996).

[3] Schauer, M; Roman, J.E.; Quintana-Ortí, E.S. and Langer, S. Parallel Computation of 3-D Soil-Structure Interaction in Time Domain with a Coupled FEM/SBFEM Approach. *Journal of Scientific Computing*, (2012), **52**:446–467.

[4] Antes, H. and Spyrakos, C. Soil-structure interaction. In: Beskos D and Anagnotopou los S. (eds) *Computer Analysis and Design of Earthquake Resistant Structures*, Computational Mechanics Publications, Southampton, (1997).

[5] Meskouris, K.; Hinzen, K.G.; Butenweg, C. and Mistler, M. *Bauwerke und Erdbeben - Grundlagen - Anwendung - Beispiele*. Vieweg+Teubner Verlag, Wiesbaden, (2007).

[6] Petersen C. *Dynamik der Baukonstruktionen*. Vieweg & Sohn Verlagsge- sellschaft mbH, Braunschweig/Wiesbaden, (2000).

[7] Lysmer, J. and Kuhlmeyer, R.L. Finite dynamic model for infinite media. *Journal of Engineering Mechanics*, **95**:859–875, (1969).

[8] Engquist, B. and Majda, A. Absorbing boundary conditions for the numerical simulation of waves. *Mathematics of Computation*, **31**:629–651, (1977).

[9] Liao, Z.P. and Wong, H.L. A transmitting boundary for the numerical simulation of elastic wave propagation. *Soil Dynamics and Earthquake Engineering*, **3**:174183, (1984).

[10] Bettess, P. *Infinite Elements*. Penshaw Press, Sunderland, U.K. (1992).

[11] Astley, R.J. Infinite elements for wave problems: a review of current formulations and a assessment of accuracy. *International Journal for Numerical Methods in Engineering*, **49**:951976, (2000).

[12] Newmark, N. A method of computation for structural dynamics. *Journal of Engineering Mechanics Division*, (1959), **85**:67–94.

[13] Schauer, M. and Roman, J.E. ScaBo Manual, Revision: 369, TU Braunschweig, Institut für Angewandte Mechanik (2010–2013), Institut für Konstruktionstechnik (2013), (2010–2014).

[14] Beck, S.; Clasen, D.; Lehmann, L.; Rurkowska, K.; Schauer, M. and Wulkau, M. elPaSo Manual, Revision: 433, TU Braunschweig, Institut für Angewandte Mechanik (2008–2013), Institut für Konstruktionstechnik (2013), (2008–2014).

[15] Blackford, L.S.; Choi, J.; Cleary, A.; DAzevedo, E.; Demmel, J.; Dhillon, I.; Dongarra, J.; Hammarling, S.; Henry, G.; Petitet, A.; Stanley, K.; Walker, D. and Whaley, R.C. *ScaLAPACK Users Guide*. Society for Industrial and Applied Mathematics, Philadelphia, (1997).

[16] Anderson, E.; Bai, Z.; Bischof, C.; Demmel, J.; Dongarra, J.; Croz, J.D.; Greenbaum, A.; Hammarling, S.; McKenney, A. and Sorensen, D. *LAPACK Users Guide*. Society for Industrial and Applied Mathematics, Philadelphia, (1992).

[17] Dongarra, J.J. and Whaley, R.C. *LAPACK working note 94: A users guide to the BLACS v1.1. Tech. Rep. UT-CS-95-281*, Department of Computer Science, University of Tennessee, (1995).

[18] Benner, P.; Quintana-Ortí, E.S. and Quintana-Ortí, G. Solving linear-quadratic optimal control problems on parallel computers. *Optimization Methods and Software*, **23**:879909, (2008).

[19] Guerrero, D.; Hernández, V. and Román, J.E. *Parallel SLICOT model reduction routines: The Cholesky factor of Grammians*. In: Proceedings of the 15th Triennal IFAC World Congress, Barcelona, Spain, (2002).

[20] Balay, S.; Buschelman, K.; Eijkhout, V.; Gropp, W.D.; Kaushik, D.; Knepley, M.; McInnes, L.C.; Smith, B.F. and Zhang, H. *PETSc users manual. Tech. Rep. ANL-95/11 - Revision 3.1*, Argonne National Laboratory, (2010).

[21] MUMPS Team. *MUMPS - a MUltifrontal Massively Parallel sparse direct Solver*, http://graal.ens-lyon.fr/MUMPS, (2012).

[22] Li, X.S. An Overview of SuperLU: Algorithms, Implementation, and User Interface. *Transactions on Mathematical Software*, **31**:302–325, (2005).

[23] Li, X.S.; Demmel, J.W.; Gilbert, J.R.; Grigori, I.L.; Shao, M. and Yamazaki, I. *SuperLU Users' Guide*, LBNL-44289, Lawrence Berkeley National Laboratory, (1999).

[24] MPI Forum (1994) *The message passing interface (MPI) standard*, http://www.mcs.anl.gov/mpi

[25] Harr, M.E. *Foundations of Theoretical Soil Mechanics*. McGraw-Hill Book Company, New York, (1966).

[26] Borsutzky, R. *Seismic Risk Analysis of Buried Lifelines*. Technische Universität Braunschweig, Braunschweiger Schriften zur Mechanik, 63, (2008).

[27] Lehmann, L. An effective finite element approach for soil-structure analysis in the time-domain, *Structural Engineering and Mechanics*, (2005), **21**:437–450.

[28] Amdahl, G.M. Validity of the single processor approach to achieving large-scale computing capabilities. *AFIPS Conference Proceedings*, (1967), **30**:483–485.

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis (Eds)

BOUNDARY ELEMENT SOLUTION OF 2D COUPLED PROBLEM IN ANISOTROPIC PIEZOELECTRIC FGM PLATES

MOHAMED ABDELSABOUR FAHMY*, †

* Jummum University College
Umm Al-Qura University
Alazizya, behind Alsalam Souq, 21955 Makkah, Saudi Arabia.
e-mail: maselim@uqu.edu.sa, Web page: https://uqu.edu.sa/staff/ar/4340548

† Faculty of Computers and Informatics
Suez Canal University
Old Campus, El-Sheikh Zayed, 41522 Ismailia, Egypt.
e-mail: Mohamed_fahmy@ci.suez.edu.eg
Web page: http://mohamed_fahmy_ci.staff.scuegypt.edu.eg

Key words: Plate; Piezoelectric; Stress; Anisotropic; functionally graded material; Dual reciprocity boundary element method.

Abstract. The mechanics of the piezoelectric functionally graded material (FGM) has received considerable research effort with their increasing usage in various applications including sensors and actuators, piezoelectric motors, reduction of vibrations and noise, infertility treatment and photovoltaics. It is hard to find the analytical solution of a problem in a general case, therefore, an important number of engineering and mathematical papers devoted to the numerical solution have studied the overall behavior of such materials. The time-stepping dual reciprocity boundary element method was proposed to solve the 2D coupled problem in anisotropic piezoelectric FGM plates. The accuracy of the proposed method was examined and confirmed by comparing the obtained results with those known previously.

1 INTRODUCTION

Numerical modelling of piezoelectric solids present certain difficulties since they exhibit not only electro-elastic coupling but anisotropic behaviour. Piezoelectric effect can only appear in crystals that lack of a centre of symmetry and that, as a consequence, are anisotropic. This anisotropy reduces in most cases to transversal isotropy. Piezoelectric ceramics are used for construction of sensors, transducers, actuators as well as adaptative structures. Lead zirconate titanate (PZT) is the most widely used piezoceramic. There are also piezopolymers as the polyvinilidene fluoride (PVDF). Owing to the coupling effects between mechanical and electric properties, piezoelectric materials (PMs) have found wide technological applications as sensors and actuators, piezoelectric motors, reduction of vibrations and noise, infertility

treatment and photovoltaics. Applications of piezoelectric materials as electro-mechanical devices have also stimulated a wide range of analytical researches [1-14]. It is usually difficult to obtain analytical solutions to problems involving finite solids or complex boundary conditions. Numerical methods, such as the finite difference method [15-19], the finite element method [20] and the boundary element method (BEM) [21-31] have also been applied to the analysis of electromechanical coupling under complicated conditions. Pan [32] derived the Green's functions for the anisotropic piezoelectric solids in an infinite plane, a half plane, and two joined dissimilar half-planes using the complex variable function method and presented a single-domain BEM analysis of 2D facture mechanics. Recent developments of 2D Green's functions and BEM analysis can be found in Refs. [33–37] for example. It is well known that the BEM presents significant advantages over other numerical techniques for the analysis of fracture mechanics problems. This fact has led to the publication of several BE approaches for the analysis of cracks in piezoelectric solids in the last few years. Presence of domain integrals in the formulation of the BEM dramatically decreases the efficiency of this technique. One of the most frequently used techniques for converting the domain integral into a boundary one developed through our paper is the so-called dual reciprocity boundary element method (DRBEM) [38-46].

In this article the DRBEM is used to solve the coupled problem in anisotropic piezoelectric FGM plates. In the case of two-dimensional, a numerical scheme for the implementation of the method is presented. The accuracy of the proposed method was examined and confirmed by comparing the obtained results with those known before.

2 FORMULATION OF THE PROBLEM

Here, we present the basic equations of the piezoelectric elasticity theory, which will be used for the solution of the problem described in the Introduction. With reference to a Cartesian coordinate system (x, y, z) as shown in Fig. 1. We shall consider a functionally graded anisotropic piezoelectric plate. The plate occupies the region $R = \{(x, y, z): 0 < x < \zeta, 0 < y < \Psi, 0 < z < \xi$ with graded material properties in the thickness direction.

In this paper, the FGM properties are graded along the thickness direction (x-direction) of the plate. The governing equations for the stress wave propagation in anisotropic functionally graded piezoelectric plate may be written in the following form

$$C_{fghi}u_{h,if} = \rho\ddot{u}_g - C_{fghi}\aleph u_{h,i} - e_{ifg}\left[\Phi_{,if} + \aleph\Phi_{,i}\right] \quad (1)$$

$$e_{fhi}u_{h,if} = \epsilon_{fi}\Phi_{,if} - e_{fhi}\aleph u_{h,i} + \epsilon_{fi}\left[\Phi_{,if} + \aleph\Phi_{,i}\right] \quad (2)$$

where σ_{fg} is the mechanical stress tensor, ε_{hi} is the strain tensor, u_h is the displacement vector, E_i is the electric field vector, Φ is the electric potential, D is the electric displacement, C_{fghi} is the elasticity tensor $\left(C_{fghi} = C_{gfhi} = C_{hifg}\right)$, e_{ifg} is the piezoelectric tensor $\left(e_{ifg} = eigf\right)$, ϵfi is the permittivity tensor $\epsilon fi = \epsilon if$, ρ is the density and $\aleph = mx + 1$.

A superposed dot denotes differentiation with respect to the time and a comma followed by a subscript denotes partial differentiation with respect to the corresponding coordinates.

3 NUMERICAL IMPLEMENTATION

Using the contracted notation, the governing equations (1) and (2) can be combined to form a single equation as follows:

$$L_{\mathrm{GH}} U_{\mathrm{H}} = \rho \delta_{GH} \ddot{U}_{\mathrm{H}} - \mathrm{B_G} \tag{3}$$

where

$$U_{\mathrm{H}} = \begin{cases} u_h & h = \mathrm{H} = 1,2,3 \\ \Phi & \mathrm{H} = 4 \end{cases} \tag{4}$$

$$\mathrm{Z_G} = \begin{cases} t_g & g = G = 1,2,3 \\ q & G = 4 \end{cases} \tag{5}$$

$$C_{f\mathrm{GH}i} = \begin{cases} C_{fghi}, & g = \mathrm{G} = 1,2,3; h = \mathrm{H} = 1,2,3 \\ \mathrm{e}_{ifg}, & g = \mathrm{G} = 1,2,3; \mathrm{H} = 4 \\ \mathrm{e}_{fhi}, & \mathrm{G} = 4; \quad h = \mathrm{H} = 1,2,3 \\ -\epsilon_{fi}, & \mathrm{G} = 4; \quad \mathrm{H} = 4 \end{cases} \tag{6}$$

$$\delta_{GH} = \begin{cases} \delta_{gh} & g = G = 1,2,3, \quad k = K = 1,2,3 \\ 0 & otherwise \end{cases} \tag{7}$$

$$L_{\mathrm{GH}} = C_{f\mathrm{GH}i} \frac{\partial}{x_i} \frac{\partial}{x_f} \tag{8}$$

$$\mathrm{B_G} = \begin{cases} C_{fghi} \bar{\bar{u}} + \mathrm{e}_{ifg} \, \bar{\bar{\Phi}}, & g = G = 1,2,3 \\ \epsilon_{fi} \Phi_{,if} - \mathrm{e}_{fhi} \bar{\bar{u}} + \epsilon_{fi} \bar{\bar{\Phi}}, & G = 4 \end{cases} \tag{9}$$

$$\bar{\bar{u}} = \aleph u_{h,\mathrm{i}}, \qquad \bar{\bar{\Phi}} = \Phi_{,if} + \aleph \Phi_{,i} \tag{10}$$

Now, we choose the fundamental solution U^*_{MH} as weighting function as follows

$$L_{GH} U^*_{MH}(x, \xi) = -\delta_{GM} \delta(x, \xi) \tag{11}$$

By integrating the weighted residual formula by parts twice we obtain the following piezoelectric reciprocity relation

$$\int_R \left(L_{\mathrm{GH}} \mathrm{U}_H U^*_{MG} - L_{\mathrm{GH}} U^*_{MH} \mathrm{U_G} \right) dR = \int_\Gamma \left(U^*_{MG} \mathrm{Z_G} - \mathrm{Z}^*_{MG} \mathrm{U_G} \right) d\Gamma \tag{12}$$

where

$$\mathrm{Z}^*_{MG} = C_{f\mathrm{G}Hi} U^*_{MH,i} n_f$$

By the use of sifting property, we obtain from equation (12) the piezoelectric integral representation formula

$$U_{\mathrm{M}}(\xi) = \int_\Gamma \left(U^*_{MG} \mathrm{Z_G} - \mathrm{Z}^*_{MG} \mathrm{U_G} \right) d\Gamma - \int_R U^*_{MG} \left(\rho \delta_{GH} \ddot{U}_{\mathrm{H}} - \mathrm{B_G} \right) dR \tag{13}$$

To transform the domain integrals into boundary integrals over the global boundary of the analyzed domain, the DRBEM can be applied to equation (13) to give the dual reciprocity representation formula of piezoelectric as

$$U_H(\xi) = \int_\Gamma \left(U^*_{HG} \mathrm{Z_G} - \mathrm{Z}^*_{HG} \mathrm{U_G} \right) d\Gamma + \sum_{q=1}^{N} \left(U^q_{HN}(\xi) + \int_\Gamma \left(T^*_{HG} U^q_{GN} - \mathrm{U}^*_{HG} T^q_{GN} \right) d\Gamma \right) \alpha^{\mathrm{q}}_{\mathrm{N}} \tag{14}$$

According to the steps described in Fahmy [43], the dual reciprocity boundary integral equation (14) can be written in the following system of equations

$$\bar{\zeta}\breve{U}(t) - \bar{\eta}\breve{T}(t) = \left(\bar{\zeta}\breve{\mho}(t) - \bar{\eta}\breve{\tau}(t)\right)\alpha(t) \tag{15}$$

where $\bar{\zeta}$, $\bar{\eta}$ are BEM system matrices, $\breve{U}$, $\breve{T}$ contain the nodal values of the generalized displacements and fluxes, and $\breve{\mho}$, $\breve{\tau}$ contain the particular solutions

The coefficient vector $\alpha_s(t)$ can be calculated by setting up a system of N equations from (15) using the point collocation procedure, which yields the system

$$M\breve{\ddot{U}}(t) + \bar{\zeta}\breve{U}(t) = \bar{\eta}\breve{T}(t) + \breve{\mathcal{B}}(t) \tag{16}$$

where the volume matrix V, piezoelectric mass matrix M and source vector $\breve{\mathcal{B}}(t)$ are as follows:

$$V = \left(\bar{\eta}\breve{\tau}(t) - \bar{\zeta}\breve{\mho}(t)\right)\mathcal{F}^{-1}, \quad M = \rho V, \quad \breve{\mathcal{B}}(t) = V\breve{B}(t). \tag{17}$$

The following matrix equation is obtained from Eq. (16).

$$\begin{bmatrix} M^{11} & M^{12} \\ M^{21} & M^{22} \end{bmatrix}\begin{bmatrix} \ddot{U}^{k} \\ \ddot{U}^{u} \end{bmatrix} + \begin{bmatrix} K^{11} & K^{12} \\ K^{21} & K^{22} \end{bmatrix}\begin{bmatrix} U^{k}(t) \\ U^{u}(t) \end{bmatrix} = \begin{bmatrix} \bar{\eta}^{11} & \bar{\eta}^{12} \\ \bar{\eta}^{21} & \bar{\eta}^{22} \end{bmatrix}\begin{bmatrix} T^{k}(t) \\ T^{u}(t) \end{bmatrix} + \begin{bmatrix} \mathcal{B}^{1}(t) \\ \mathcal{B}^{2}(t) \end{bmatrix} \tag{18}$$

The unknown fluxes $T^{u}(t)$ are obtained from the first row of matrix equation (18) and are expressed as follows.

$$T^{u}(t) = (\eta^{12})^{-1}\big[M^{11}\ddot{U}^{k}(t) + M^{12}\ddot{U}^{u}(t) + K^{11}U^{k}(t) + K^{11}U^{k}(t) + K^{12}U^{u}(t) - \bar{\eta}^{11}T^{k}(t) - \mathcal{B}^{1}(t)\big] \tag{19}$$

Making use of Eq. (19), we can write the second row of matrix equation (18) as

$$M^{u}\ddot{U}^{u}(t) + K^{u}U^{u}(t) = Q^{k}(t) \tag{20}$$

where

$$Q^{k}(t) = \bar{\mathcal{B}}^{k}(t) + \bar{\eta}^{k}T^{k}(t) - M^{k}\ddot{U}^{k}(t) - K^{k}U^{k}(t)$$
$$M^{u} = M^{22} - \bar{\eta}^{22}(\bar{\eta}^{12})^{-1}M^{12}$$
$$M^{k} = M^{21} - \bar{\eta}^{22}(\bar{\eta}^{12})^{-1}M^{11}$$
$$K^{u} = K^{22} - \bar{\eta}^{22}(\bar{\eta}^{12})^{-1}K^{12}$$
$$K^{k} = K^{21} - \bar{\eta}^{22}(\bar{\eta}^{12})^{-1}K^{11}$$
$$\bar{\eta}^{k} = \bar{\eta}^{21} - \bar{\eta}^{22}(\bar{\eta}^{12})^{-1}\eta^{11}$$
$$\bar{\mathcal{B}}^{k}(t) = \mathcal{B}^{2}(t) - \bar{\eta}^{22}(\bar{\eta}^{12})^{-1}\mathcal{B}^{1}(t)$$

We now split the system (20) into elastic and electric parts as follows:

$$\begin{bmatrix} M^{u}_{uu} & 0 \\ M^{u}_{\varphi u} & 0 \end{bmatrix}\begin{bmatrix} \ddot{u}^{u}(t) \\ \ddot{\varphi}^{u}(t) \end{bmatrix} + \begin{bmatrix} K^{u}_{uu} & K^{u}_{u\varphi} \\ K^{u}_{\varphi u} & K^{u}_{\varphi\varphi} \end{bmatrix}\begin{bmatrix} u^{u}(t) \\ \varphi^{u}(t) \end{bmatrix} = \begin{bmatrix} Q^{k}_{u}(t) \\ Q^{k}_{\varphi}(t) \end{bmatrix} \tag{21}$$

The unknown electric potential φ^{u} can be obtained from the second row of Eq. (21) as

$$\varphi^{u}(t) = \left(K^{u}_{\varphi\varphi}\right)^{-1}\left[Q^{k}_{\varphi}(t) - M^{u}_{\varphi u}\ddot{u}^{u}(t) - K^{u}_{\varphi u}u^{u}(t)\right] \tag{22}$$

With the aid of Eq. (22) into the first row of Eq. (21) we obtain

$$\bar{M}^{u}\ddot{u}^{u}(t) + \bar{K}^{u}u^{u}(t) = \bar{Q}^{k}(t) \tag{23}$$

where

$$\bar{Q}^{k}(t) = Q^{k}_{u}(t) - K^{u}_{u\varphi}\left(K^{u}_{\varphi\varphi}\right)^{-1}Q^{k}_{\varphi}(t)$$
$$\bar{M}^{u} = M^{u}_{uu} - K^{u}_{u\varphi}\left(K^{u}_{\varphi\varphi}\right)^{-1}M^{u}_{\varphi u}$$
$$\bar{K}^{u} = K^{u}_{uu} - K^{u}_{u\varphi}\left(K^{u}_{\varphi\varphi}\right)^{-1}K^{u}_{\varphi u}$$

We can write Eq. (23) at time step $n+1$

$$\bar{M}^u \ddot{u}^u_{n+1} + \bar{K}^u u^u_{n+1} = \bar{Q}^k_{n+1} \tag{24}$$

where

$$\bar{Q}^k_{n+1} = \bar{B}^k_{n+1} + \eta^k T^k_{n+1} - M^k \ddot{u}^k_{n+1} - K^k u^k_{n+1} \tag{25}$$

The displacements u_{n+1} and velocities $\dot{u}_{n+1}$ used in this algorithm are approximated at time step n+1 as follows:

$$\dot{u}_{n+1} \approx \dot{u}_n + [(1-\delta)\ddot{u}_n + \delta \ddot{u}_{n+1}]\Delta t \tag{26}$$

$$u_{n+1} \approx u_n + \dot{u}_n \Delta t + \left[\left(\frac{1}{2} - \alpha\right)\ddot{u}_n + \alpha \ddot{u}_{n+1}\right]\Delta t^2 \tag{27}$$

The acceleration at time step n+1 may be expressed from Equation (27) as:

$$\ddot{u}_{n+1} \approx \frac{1}{\alpha \Delta t^2}(u_{n+1} - u_n) - \frac{1}{\alpha \Delta t}\dot{u}_n - \left(\frac{1}{2\alpha} - 1\right)\ddot{u}_n \tag{28}$$

Upon substitution of (28) into (24) we obtain the following algebraic system

$$\mathbb{R} u^u_{n+1} = \mathcal{M}_{n+1} \tag{29}$$

where the stiffness matrix $\mathbb{R}$ and effective load vector $\mathcal{M}_{n+1}$ are given by

$$\mathbb{R} = \frac{1}{\alpha \Delta t^2}\bar{M}^u + \bar{K}^u \tag{30}$$

$$\mathcal{M}_{n+1} = \bar{Q}^k_{n+1} + \bar{M}^u\left[\frac{1}{\alpha \Delta t^2} u^u_n + \frac{1}{\alpha \Delta t}\dot{u}^u_n + \left(\frac{1}{2\alpha} - 1\right)\ddot{u}^u_n\right] \tag{31}$$

Once we have solved (29) for the unknown displacements at time step n+1, we can compute the accelerations and velocities from equations (28) and (26) respectively. Finally, the electric potential $\varphi^u(t)$ can be obtained from (22) and the unknown generalized tractions $T^u(t)$ can be determined using equation (19).

4 NUMERICAL RESULTS AND DISCUSSION

With the view of illustrating the numerical results, the material chosen for the plate is the piezoelectric ceramic Lead Zirconate Titanate (PZT), and the physical data for which is given as follows:

The elasticity tensor C_{fghi}, piezoelectric tensor e and relative permittivity ϵ^{rel}

$$C_{fghi} = \begin{pmatrix} 107.6 & 63.10 & 63.90 & 0.000 & 0.000 & 0.000 \\ 63.10 & 107.6 & 63.90 & 0.000 & 0.000 & 0.000 \\ 63.90 & 63.90 & 100.4 & 0.000 & 0.000 & 0.000 \\ 0.000 & 0.000 & 0.000 & 19.60 & 0.000 & 0.000 \\ 0.000 & 0.000 & 0.000 & 0.000 & 19.60 & 0.000 \\ 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 22.20 \end{pmatrix}$$

$$e = \begin{pmatrix} 0.00 & 0.00 & 0.00 & 0.00 & 12.0 & 0.00 \\ 0.00 & 0.00 & 0.00 & 12.0 & 0.00 & 0.00 \\ -9.6 & -9.6 & 15.1 & 0.00 & 0.00 & 0.00 \end{pmatrix}$$

$$\epsilon^{rel} = \begin{pmatrix} 1936 & 0.00 & 0.00 \\ 0.00 & 1936 & 0.00 \\ 0.00 & 0.00 & 2109 \end{pmatrix}$$

the results are plotted in Figs. 2–4 to show the validity of the DRBEM. These results obtained with the DRBEM have been compared graphically with those obtained using the Meshless Local Petrov–Galerkin (MLPG) method of Sladek et al. [47] are shown graphically in the same figures to confirm the validity of the proposed method. It can be seen from these figures that the DRBEM results are in excellent agreement with the results obtained by MLPG. The effects of the number of elements used were also examined. It was found that a further increase of boundary elements in the DRBEM led to improved numerical results (see Figures 2, 3 and 4).

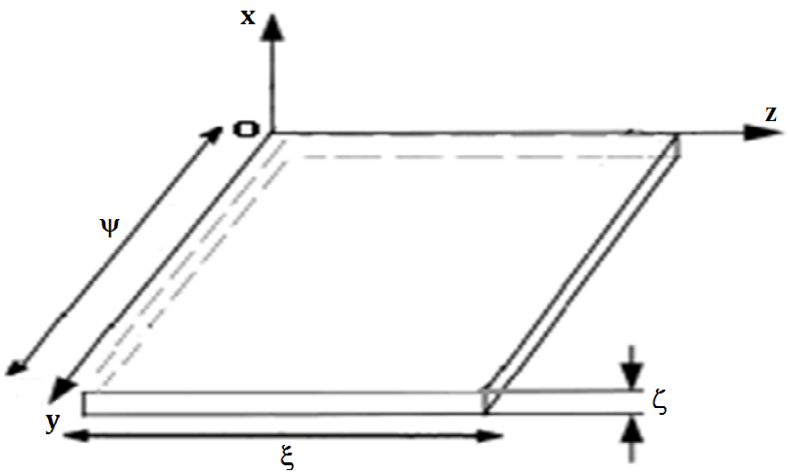


Fig. 1. The coordinate system of the plate.

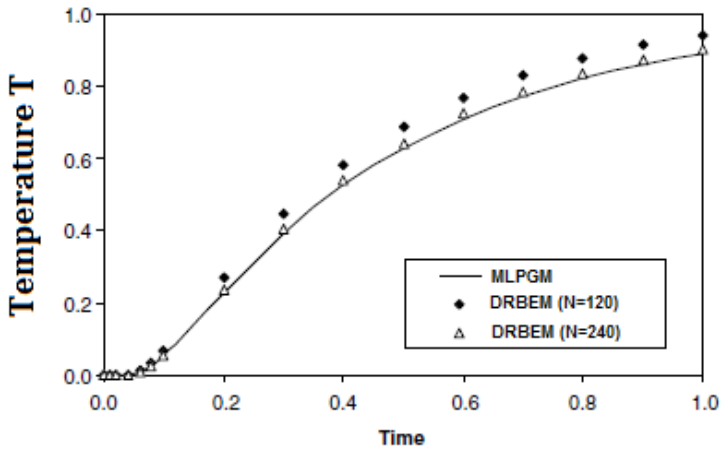


Fig. 2. Variation of the temperature T with time t for MLPGM and DRBEM

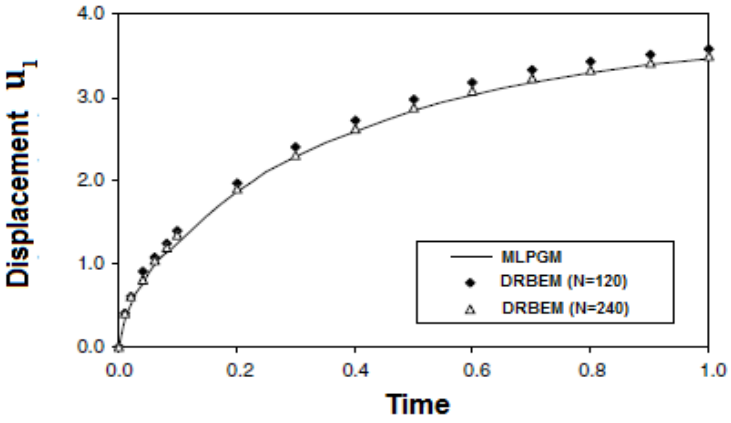


Fig. 3. Variation of the displacement u_1 with time t for MLPGM and DRBEM

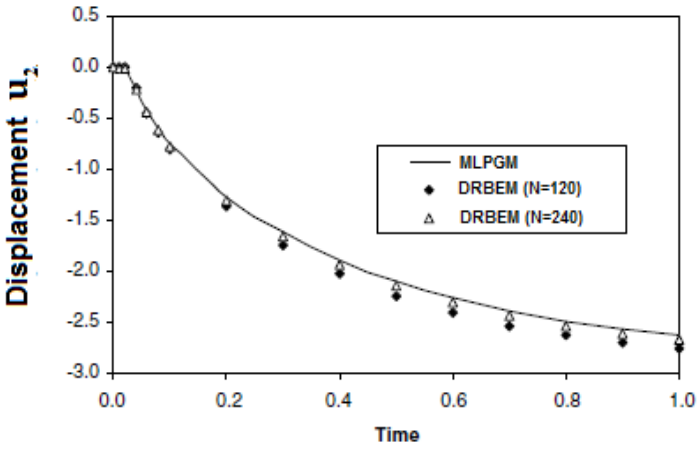


Fig. 4. Variation of the displacement u_2 with time t for MLPGM and DRBEM

REFERENCES

[1] Suo Z, Kuo CM, Barnett DM, Willis JR. Fracture mechanics for piezoelectric ceramics. Journal of the Mechanics and Physics of Solids 1992;40:739–765.
[2] Zhang N, Gao CF. Effects of electrical breakdown on a conducting crack or electrode in electrostrictive solids. European Journal of Mechanics-A/Solids 2012;32:62–68.
[3] Chen YH, Lu TJ. Cracks and fracture in piezoelectric materials. Advances in Applied Mechanics 2002;39:121–215.
[4] Abd-Alla, A.N., Alsheikh, F.A., Reflection and refraction of plane quasi-longitudinal waves at an interface of two piezoelectric media under initial stresses, Archive of Applied Mechanics 79, pp.843–857. (2009).
[5] Abd-Alla, A.N., Al-sheikh, F., Al-Hossain, A.Y., Effect of initial stresses on dispersion relation of transverse waves in a piezoelectric layered cylinder Materials Science and Engineering B: Solid-State Materials for Advanced Technology 162 pp.147–154, (2009).
[6] Abd-Alla, A.N.,Al-sheikh, F., The effect of the initial stresses on the reflection and transmission of plane quasi-vertical transverse waves in piezoelectric materials, World Academy of Science, Engineering and Technology, 50, pp.660-668 (2009).
[7] Suo Z, Kuo CM, Barnett DM, Willis JR. Fracture mechanics for piezoelectric ceramics. J Mech Phys Solids 1992:40:739–65.
[8] Pan EA. BEM analysis of fracture mechanics in 2D anisotropic piezoelectric solids. Eng Anal Bound Elem1999:23:67–76.
[9] Zhang TY, Zhao MH, Tong P. Fracture of piezoelectric ceramics. Adv Appl Mech 2002:38:147–289.
[10] Jin B, Zhong Z. A moving mode-III crack in functionally graded piezoelectric material: permeable problem. Mech Res Commun 2002:29:217–224.
[11] Lin S, Narita F, Shindo Y. Electroelastic analysis of a penny-shaped crack in a piezoelectric ceramic under mode I loading. Mech Res Commun 2003:30: 371–86.
[12] Fang DN, Wan YP, Soh AK. Magnetoelastic fracture of soft ferromagnetic materials. Theor Appl Fract Mech 2004:42:317–34.
[13] Kuna M. Fracture mechanics of piezoelectric materials – where are we right now? Eng Fract Mech 2010:77:309–26.
[14] Zhong XC, Zhang KS. Electroelastic analysis of an electrically dielectric Griffith crack in a piezoelectric layer. Int J Eng Sci 2010:48:612–23.
[15] Abd-Alla AM, El-Naggar AM and Fahmy MA, Magneto-thermoelastic problem in non-homogeneous isotropic cylinder, Heat and Mass Transfer, 2003:39:625-629.
[16] El-Naggar AM, Abd-Alla AM and Fahmy MA, The propagation of thermal stresses in an infinite elastic slab, Applied Mathematics and Computation, 2003:12:220-226.
[17] El-Naggar AM, Abd-Alla AM, Fahmy MA and Ahmed SM, Thermal Stresses in a rotating non-homogeneous orthotropic hollow cylinder, Heat and Mass Transfer, 2002:39:41-46.
[18] Fahmy MA, Thermal stresses in a spherical shell under three thermoelastic models using FDM, International Journal of Numerical methods and Applications, 2009:2:123-128.
[19] Fahmy MA, Finite difference algorithm for transient magneto-thermo-elastic stresses in a non-homogeneous solid cylinder, International Journal of Materials Engineering and Technology, 2010:3:87-93.

[20] Ha SK, Keilers C, Chang FK. Finite element analysis of composite structures containing distributed piezoceramic sensors and actuators. AIAA Journal 1992;30:772–780.
[21] Lee JS. Boundary element method for electroelastic interaction in piezoceramics. Eng Anal Bound Elements 1995;15:321–328.
[22] Abd-Alla AM, El-Shahat TM and Fahmy MA, Thermal stresses in a rotating non-homogeneous anisotropic elastic multilayered solids, Far East Journal of Applied Mathematics 2007:27: 223-243.
[23] Abd-Alla AM, El-Shahat TM and Fahmy MA, Effect of inhomogeneity on the thermoelastic stresses in micro-engineering anisotropic solid, Far East Journal of Applied Mathematics 2007:27:245-264.
[24] Abd-Alla AM, El-Shahat TM and Fahmy MA, Thermoelastic stresses in inhomogeneous anisotropic solid in the presence of body force, International Journal of Heat & Technology 2007:25:111-118.
[25] Abd-Alla AM, Fahmy MA and El-Shahat TM, Magneto-thermo-elastic stresses in inhomogeneous anisotropic solid in the presence of body force, Far East Journal of Applied Mathematics, 2007:27:499-516.
[26] Abd-Alla AM, Fahmy MA and El-Shahat TM, Transient piezothermoelastic stresses in a rotating non-homogeneous composite structure, Far East Journal of Applied Mathematics, 2007:27:489-497.
[27] Fahmy MA, Effect of initial stress and inhomogeneity on magneto-thermo-elastic stresses in a rotating anisotropic solid, JP Journal of Heat and Mass Transfer, 2007:1:93-112.
[28] Abd-Alla AM, Fahmy MA and El-Shahat TM, Magneto-thermo-elastic problem of a rotating non-homogeneous anisotropic solid cylinder, Archive of Applied Mechanics 2008:78:135-148.
[29] Fahmy MA, Thermoelastic stresses in a rotating non- homogeneous anisotropic body, Numerical Heat Transfer, Part A: Applications, 2008:53:1001-1011.
[30] Fahmy MA and El-Shahat TM, The effect of initial stress and inhomogeneity on the thermoelastic stresses in a rotating anisotropic solid, Archive of Applied Mechanics, 2008:78:431-442.
[31] Fahmy MA and Salama SA (2010), Boundary element solution of steady-state temperature distribution in non-homogeneous media, Far East Journal of Applied Mathematics 2010:43:31-40.
[32] Pan EN. A BEM analysis of fracture mechanics in 2D anisotropic piezoelectric solids. Engng Anal Bound Elem 1999;23:67–76.
[33] Denda M, Lua J. Development of the boundary element method for 2D piezoelectricity. Compos, Part B-Engng 1999;30:699–707.
[34] Qin QH, Meng L. BEM for crack-inclusion problems of plane thermopiezoelectric solids. Int J Numeric Meth Engng 2000;48:1071–1088.
[35] Qin QH. Thermoelectroelastic analysis of cracks in piezoelectric halfplane by BEM. Comput Mech 1999;23:353–360.
[36] Gao CF, Wang MZ. Green's functions of an interfacial crack between two dissimilar piezoelectric media. Int J Solids Struct 2001;38: 5323–5334.
[37] Liu Y, Fan H. Analysis of thin piezoelectric solids by the boundary element method. Comput Methods Appl Mech Engng 2002;191:2297–2315.

[38] Fahmy MA, Application of DRBEM to non steady-state heat conduction in non-homogeneous anisotropic media under various boundary elements, Far East Journal of Mathematical Sciences, 2010:43:83-93.
[39] Fahmy MA, Transient magneto-thermoviscoelastic plane waves in a non-homogeneous anisotropic thick strip subjected to a moving heat source, Applied Mathematical Modelling, 2012:36:4565-4578.
[40] Fahmy MA, The effect of rotation and inhomogeneity on the transient magneto-thermoviscoelastic stresses in an anisotropic solid, ASME Journal of Applied Mechanics, 2012:79:1015.
[41] Fahmy MA, Transient magneto-thermo-viscoelastic stresses in a rotating nonhomogeneous anisotropic solid with and without a moving heat source, Journal of Engineering Physics and Thermophysics 2012:85:950-958.
[42] Fahmy MA, Transient magneto-thermo-elastic stresses in an anisotropic viscoelastic solid with and without moving heat source, Numerical Heat Transfer, Part A: Applications 2012:61:547-564.
[43] Fahmy MA, Implicit-Explicit time integration DRBEM for generalized magneto-thermoelasticity problems of rotating anisotropic viscoelastic functionally graded solids, Engineering Analysis with Boundary Elements 2013:37:107-115.
[44] Fahmy MA, Generalized magneto-thermo-viscoelastic problems of rotating functionally graded anisotropic plates by the dual reciprocity boundary element method, Journal of Thermal Stresses 2013:36:1-20.
[45] Fahmy MA, A three-dimensional generalized magneto-thermo-viscoelastic problem of a rotating functionally graded anisotropic solids with and without energy dissipation, Numerical Heat Transfer, Part A: Applications 2013:63:713-733.
[46] Fahmy MA, A Computerized DRBEM model for generalized magneto-thermo-visco-elastic stress waves in functionally graded anisotropic thin film/substrate structures, Latin American Journal of Solids and Structures 2014:11:386-409.
[47] J. Sladek, V. Sladek, P. Stanak, C. Zhang, M. Wünsche, Analysis of the bending of circular piezoelectric plates with functionally graded material properties by a MLPG method, Engineering Structures 2013:47:81–89.

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis (Eds)

COUPLED NUMERICAL MULTIPHYSICS SIMULATION METHODS IN INDUCTION SURFACE HARDENING

DIRK SCHLESSELMANN*, BERNARD NACKE*, ALEXANDER NIKANOROV* AND SERGEY GALUNIN†

* Institute of Electrotechnology (ETP)
Leibniz Universität Hannover
Wilhelm-Busch-Str. 4, D-30167 Hannover, Germany
e-mail: etp@etp.uni-hannover.de, www.etp.uni-hannover.de

† Department of Electrotechnology and Converter Engineering
Electrotechnical State University St. Petersburg
Prof. Popov Str. 5, 197376 St. Petersburg, Russia
email: sagalunin@gmail.com

Key words: Induction Surface Hardening, Induction Heating, Numerical Simulation, Coupled Problems, Multiphysics Problems

Abstract. Numerical simulation is a valuable tool to help investigate complex multiphysics problems of engineering and science. This also applies to inductive surface hardening with its coupled electromagnetic and temperature fields as well as the microstructure changes of the hardened material. In this field, numerical simulation is a well-established approach for effective process design. This is particularly true since an analytical approach usually fails because of the complexity of the problems. Also, experiments oftentimes are not leading to a solution in an acceptable period of time because of the big number of process parameters. Furthermore, numerical simulation can help to investigate effects that could not have been observed otherwise. An example is the Joule heat distribution within a heated work piece during inductive heating. However, the fields of application as well as the methods of numerical simulation have to keep pace with technological progress. Two examples of new applications and methods for numerical simulation in induction hardening are presented in this paper: A complex 3D model of a large bearing and a new approach for the numerical simulation of the martensite microstructure.

1 INTRODUCTION

In the first part of this paper, an application-oriented strategy to numerically model and investigate scan hardening processes for large bearings is introduced. It will be shown how the 3D model can be applied to different hardening setups by a user without deeper knowledge of the numerical software. The advantages of using the model for the development of complex inductor geometries compared to an experiment based approach will be pointed out. The model calculates the temperature profile within a work piece, which is shown exemplarily for an inner ring of a main bearing used in wind power systems.

In the second part of the paper, a new numerical model for calculating martensite microstructure in induction surface hardening processes is introduced. It takes into account the heating as well as the quenching process and uses the temperature history of a work piece to calculate martensite formation. The calculation is based on an empirical equation found by Koistinen and Marburger. A comparison between the heat distribution within a work piece at the end of the heating process and the distribution of martensite after quenching is performed for different process parameters. Thus, it is determined, in which case the temperature distribution is sufficient to predict the hardened layer and in which case the microstructure has to be calculated to receive accurate results. The model is verified by comparing simulation results with experiments.

2 NUMERICAL 3D MODELLING AND VERIFICATION OF INDUCTION SURFACE HARDENING PROCESSES FOR LARGE BEARINGS

The rapid growth of renewable energy all over the world goes hand in hand with technological progress in this field. This also applies to wind power systems with their constantly increasing size and nominal output power of recently up to 8 MW (Vestas V164-8.0). The main bearings of such systems usually have a diameter of several meters. To prevent the bearings from wearing off, their running surfaces have to be hardened. A valid and often used process for this is inductive scan hardening. The process development requires a big financial effort as well as a lot of time because of the size of the bearings and the complexity of the inductors (Figure 1). Therefore, destructive material testing as well as design adaptions of the inductor have to be minimized. This can be achieved with the help of numerical simulation. A numerical 2D model for hardening of the main bearing of a wind power system can be found in [1]. However, to fully understand and calculate the inductive heating process accurately, a full 3D model is required necessarily.

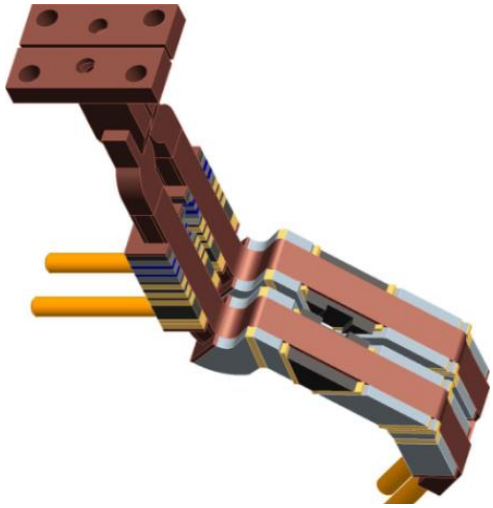


Figure 1: Scanning inductor

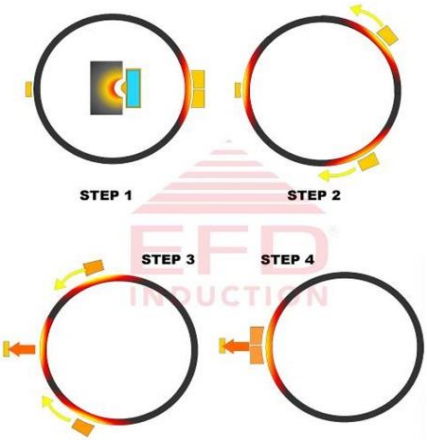


Figure 2: Hardening process

It is crucial for the quality of the hardening process that no soft zones occur. This means an even and seamless hardening profile along the circumference of the bearing is required. To achieve this, the hardening process is divided into four steps which can be seen in Figure 2. In step 1, two scanning inductors remain stationary and heat up the work piece until hardening temperature is reached. After this, for step 2, quenching is initiated and both inductors are moving counter wise along the circumference to harden the work piece simultaneously. With some delay,

depending on the size of the bearing and the feeding speed, a third inductor starts to preheat the zone where the scanning inductors meet again. In step 3, the third inductor is removed right before the scanning inductors reach this final zone. When the scanning process stops the inductors remain motionless and heat up the preheated zone to hardening temperature before they are removed as well. After quenching this area, step 4 and hence the full hardening process is completed [2]. In this paper, only the scanning phase in step 2 will be investigated. It can be regarded as a quasi-stationary process in terms of the temperature profile within the work piece.

2.1 Development of the numerical model

For all calculations the FEM software package ANSYS is used. The algorithm applied to calculate the heating process is shown in Figure 3. Before the calculation starts, the geometry and the mesh are created. Also, a matrix containing the coordinates of all nodes of the work piece is stored. This is required to take into account the relative motion of the inductor. The idea is to move the temperature profile within the work piece along the direction of feeding as explained in [3]. The elements along the direction of movement can be of different size. In order to achieve this, temperatures, usually calculated at each node of an element, are interpolated if necessary.

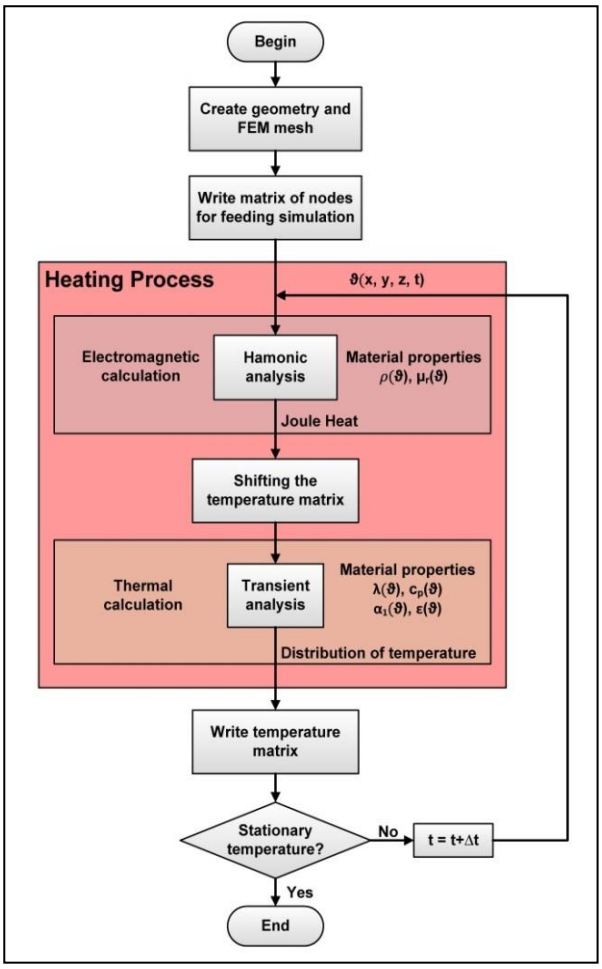


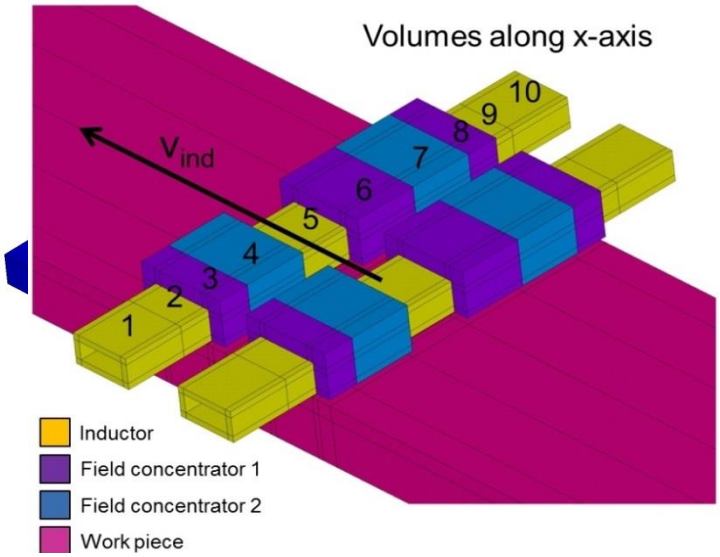
Figure 3: Calculation algorithm

The heating process is divided into sufficiently small time steps and performed as a coupled harmonic-electromagnetic and transient-thermal calculation. The material properties are adjusted after each time step according to the recent temperature within the elements. If a stationary temperature profile is reached, the calculation ends. The result is the quasi-stationary 3D temperature distribution within the work piece. The calculation always starts with the whole system at room temperature. The transient transition period till the stationary state is reached has no relevance and physical significance since the real starting process is not part of the calculation.

The general idea for the 3D model of a complex large scale inductor and bearing is to create the system's geometry based on layers. All layers consist of the same number of geometric points. Their spatial coordinates are stored in simple text files, one for each layer. This way, the user is able to change the geometry by simply editing the text files without

having to use ANSYS. A layer does not have to be plane; the points defining it can differ in all three dimensions. The number of layers is unlimited in principal. The process of building the model from the bottom up at the start of a calculation is automated and does not require any action by the user. The points within each layer are used to define lines and areas. By finally connecting the layers, a 3D volume model is created. However, all geometries investigated have to have the same principal topology, which will be explained further below.

To check if the automated generation of geometry and the calculation are working correctly, a simple setup was implemented and tested (Figure 4). Only the conductors, their field concentrating materials and the work piece are depicted, the surrounding air is suppressed. The setup consists of a flat steel work piece and two current carrying conductors. In this case, the geometry is not varying along the conductors and all points of a layer lie within the same plane. The 11 different layers were created by simply changing the x-coordinate of their points. The amount of volumes between the layers sums up to 81. This topology can only be changed by editing the scripts for the automated creation of geometry. However, almost any setup with two conductors and field concentrators above a work piece can be modelled, which is the typical design approach for scanning inductors of large bearings.



Volume	Field conc. left	Field conc. right	Work piece
1	4	4	4
2	4	4	4
3	2	2	1
4	3	3	1
5	4	4	1
6	2	2	1
7	3	3	1
8	2	2	1
9	4	4	4
10	4	4	4

1 – Work piece
2 – Field concentrator 1
3 – Field concentrator 2
4 – Air (not shown)

Figure 4: Simple test setup

After creating the volumes and the elements of the model, material properties have to be assigned. To keep this procedure as simple as possible, a text file containing the required information is used. The file contains a matrix as shown on the right side of Figure 4: There is a line for each volume along the inductor and four rows. Two rows are used to specify the flux concentrating material of the left and right conductor, one defines if the work piece is present in this layer of volumes. The numbers within the matrix are pointing to the properties of the material in a database. This database can be edited at any time and new materials can be added. The material database and the assignment are completely independent of the geometry. This is a key aspect of the model: It is possible to test different field concentrator variants with the same inductor geometry by only editing the matrix of material parameters.

2.2 Experimental setup

Figure 5: Experimental setup inner ring

The hardening of an inner ring of a large bearing can be seen in Figure 5. A numerical investigation is especially interesting for the inner ring since the design of the inductor is by far more complex compared to the inductor of the outer ring of the bearing. The process used a frequency of 3.8 kHz, the overall system power was about 160 kW. A feeding speed of 5 mm/s was chosen for the inductor. The quenching device is not part of the numerical model since only the temperature distribution within the work piece is analysed in this investigation.

2.3 Verification

The volumetric model of the inductor and the inner ring from the experiment are shown in Figure 6. It consists of the inductor, including all field concentrators, the inner ring and the surrounding air, which again is suppressed in this picture. Additional constructive elements of the inductor are not included in the model since they are irrelevant for the magnetic field distribution. The conductors are not straight, which means that some of the 44 layers that were defined for this model have points in more than one plane. To limit the size of the model, only the volume of the inner ring, which is under influence of the electromagnetic field, is implemented. The curvature of the bearing is relatively small because of its large diameter and is hence neglected. This means that the inner ring in the model is straight with regard to the feeding direction of the inductor. All in all, the model consists of 202,288 elements, which results in a calculation time of some 24 hours till the steady temperature state is reached if a standard PC is used. Thus, a new flux concentrator configuration can be tested within one

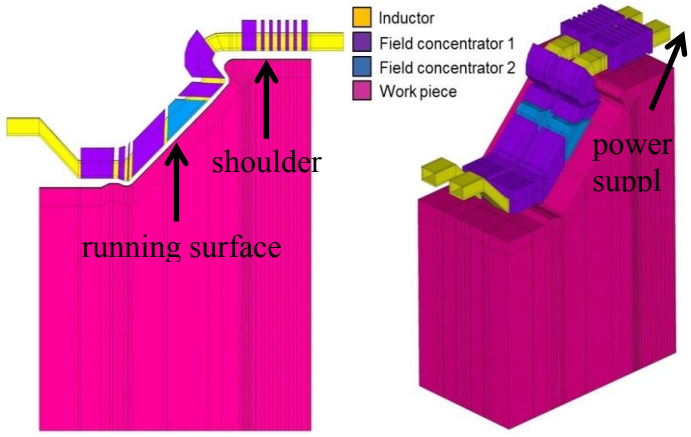


Figure 6: Volumetric model of the experimental setup

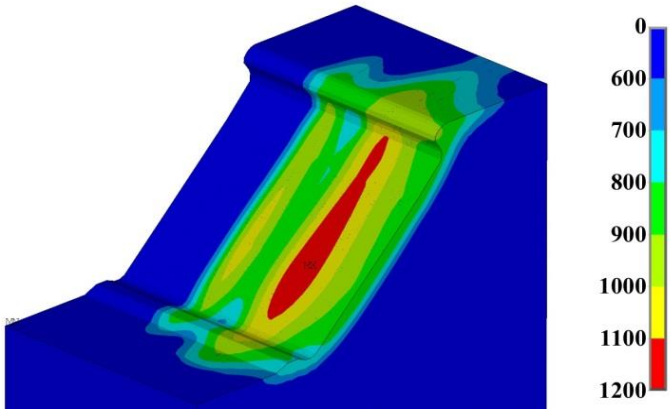


Figure 7: 3D quasi-steady temperature profile

day. For an experiment, field concentrating material would have to be removed and replaced, which is by far more time consuming and expensive.

The model is verified by comparing microsections of a hardened work piece with the quasi-steady temperature profile. The temperature for austenitization is about 850°C in this process and was derived from temperature measurements during experiments. A surface hardening depth (SHD) of about 6 mm is required along the running surface of the inner ring. Usually, the accurate calculation of an SHD of several millimetres has to take into account a calculation of martensite [3]. However, the temperature profile is sufficient to evaluate the hardened zone qualitatively.

Figure 7 shows the quasi-steady temperature profile. The maximum of 1185°C occurs at the edge of the second conductor with regard to the scanning direction where the heated work piece leaves the inductor.

A microsection and a cross-section of the temperature profile are compared in Figure 8. The cross-section is taken from where the maximum temperature occurs. The grey color of the temperature profile shows the area where austenitization was reached. The calculated SHD along the shoulder is smaller in comparison to the microsection. This can be explained by taking a closer look at the gap between inductor and work piece. In this area the gap was 4 mm in the experiment as opposed to 5 mm in the simulation. Furthermore the power connection part of the inductor is not fully implemented in the model. These deviations result in less power being induced in the shoulder area. However, the model is verified with the help of the running surface: Temperature profile and microsection show good agreement since the numerical model is accurate there.

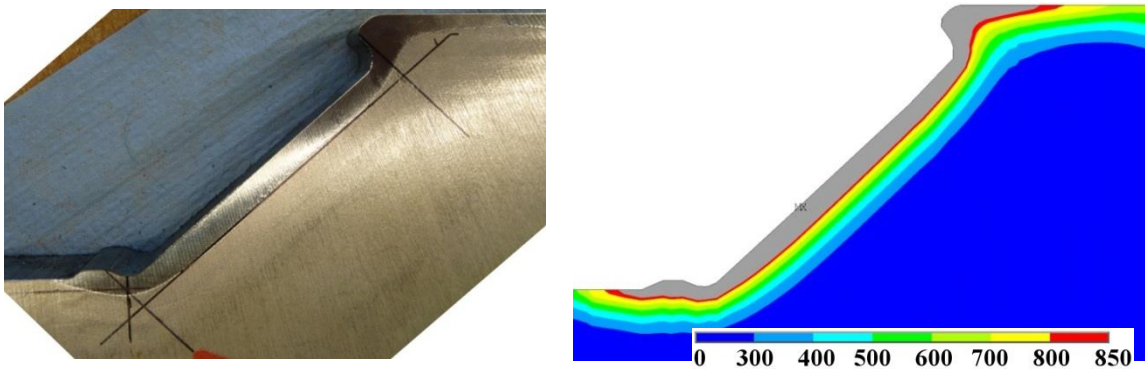


Figure 8: Comparison of microsection and temperature profile

2.4 Conclusions

The model offers an easy way to numerically investigate induction hardening processes of large bearings. The data required for the automated creation of the complex geometry as well as all information about the position and type of flux concentrators is stored in simple text files. This way, no deeper knowledge of the simulation software and calculation algorithm is required to apply the model, given that the topology of the geometry remains unchanged. Results for a new flux concentrator of an existing setup or a modified inductor shape are available within one day. The model was verified by comparing its results to microsections and will be used to effectively develop hardening processes. The costly repetitive process of redesigning the inductor and conducting experiments can be avoided that way.

3 NUMERICAL CALCULATION OF MARTENSITE MICROSTRUCTURE IN INDUCTION SURFACE HARDENING PROCESSES

In the second part of the paper, a new numerical model for calculating martensite microstructure in induction surface hardening processes is introduced. It takes into account the heating as well as the quenching process and uses the temperature history of a work piece to calculate martensite formation. The calculation is based on an empirical equation found by Koistinen and Marburger [4]. Before only coupled electromagnetic and thermal models have been used for the investigation and design of complex induction hardening processes for several years. Instead of calculating martensite formation during quenching, the models have used the temperature distribution at the end of the heating process to predict the microstructure within the work piece [3]. This approach is valid for surface hardening depths which do not exceed a few millimetres. If bigger hardening depths have to be investigated, heat transfer from the surface of the work piece to its core becomes increasingly important. In this case, the temperature distribution might lead to inaccurate predictions of the hardened profile. Furthermore, heat transfer has to be considered, if there is a delay between heating and quenching. This undesirable situation occurs in many industrial hardening processes. To investigate situations as described above, a new algorithm for computing martensite microstructures is applied. The aim is to determine, if there is a critical depth, which requires a martensite calculation. In addition, situations are identified, in which a delay of quenching cannot be neglected. For this, the influence of different quenching situations on the critical depth is investigated.

3.1 Algorithm for calculating martensite microstructures

Subsequent to heating, the quenching process has to be taken into consideration. Therefore, a thermal calculation based on the temperature distribution at the end of the heating process has to be performed. The heat transfer is represented by a coefficient as a function of temperature, which is assigned to all locations on the work piece's surface quenched with cooling agent. Coefficients for different agents and parameters like pressure and flow rate can be found in respective literature [6].

For calculating martensite microstructures the algorithm shown in Figure 9 has to be performed for every node of a FEM model. The algorithm starts with an analysis of each node's temperature

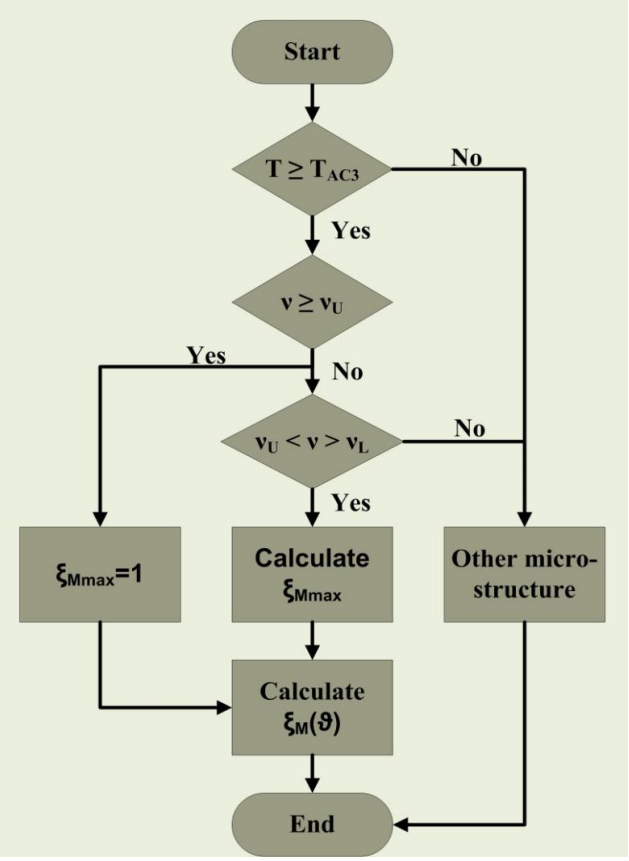


Figure 9: Algorithm for calculating martensite

history. Two conditions have to be met necessarily to receive a fully martensitic microstructure: First, the austenitizing temperature T_{AC3} has to be reached during heating. This is required for a complete austenitic transformation. Secondly, the upper critical cooling rate v_U has to be reached during quenching. Curve 1 in Figure 10 shows a typical time-temperature curve of an inductive hardening process including heating and quenching, which fulfills these conditions. In this case, diffusion of carbon and iron is not possible and microstructures other than martensite cannot form. The fraction of martensite ξ_M can be calculated according to an equation of Wildau and Hougardy [7]:

$$\xi_M(\vartheta) = \left[1 - \exp\left(-c_1(T_{MS} - \vartheta)^{c_2}\right)\right] * 100\% \quad (1)$$

(1) is based on an empirical approach by Koistinen and Marburger [4]. By introducing an additional constant c_2 and taking into account that c_1 and c_2 strongly depend on the material's carbon content, Wildau and Hougardy suggest an equation suitable for a wide range of steels.

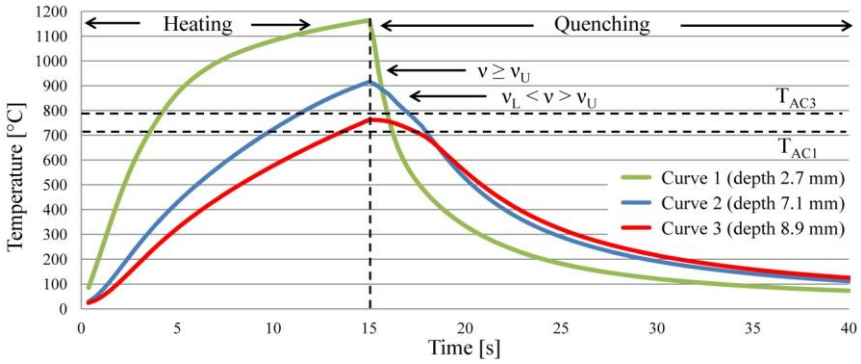


Figure 10: Typical time-temperature curve

The martensite fraction is only a function of the material's temperature ϑ after quenching. The formation of martensite begins below the martensite starting temperature T_{MS}. If the speed of quenching is between v_U and the lower critical cooling rate v_L, diffusion of carbon and iron occurs. Therefore the resulting microstructure contains bainite, ferrite and perlite besides martensite. An example of such time-temperature profiles is given by Curve 2 in Figure 10. (1) has to be extended with a factor ξ_{Mmax}, which represents the maximum martensite fraction at a certain cooling speed v.

$$\xi_M(\vartheta) = \xi_{Mmax}(v) * \left[1 - \exp\left(-c_1(T_{MS} - \vartheta)^{c_2}\right)\right] * 100\% \quad (2)$$

This factor is unity for cooling rates higher than v_U and zero for rates lower than v_L where martensite formation does not occur. To determine ξ_{Mmax}, the contents of bainite ξ_B, ferrite ξ_B and perlite ξ_B are calculated. The residual fraction of martensite is then given by [8]:

$$\xi_{M\_max} = 1 - (\xi_B + \xi_B + \xi_B) \quad (3)$$

Several authors suggest using a semi-empirical approach by Johnson, Mehl and Avrami or Leblond to calculate microstructures occurring in diffusion processes [9-11]. Here, data that is derived from time-temperature transformation (TTT) and continuous cooling transformation (CCT) diagrams or determined by experiments is used to describe the formation. This approach is applied in simulations of many different processes, for example precision forging [12]. However, this paper is focused on calculating the martensite content within a relatively thin layer in induction surface hardening applications. Therefore, a more direct way to determine ξ_{Mmax} is suggested. By using a TTT diagram, ξ_{Mmax} can be determined for different quenching rates. It was found that the relation can be described as:

$$\xi_{Mmax}(v) = 1 - \exp\left[-(v - c_5)/c_6\right] \quad (4)$$

For quenching rates lower then v_U first fractions of martensite occur at lower temperatures T_{MS}. Using data from the same TTT diagram T_{MS} is given as:

$$T_{MS}(v) = -c_3 * \ln(v) + c_4 \quad (5)$$

Both empirical equations are determined by using the method of least squares and show good agreement with data derived from TTT diagrams of various steels. The suggested approach can only be used for quenching processes typical for induction surface hardening. More complex temperature time profiles require the approach of Johnson, Mehl and Avrami or Leblond. Furthermore, the direct approach does not take into consideration that austenite formation already starts at temperatures above T_{AC1} [13]. The calculation only takes into account fully austenitized areas of material.

3.2 Investigation using a basic 1D model

The numerical model used for all basic investigations is one-dimensional and represents the infinite half space. It consists of a current fed inductor 2 mm above a work piece. Carbon steel C45 and alloyed steel 42CrMo4 are investigated exemplarily with their different material parameters [14]. To cover a wide range of hardening depths, frequencies of 62 kHz and 3 kHz are used in the simulation. In addition, a depth d_{TH} (seen from the surface of the work piece) is defined for each simulation. The hardening temperature T_H has to be reached within this depth, meaning that the material is fully austenitized. A constant value T_H, sufficient for the shortest heating time, is assumed to allow better comparability. However, T_H is varied for both frequencies because the initial heating time differs significantly. The parameters for investigating quenching are the delay time t_{delay} between heating and quenching and the intensity $q_{int.}$ of quenching. Water is used for quenching. Quenching is maintained until the maximum temperature within austenitized areas is below 50°C. To be able to compare temperature and martensite distribution, a depth d_M is defined. Within this depth the martensite fraction amounts to 95 % and is therefore considered to be fully martensitic.

Table 1: Investigation with f = 62 kHz and h_{WP} = 100 mm

Material	T_H [°C]	d_{TH} [mm]	d_M [mm]	t_{heat} [s]	t_{delay} [s]	Quenching agent	$q_{int.}$ [m³/s*m²]
C45	925	1,00	1,00	0,27	0	water	2,00
C45	925	2,00	2,00	1,75	0	water	2,00
C45	925	3,00	3,00	6,85	0	water	2,00
42CrMo4	925	3,00	3,00	6,30	0	water	2,00
42CrMo4	925	3,00	3,00	6,30	1,0	water	2,00
42CrMo4	925	3,00	3,00	6,30	0	water	0,28
42CrMo4	925	3,00	3,00	6,30	1,0	water	0,28

Table 1 shows simulation results for a frequency of 62 kHz and a work piece height of 100 mm. The large height of the work piece in relation to the hardening depth of a few millimeters means that the core temperature does not increase during heating. For a value of d_{TH} up to of 3 mm there is no deviation of d_M for C45 as well as 42CrMo4. Larger values of d_{TH} have not been investigated with this frequency because of the small penetration depth. A decrease of

the quenching rate from 2 m³/s*m² to 0.28 m³/s*m² and the influence of a delay of 1 s was checked for 42CrMo4 additionally.

Table 2: Investigation with f = 62 kHz and h_{WP} = 6 mm

Material	T_H [°C]	d_{TH} [mm]	d_M [mm]	t_{heat} [s]	t_{delay} [s]	Quenching agent	$q_{int.}$ [m³/s*m²]
42CrMo4	925	3,00	3,00	2,90	0	water	2,00
42CrMo4	925	3,00	3,00	2,90	0	water	0,28
42CrMo4	925	3,00	3,40	2,90	0,5	water	0,28
42CrMo4	925	3,00	3,72	2,90	0	water	0,28
42CrMo4	925	3,00	4,68	2,90	1,5	water	0,28
42CrMo4	870	3,00	4,44	2,90	1,5	water	2,00

In Table 2 simulation results for a frequency of 62 kHz and a reduced work piece height of 6 mm are shown for 42CrMo4. The austenitization depth d_{TH} has a fixed value of 3 mm, which means that 50 % of the work piece is austenitized after heating. In this case, the heating time of 2.9 is significantly lower than 6.3 s in case of a work piece with a height of 100 mm.

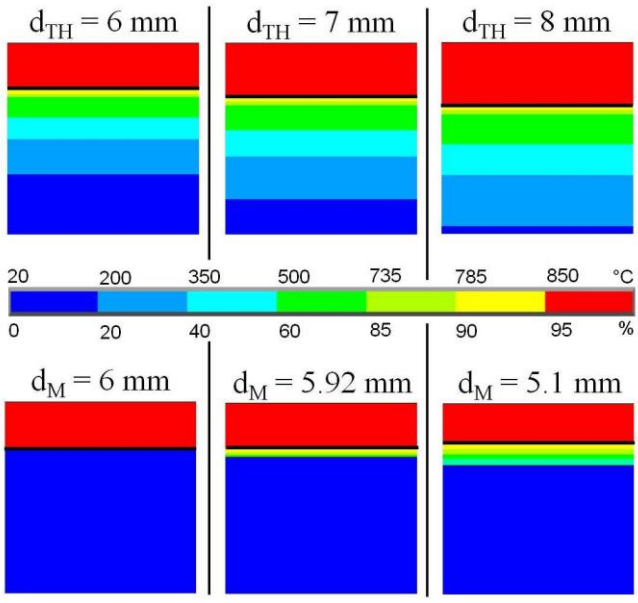


Figure 11: Investigation with f = 3 kHz, h_{WP} = 100 mm and d_{TH} with a value of (a) 6mm, (b) 7 mm and (c) 8 mm

A decrease of the quenching intensity does not affect d_M. A deviation only occurs, if a delay is taken into consideration as well. A delay time of 1.5 s results in 78 % percent of the work piece having a fully martensitic microstructure. If adjusted to the heating time of 2.9 s, T_H has a value of about 870°C. In this case, the deviation decreases and only 74 % of the work piece is fully martensitic. However, the influence of a delay is also evident here. In contrast to a work piece with a height of 100 mm, self-quenching does not occur. The average temperature after heating is much higher. Areas of the work piece further beneath the surface than d_{TH} are austenitized between heating and quenching because the heat cannot be transferred to a cold core.

Table 3: Investigation with f = 3 kHz, h_{WP} = 100 mm and variation of d_{TH}

Material	T_H [°C]	d_{TH} [mm]	d_M [mm]	t_{Heat} [s]	t_{delay} [s]	Quenching agent	$q_{int.}$ [m³/s*m²]
C45	850	6,00	6,00	7,70	0	water	1,20
C45	850	7,00	5,92	11,00	0	water	1,20
C45	850	8,00	5,10	15,00	0	water	1,20

The investigation of hardening with a frequency of 3 kHz is shown in Table 3. The austenitization temperature was defined as 850°C and a quenching intensity of 1.2 m³/s*m² was used. Up to an austenitization depth of 6 mm there is no difference between the temperature and martensite profile. Bigger austenitization depths result in a decrease of the fully martensitic area. This can be also seen in Figure 11, where temperature and martensite profiles are shown for different values d_{TH} from the left to the right. The deviation can be explained by taking a look at the heating times required for achieving these depths. For example, an austenitization depth of 8 mm requires a heating time of 15 s. Temperature curves for this case are illustrated in Figure 10. The amount of heat energy induced into the work piece in that period of time can neither be dissipated fast enough by quenching of the surface with water nor by self-quenching. Therefore, cooling rates are lower than v_U and only fractions of material become martensitic (compare Curve 2 in Figure 10). It can be stated that longer heating times and bigger depths d_{TH} have a negative impact on the depth of full martensite.

3.3 Verification

Figure 12 shows the microsection and simulation results of a hardened straight-cut gear. As described in [5], the gear was heated using simultaneous dual frequency (SDF®) and quenched with a water-air spray cooling afterwards. The temperature profile at the end of the heating process and the microsection show good agreement at the root and the flank of the tooth. However, the microsection shows that the tip of the tooth is almost through hardened, whereas the temperature distribution indicates a hardening profile following the gear's contour. This can be explained by taking a closer look at the assumptions that are made when a temperature distribution is used to predict a hardening profile: As mentioned before,

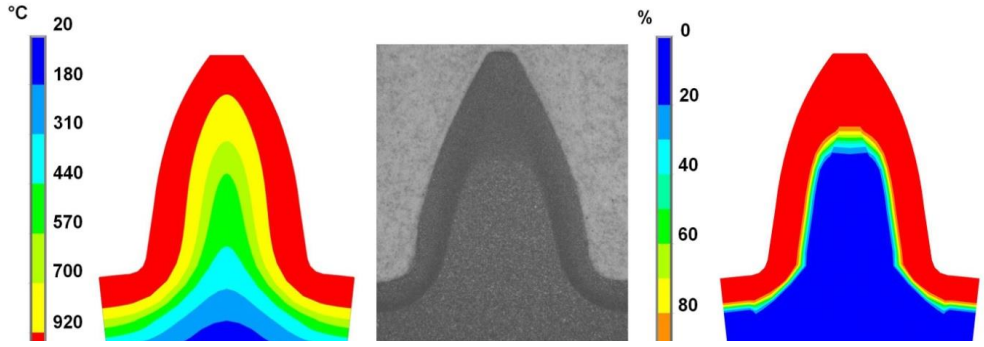


Figure 12: (a) Temperature profile, (b) microsection and (c) martensite profile of a straight cut gear

quenching is supposed to start without any delay after heating. If there is a delay, heat transfer from the surface of the work piece to its core has to be taken into account. The zone of austenitization might increase and therefore the hardened layer might as well. This effect is more pronounced at the tip of a tooth, because heat accumulates in this area, whereas self-quenching occurs in the root area and heat is transferred to the core of the work piece. The actual delay between the end of heating and the start of quenching was 0.4 s in the experiment. If this is considered in the simulation and the new algorithm to compute the martensite microstructure is used, a better agreement of microsection and simulation is evident for the tip: An increased hardening depth as well as the contour being less sharp can be seen if Figure 12 (c) and (a) are compared. The results for root and flank stay unchanged.

3.4 Conclusions

A module for the calculation of quenching and martensite fraction was implemented in the existing simulation tool used for computing inductive surface hardening processes. Investigations with a 1D model have shown that the temperature distribution at the end of the heating process usually is sufficient for a good prediction of the hardened profile if a frequency of several 10's of kHz is used and the height of the work piece is much bigger than the hardening depth. In case of a relatively thin work piece, a delay between heating and quenching as well as the quenching parameters have an influence on the hardening depth and martensite formation has to be calculated to receive accurate results. The verification of the extended simulation tool made evident that a deviation of temperature and martensite profile might also occur when more complex work pieces like gears are hardened with induction. A calculation of martensite should always be performed in such investigations to check, if there is deviation compared to the temperature profile.

REFERENCES

[1] Stiele, H., Brand, M., Asperheim J.-I.: Unterstützung der Prozessentwicklung des induktiven Randschichthärtens großer Lagerringe durch Simulation. Elektrowärme International, Vol. 68, 2010, No. 3, pp. 203-206.
[2] Stiele, H., Schulte, H.: Induction hardening of wind energy plants and components. Heat Processing, Vol. 7, 2009, No. 1, pp. 1-6.
[3] Zedler, T.: Numerische Modellierung, Analyse und Design von induktiven Systemen für das Randschichthärten komplexer Werkstückgeometrien. Sierke-Verlag, Göttingen, 2010, 50 pp.
[4] Koistinen, D. P., Marburger, R. E.: A General Equation Prescribing the Extent of the Austenite-Martensite Transformation in Pure Iron-Carbon Alloys and Plane Carbon Steels. Acta Metallurgica, 7 (1959) p. 59-60.
[5] Schlesselmann, D., Jestremski, M., Rodman, D., Nacke, B.: Numerical Simulation Methods for Complex Induction Hardening Processes. Proceedings of UIE Congress, XVII, St. Petersburg, Russia, (2012) p. 98-103.
[6] Golovin, G., Zamyatin, M.: High frequency heat treatment. Mashinostroenie, Leningrad (1990).
[7] Wildau, M., Hougardy, H.: Zur Auswirkung der Ms-Temperatur auf Spannungen und Maßänderungen. Härterei-Technische Mitteilungen, 42 (1987) p. 261-268.
[8] Börnsen M.: Zum Einfluss von Gefügeumwandlungen auf Spannungen und Formänderung bei thermischer und mechanischer Belastung. VDI-Verlag, Düsseldorf, 1 edition (1989).
[9] Johnson, W. A., Mehl, R. F.: Reaction Kinetics in Process of Nucleation and Growth. Transactions of the AIME, 135 (1939) p. 416-458.
[10] Avrami, M.: Kinetics of Phase Change III. Journal of Chemical Physics, 9 (1941) p. 177-184.
[11] Leblond, J. B., Devaux, J.: A New Kinetic Model for Anisothermal Metallurgical Transformations in Steels Including Effect of Austenite Grain Size. Acta Metallurgica, 32 (1984) p. 137-146.
[12] Behrens, B. A., Bach, F. W., Bouguecha, A., Nürnberger, F., Schaper, M., Yu, Z., Klassen, A.; Numerische Berechnung einer integrierten Wärmebehandlung für präzisionsgeschmiedete Bauteile. Journal of Heat Treatment and Materials, 67 (2012) p. 337-343.
[13] Läpple, V.: Wärmebehandlung des Stahls. Verlag Europa-Lehrmittel, Haan-Gruiten, 10 edition (2010).
[14] Wever, F., Rose, A., Peter, W., Strassburg, W., Rademacher, L.: Atlas zur Wärmebehandlung der Stähle. Verlag Stahleisen, Düsseldorf (1961).

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis (Eds)

MULTI-PHYSICS AND MULTI-OBJECTIVE DESIGN OF A BENCHMARK DEVICE: A PROBLEM OF INVERSE INDUCTION HEATING

P. DI BARBA*, F. DUGHIERO†, M. FORZAN† AND E. SIENI†

* Department of Electrical, Computer and Biomedical Engineering
University of Pavia
27100 Pavia, Italy
e-mail: paolo.dibarba@unipv.it, web page: http://www. unipv.it

†Department of Industrial Engineering
University of Padova,
Via Gradenigo 6/a
35131 Padova, Italy
email: {fabrizio.dughiero,michele.forzan,elisabetta.sieni}@unipd.it - Web page: http://www.unipd.it

Keywords: induction heating, optimization, magnetic and thermal problem, multi-physics.

Abstract. In the paper, a bi-objective optimization problem characterized by a multi-physics field analysis is investigated. The optimal design of a pancake inductor, related to the design of industrial devices for the controlled heating of a graphite disk is considered as the benchmark problem. The expected goal of the optimization process is to improve temperature uniformity in the disk as well as electrical efficiency of the inductor. The optimized device is designed using a multi-physics problem: magnetic problem for electrical efficiency computation and thermal problem for temperature uniformity computation. The solution of the relevant bi-objective optimization problem is based on a modified multi-objective genetic algorithm in the class of Non-dominated Sorting Genetic Algorithm. The proposed algorithm exploits the migration concept to vary the population genetic characteristics during optimization process in order to improve the Pareto front approximation.

1. INTRODUCTION

Induction heating is used in thermal processes to heat the workpiece at a prescribed temperature with high efficiency and accurate temperature control. In this area the solution of coupled electromagnetic and thermal fields is mandatory, as well as the use of optimization algorithms to identify the best device [1–6].

In the paper, a benchmark model to approach a multiphysics bi-objective optimal design is presented [2–5,7–9]. A finite-element analysis (FEA) is used to solve the inverse problem, whereas the optimization is performed by means of a modified NSGA algorithm. The proposed optimization algorithm uses the migration concept [10–12] to vary the population genetic heritage in order to modify the Pareto front through a better approximation in the objective space. It is well known that the concept of migration is observable in nature when groups of people move to a new country and mix with local population. For instance, mathematical models of bio-geography inspired a class of derivative-free optimization

algorithms aiming to find an optimum balance between immigrating and emigrating populations in an island [10,13–17]. The concept of migration has been already implemented in other algorithms that make use of parallel computing. In these algorithms migration concept is referred to an exchange of individuals between independent islands that evolve autonomously [14,18–20]. In the proposed strategy the migration concept is used to modify the genetic characteristics of the current population.

The benchmark model represents an industrial device for the epitaxial processing of silicon wafer [2,7,9]. This device, meaningful from the industrial process viewpoint, has been already proposed as a test model for new optimization algorithms [2,3,5,7,9].

The designed device includes a graphite disk, a pancake inductor and a ferrite yoke. Since a pancake inductor does not induce power on the axis of the load, temperature close to the disk axis can be significantly lower. The magnetic concentrator is used to increase the induced power density and as consequence the temperature in the center of the disk.

The direct problem solves a time-harmonic magnetic problem to evaluate the power density in the graphite disk coupled to a steady-state thermal problem to evaluate the temperature profile.

2. PROBLEM DESCRIPTION

2.1. Direct problem

The 2D benchmark model, sketched in Figure 1 (a), includes a graphite disk with a radius of 357.5 mm, an inductor with 12 copper turns (a pancake inductor) and a ferrite ring, magnetic field concentrator, under the most internal turns (one or two) that are located at the same height. Moreover, the two most external turns are fixed at the same height. In Figure 1 (a) the design variables are also shown.

All turns, series connected, carry a current in the order of 500-600 A_{rms} at 4,250 Hz [7,21,22]. A total power of about 60 kW is prescribed in the device so that the disk reaches a steady state average temperature of 1050-1100 °C, as required by the industrial process. The corresponding inductor current is tuned in each FEA simulation: the FEA solution is updated with a new value of the inductor current (the source of electromagnetic model) when nonlinear material properties are taken into account in the model. Thermal and electrical properties of materials are in Table 1. In Figure 2, magnetic relative permeability and magnetization curve of the magnetic concentrator are presented.

The target of the multi-physics design is to evaluating the graphite disk temperature using a given inductor geometry. The magnetic analysis evaluates the power density in the graphite disk starting from the inductor geometry and the supply current. The power density is the source for the thermal problem.

The magnetic problem is solved in time-harmonic conditions by means of a commercial FEA code using the well-known **A**-V formulation, on second-order elements [23,24]. The current distribution in each turn is taken into account to correctly evaluate the inductor efficiency [2,25,26]. In particular, the magnetic problem is solved in terms of the phasor of the magnetic vector potential, $\boldsymbol{A}$ [2,3,25]:

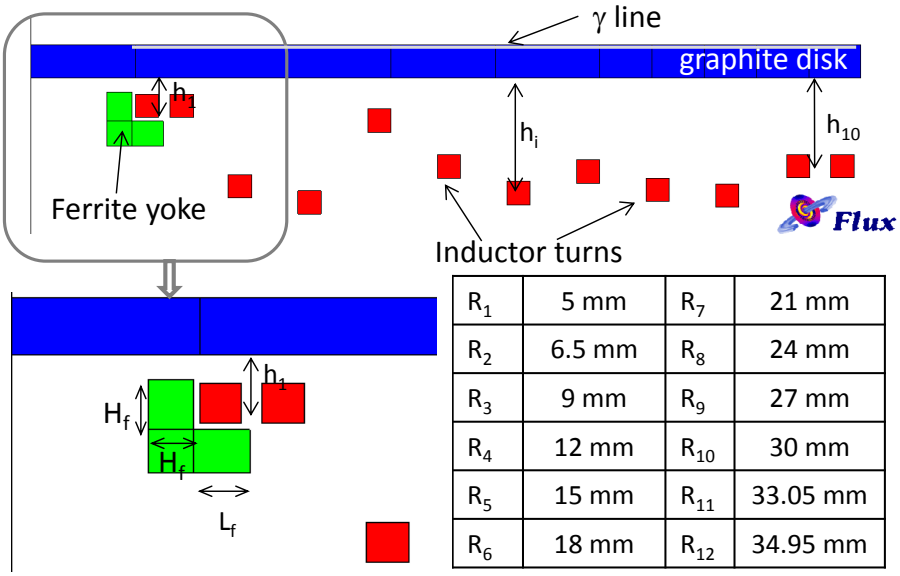
$$\nabla^2 \dot{\boldsymbol{A}} - j\omega\mu\rho^{-1}\dot{\boldsymbol{A}} = -\mu\dot{\boldsymbol{J}} \tag{1}$$

where $\dot{\boldsymbol{J}}$ and $\dot{\boldsymbol{A}}$ are the phasors of the current density and magnetic vector potential, respectively, μ is the material magnetic permeability, ρ the material electrical resistivity and ω magnetic field pulsation. The resistivity depends on materials as reported in Table I and is the one of the graphite ($\rho=\rho_g$) in the disk, the one of the copper ($\rho=\rho_c$) in the inductor turns. Conductivity of the air and magnetic concentrator is null.

The electrical efficiency, η, is, then, computed from the power density as follow:

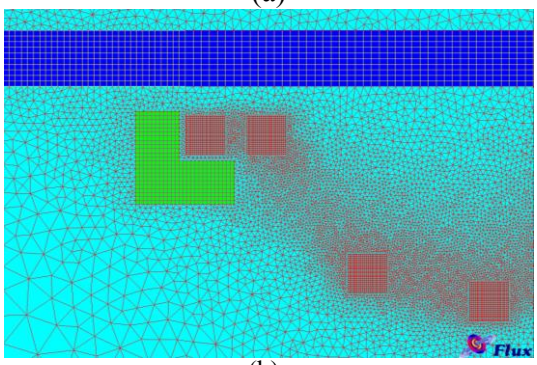
$$\eta = \frac{\int_{V_g} \rho_g^{-1} \omega^2 \left\| \dot{\boldsymbol{A}} \right\|^2 dV}{\int_{V_g} \rho_g^{-1} \omega^2 \left\| \dot{\boldsymbol{A}} \right\|^2 dV + \int_{V_c} \rho_c^{-1} \omega^2 \left\| \dot{\boldsymbol{A}} \right\|^2 dV} \tag{2}$$

where V_g and V_c are the volume of the graphite disk and the copper turns, respectively.



R_1	5 mm	R_7	21 mm
R_2	6.5 mm	R_8	24 mm
R_3	9 mm	R_9	27 mm
R_4	12 mm	R_{10}	30 mm
R_5	15 mm	R_{11}	33.05 mm
R_6	18 mm	R_{12}	34.95 mm

(a)



(b)

Figure 1: (a) Geometry of the benchmark model with 12 design variables. In the enlarged image the magnetic concentrator (ferrite ring), and the relevant design variables are also shown. Finally, in the table the radii of the turns are reported. (b) Detail of the mesh.

The thermal problem is solved in steady-state condition, assuming the power density in the disk computed by means of the magnetic problem as the source term [2,3,27]. The thermal

domain is the graphite disk. Along the domain profile a boundary condition of heat exchange along is imposed. The thermal conductivity of the graphite is reported in Table I [2,28,29]. From these assumptions, it results that the solutions of magnetic and thermal problem are weakly coupled by means of the source term of thermal equation:

$$-\nabla\cdot(\lambda\nabla T)=\rho^{-1}\omega^2\left\|\dot{\boldsymbol{A}}\right\|^2 \tag{3}$$

in which λ is the thermal conductivity of the material. Along the disk surface these boundary conditions subsist:

$$\frac{\partial T}{\partial n}=0 \tag{4}$$

at r = 0, and

$$-\lambda\frac{\partial T}{\partial n}=h\left(T-T_0\right)+\varepsilon k_B\left(T^4-T_0^4\right) \tag{5}$$

elsewhere, where h is the convective exchange coefficient (h=10 $Wm^{-2}K^{-1}$), ε emissivity coefficient (ε=0.6) and k_B Stefan-Boltzmann constant. The external temperature, T_0, is equal to 850 °C. These parameters of the thermal model have been tuned in order to fulfill experimental results of a real device considering that the expected average temperature is 1050-1100 °C.

Table 1: Electrical and thermal material properties of model materials

Element	**Electrical properties**		
Disk	Graphite (at 1200°C)	ρ_g=7.76 10^{-6} Ωm	μ_r = 1
Inductor	Copper	ρ_c=1.6 10^{-8} Ωm	μ_r = 1
Ferrite ring	Ferrite	--	Nonlinear. Relative permeability in Figure 2
	Thermal properties		
Disk	Graphite	λ=60 $Wm^{-1}K^{-1}$	

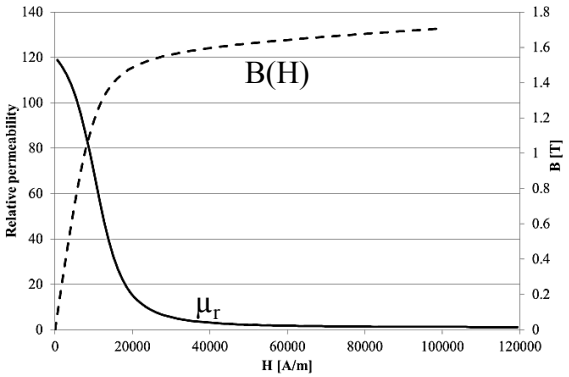


Figure 2: Relative permeability and magnetization curve of ferrite ring material.

A typical second-order mesh to solve the magnetic problem has 124,000 nodes and 57,000 elements. An example of the domain mesh is in Figure 1(b). The mesh of the thermal model is the one of the ferrite disk. The temperature has been evaluated on 201 sample points,

regularly spaced along the γ line showed in Figure 1(a).

2.2. Inverse problem

A 12-dimensional vector $\boldsymbol{x}$ of geometric design variables has been defined for the model in Figure 1 (a); the list of considered design variables with corresponding variation ranges is in Table 2. In particular, the design variables are the vertical positions of the inductor turns and the size of the magnetic yoke. The design problem is characterized by two conflicting objectives: the maximization of the electrical efficiency, η, defined as the ratio of active power transferred to the disk to the one transferred to the entire device using (2), and maximization of the temperature uniformity along the surface of the graphite disk at thermal steady state. In practice the inverse problem has been implemented as the simultaneously minimization of the following two objective functions:

$$f_1(\boldsymbol{x}) = 1 - \eta(\boldsymbol{x}) \tag{6}$$

$$f_2(\boldsymbol{x}) = \left[N_{\max} - \sup_j N_j(|T_j(\boldsymbol{x},\gamma) - T_i(\boldsymbol{x},\gamma)| < \frac{\Delta T^*}{2}) \right], \quad j \neq i \tag{7}$$

where f_1 is the complementary value of the electrical efficiency and f_2 measure the temperature in-homogeneity using the "criterion of proximity" [2,4,30]. This criterion is based on a tolerance interval, ΔT (=10°C), around a give temperature value. In practice for each solution the greatest number of points included in the tolerance band is searched for (that corresponds to search for the minimum number of points outside the tolerance band). A detailed description of this criterion is in [2,4,30]. An example of 'proximity criterion' is in Figure 3: considering the reference path γ, shown in Figure 1(a), 9 points satisfy the criterion while 14 points are outside the tolerance band ΔT; therefore 14 is the value assigned to f_2.

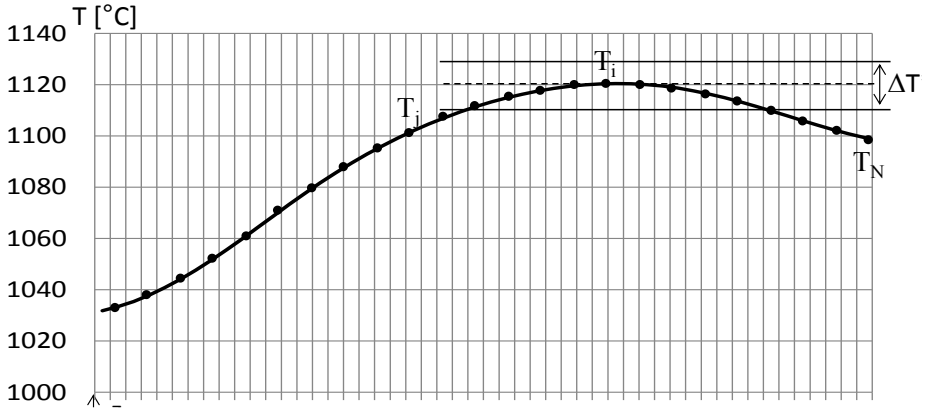


Figure 3: Example of "criterion of proximity" applied to the temperature along the γ path shown in Fig 1(a).

Table 2: Design variable ranges

	min	**max**
h_1,....., h_{10} [mm]	0	60
L_f [mm]	1	40
H_f [mm]	10	25

In the optimization problem, both functions (6) and (7) have to be minimized with respect to design variables shown in Figure 1 (a) and Table 2: the objective (6) refers to the magnetic domain, while the (7) refers to the thermal domain and a multi-physics and multi-objective inverse problem is originated.

1.1. Multi-physics optimization problem

The direct and inverse problems are coupled as in Figure 4 in order to create an optimization tool to design the improved geometry of the device in Figure 1(a). The multi-physics problem includes the solution in two steps of a magnetic problem and a thermal problem. The magnetic problem step includes two solutions of the same geometry if non-linear material properties are considered. The first solution is used to evaluate the amplitude of the current source in order that the power dissipated by the device is close to the prescribed one (e.g. 60 kW); whereas the power density that is the source term of the thermal problem, and the electrical efficiency that is the input of the optimization algorithm, are evaluated from the second solution. From the thermal problem the second input for the optimization algorithm, the temperature uniformity, is evaluated. Then, the solution of the multi-physics direct problem in terms of electric efficiency and temperature uniformity are the inputs of the optimization algorithm that generates a new set of design variables vectors.

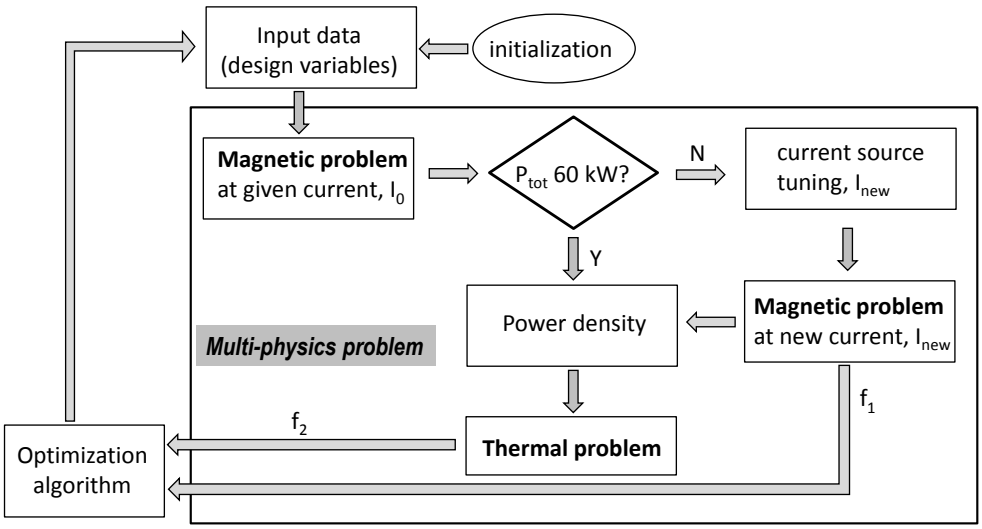


Figure 4 Coupling of the direct and inverse problem.

2. OPTIMIZATION ALGORITHM

Optimization process uses a modified NSGA-II algorithm, named Migration-NSGA, MNSGA, that implements the concept of the migration of groups of individuals that belong to populations with different genetic characteristics [8,31,32]. This algorithm mimics the event of the arrival of a group of individuals that mixes with the current population and has different characteristics with respect to originating population. The new individuals carry different genes that can improve the original population. This way, the genetic heritage of the population can be mutated.

MNSGA with the strategies sketched in Figure 5. A group of P_m individuals, generated

through a random process like the one applied to create the initial population, the 'immigrated population', is added to the current population after the main genetic operator (i.e. cross-over + mutation) and before the selection operator. This way, the immigrant individuals are added to the existing population of parents and off-springs and selection operates in such a way that, among the new individuals, only the ones with better characteristics are preserved in the population survived after selection.

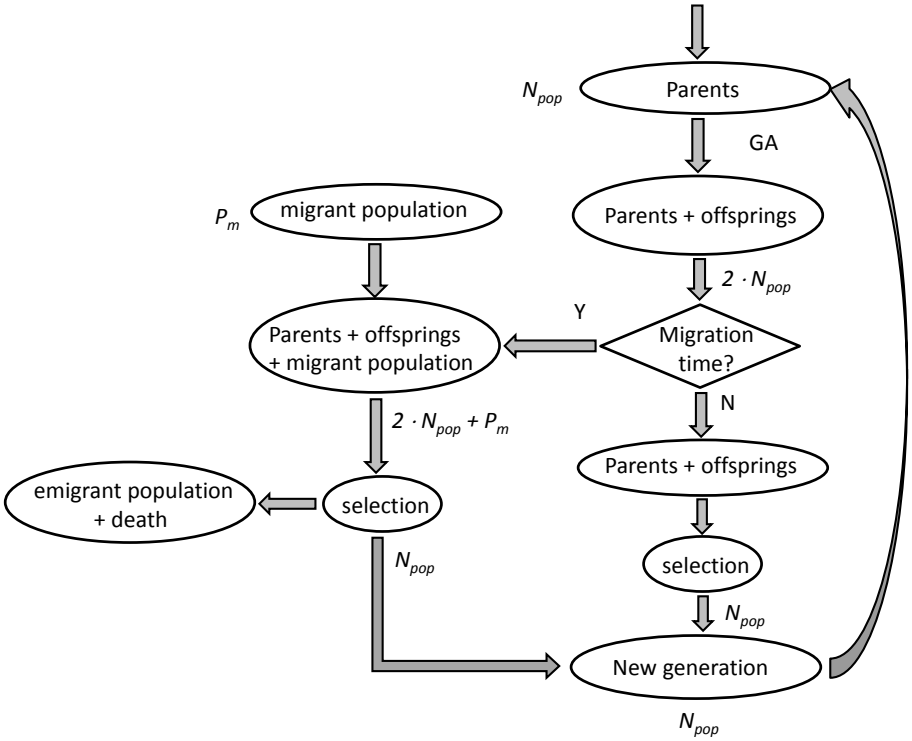


Figure 5 MNSGA algorithm

The concept of the population migration in a population with N_{pop} individuals is ruled by two parameters:

- the period of migration, T_m, with respect to the number of generations, corresponding to the number of iterations (e.g. T_m=1 is equivalent to introduce a migration at every iteration)
- the number of individuals in the immigrated population, P_m, between 1 and N_{pop}.

The number of individuals after selection is maintained constant to N_{pop} by applying natural selection and emigration events. Both the steps that reduce the population dimension managed in the selection algorithm intended as: a subgroup of individuals dies and a subgroup emigrates. In both cases some individuals come out from the population during the selection step of the NSGA algorithm.

3. RESULTS

The device in Figure 1(a) has been optimized using both standard NSGA-II and new MNSGA algorithm. In the MNSGA algorithm different values for the migration parameters, T_m and P_m, have been chosen (Table 3). The number of iterations of both algorithms, NSGA-

II and MNSGA, has been fixed to 50 with a population of 20 individuals. In Table 3 the maximum number of new individuals evaluated in each case during the optimization process is reported. It is evident that MNSGA algorithm evaluates a higher number of individuals and therefore implies an increased number of calls to the objective functions. Nevertheless, the increment of number of individuals in the order of 10-25 % improves the Pareto front approximation including more solutions that are not found using NSGA-II algorithm.

Table 3: Description of the optimization algorithm set-up.

#case	T_m	P_m	#NSGA-generated individuals	#immigrated individuals (increment[%])	#total individuals
NSGA	--	--	1020	--	1020
MNSGA_T2N10	2	10	1020	250 (+25%)	1270
MNSGA_T5N10	5	10	1020	100 (+10%)	1120
MNSGA_T5N20	5	20	1020	200 (+19%)	1220

In Figure 6 the comparison of the Pareto front for the cases NSGA and MNSGA_T5N20 (Table 3), obtained starting from the same initial population, is reported. The Pareto fronts found using MNSGA algorithm are broader with respect to the ones obtained using NSGA-II algorithm. Moreover, in some cases solutions found with NSGA-II are dominated by the ones obtained using MNSGA. It appears that the inclusion of new individuals (i.e. the immigration event) during the population evolution improves substantially the approximation of the Pareto front.

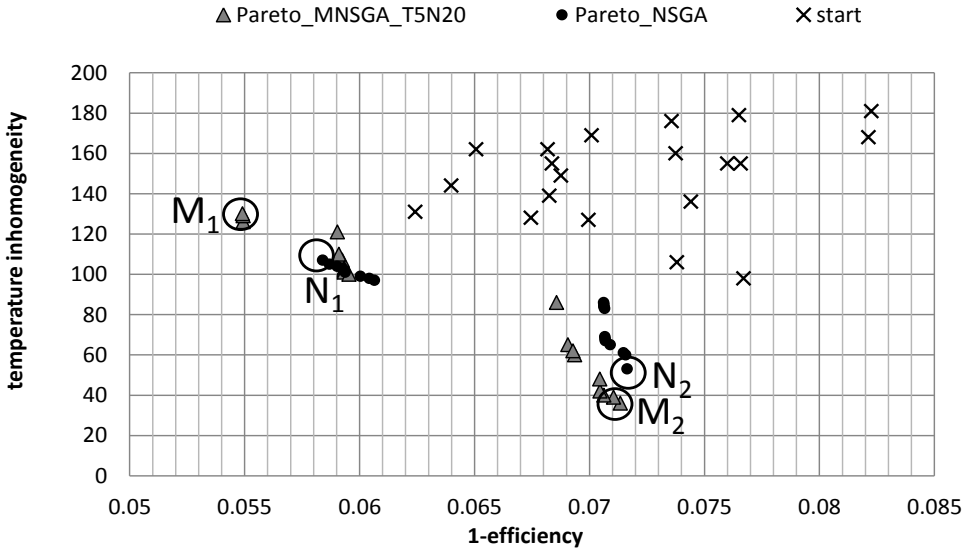


Figure 6 Starting populations and Pareto fronts obtained using NSGA-II and MNSGA algorithm.

Figures 7 and 8 show the geometries corresponding to the solutions outlined on the approximated Pareto front in Figure 6 and reported in Table 4. They have been obtained using NSGA-II algorithm and MNSGA algorithm, respectively. In each figure the solution of magnetic problem is shown in terms of the magnetic flux lines, whereas the thermal problem solution is shown in terms of the temperature along the disk profile. The difference in terms of temperature uniformity for the two solutions at the two ends of the Pareto front is evident.

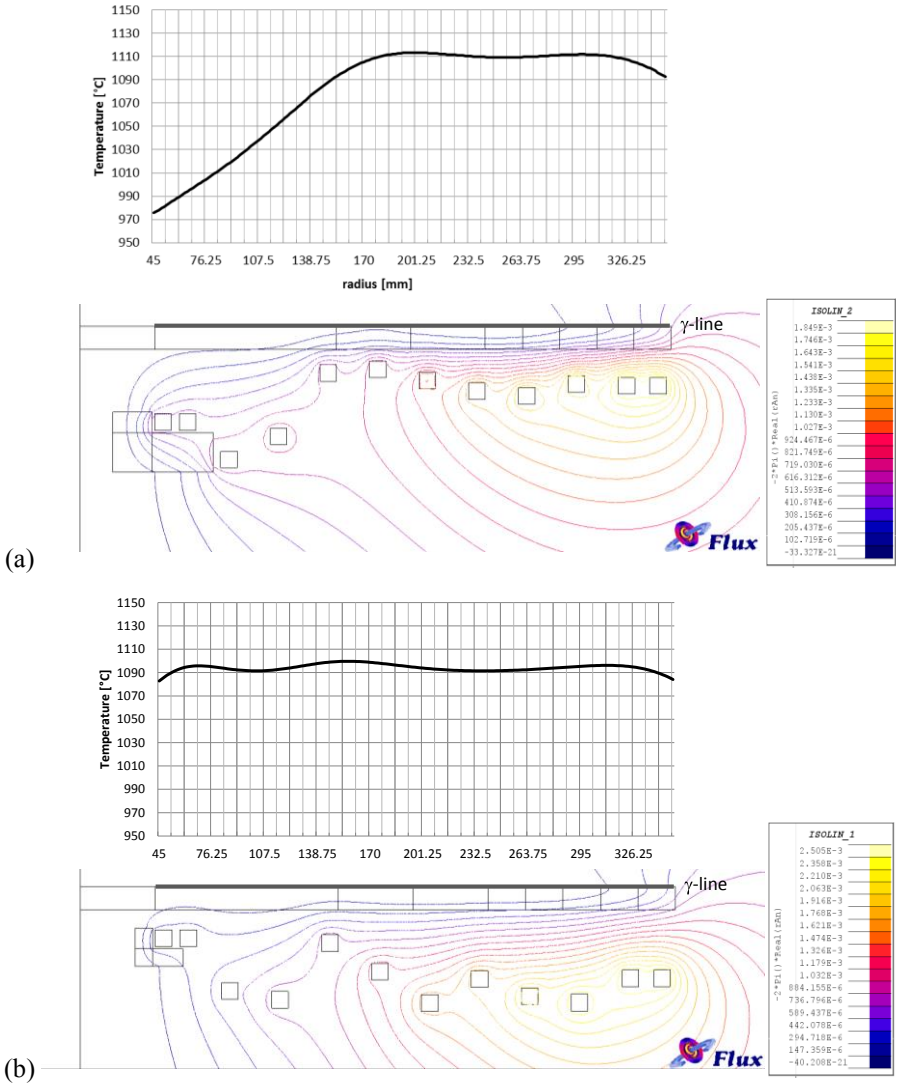


Figure 7 Geometry, magnetic flux lines and temperature on the disk surface of two solutions on Pareto front obtained using NSGA-II algorithm. (a) Point N1 and (b) point N2.

Table 4: turn positions [mm], ferrite sizes [mm] and values of objective functions (6) and (7) for solutions highlighted in Figure 6.

	h_1	h_2	h_3	h_4	h_5	h_6	h_7	h_8	h_9	h_{10}	h_f	L_f	f_1	f_2
N1	28.3	5.6	19.6	57.9	60.0	53.2	47.0	44.1	51.1	50.1	23.8	37.0	0.058	107
N2	55.6	23.9	18.6	52.6	35.5	16.8	31.0	20.9	17.0	31.6	10.5	18.3	0.072	53
M1	46.2	45.9	59.7	52.1	40.9	53.7	26.0	49.7	58.0	59.6	14.7	5.4	0.055	130
M2	55.6	24.1	21.3	49.9	32.4	22.4	32.1	22.0	14.8	32.2	11.1	21.7	0.071	36

Comparing Figure 7(b) and 8(b), that correspond to the points M2 and N2 in Figure 6 for which (7) is minimum, it is evident that the found solutions in terms of turns positions and ferrite sizes are very similar, as well as it appears in Table 4. In contrast, the geometry in Figure 7 (a) and 7(b) are very different in terms of both ferrite sizes and turn positions.

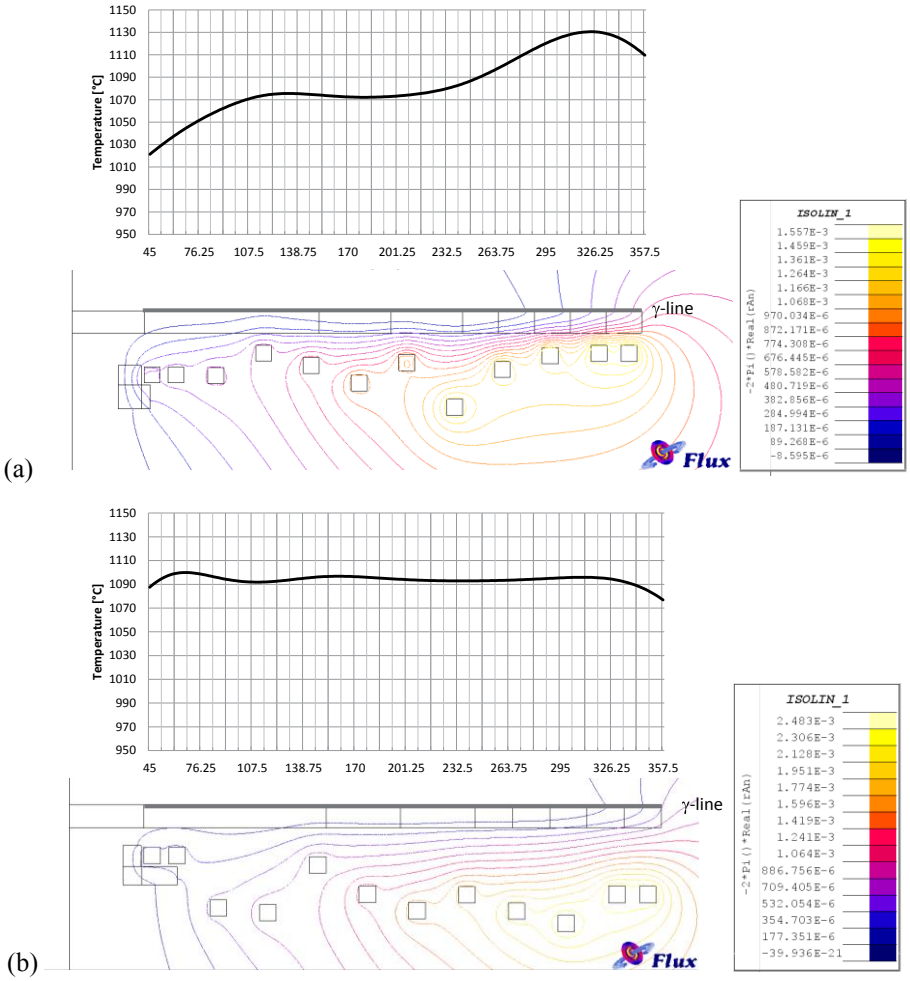


Figure 8 Geometry, magnetic flux lines and temperature on the disk surface of two solutions on Pareto front obtained using MNSGA algorithm. (a) Point M1 and (b) point M2.

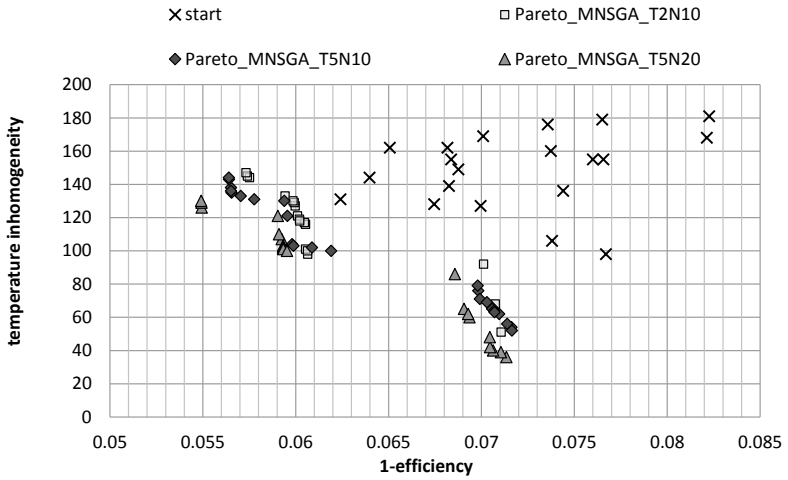


Figure 9 Comparison of the Pareto fronts obtained using MNSGA with different set-up of the immigration parameters (as in Table 3).

Figure 9 shows the Pareto front obtained starting from the same initial population and tuning the MNSGA algorithm as in Table 3: a different tuning of the optimization algorithm gives rise to different approximated Pareto fronts.

4. CONCLUSIONS

The paper shows the results obtained solving a multi-physics benchmark problem and optimizing two objectives functions stemming from two different physical domains. The approximated Pareto fronts prove a substantial improvement of the objective functions with respect to the initial sets. In particular, optimal solutions with very good temperature uniformity have been found.

REFERENCES

[1] Di Barba P, Dughiero F, Lupi S, Savini A. Optimal shape design of devices and systems for induction-heating: Methodologies and applications. *COMPEL - Int J Comput Math Electr Electron Eng* (2003) **22**:111–22.

[2] Di Barba P, Forzan M, Sieni E. Multi-objective design of a power inductor: a benchmark problem of inverse induction heating. *COMPEL - Int J Comput Math Electr Electron Eng* (2014) **33**:1990–2005.

[3] Di Barba P, Dughiero F, Forzan M, Sieni E. Parametric vs non-parametric optimal design of induction heating devices. *Int J Appl Electromagn Mech* (2014) **44**:193–9.

[4] Di Barba P, Dughiero F, Forzan M, Sieni E. A Paretian Approach to Optimal Design With Uncertainties: Application in Induction Heating. *Magn IEEE Trans On* (2014) **50**:917–20.

[5] Di Barba, P, Pleshivtseva Y, Rapoport E, Forzan M, Lupi S, Sieni E, Nacke B, Nikanorov A, Multi–objective optimisation of induction heating processes: methods of the problem solution and examples based on benchmark model. *Int J Microstruct Mater Prop* (2013) **8**:357–72.

[6] Pleshivtseva Y, Di Barba, P., Rapoport E, Nacke B, Nikanorov A, Lupi S, Sieni E, Forzan, M. Multi-objective optimisation of induction heaters design based on numerical coupled field analysis. *Int J Microstruct Mater Prop* (2014) **9**:532–551.

[7] Forzan, M., Maccalli, G., Valente, G., Crippa, D. Design of an innovative heating process system for the epitaxial growth of silicon carbide layers wafer, *Proc. of MMP-Modelling for Material Processing* (2006):101-107.

[8] Di Barba P. *Multiobjective shape design in electricity and magnetism*. Doredrecht ; New York: Springer; (2010).

[9] Di Barba, Paolo, Forzan M., Sieni E. Multiobjective design optimization of an induction heating device: a benchmark problem. *Int J Appl Electromagn Mech* (in press).

[10] Simon D. Biogeography-Based Optimization. *Evol Comput IEEE Trans On* (2008) **12**:702–13.

[11] Forrest S. Genetic Algorithms: Principles of Natural Selection Applied to Computation. *Science* (1993) **261**:872–8.

[12] Schmitt LM. Theory of genetic algorithms. *Theor Comput Sci* (2001) 259:1–61.

[13] Xu S, Ji Z, Pham DT, Yu F. Binary Bees Algorithm – bioinspiration from the foraging mechanism of honeybees to optimize a multiobjective multidimensional assignment problem. *Eng Optim* (2011) **43**:1141–59.
[14] Dos Santos Coelho L, Alotto P. Electromagnetic Optimization Using a Cultural Self-Organizing Migrating Algorithm Approach Based on Normative Knowledge. *Magn IEEE Trans On* (2009) **45**:1446–9.
[15] Hu X-M, Zhang J, Li Y. Orthogonal Methods Based Ant Colony Search for Solving Continuous Optimization Problems. *J Comput Sci Technol* (2008) **23**:2–18.
[16] Palupi Rini D, Mariyam Shamsuddin S, Sophiyati Yuhaniz S. Particle Swarm Optimization: Technique, System and Challenges. *Int J Comput Appl* (2011) **14**:19–27.
[17] Kennedy J, Eberhart R. Particle swarm optimization. *Neural Netw 1995 Proc IEEE Int Conf On* (1995) **4**:1942–8.
[18] Märtens M, Izzo D. The asynchronous island model and NSGA-II: study of a new migration operator and its performance. *Proc. 15th Annu. Conf. Genet. Evol. Comput., Amsterdam, The Netherlands: ACM*; (2013):1173–80.
[19] Deep K, Dipti. A self-organizing migrating genetic algorithm for constrained optimization. *Appl Math Comput* (2008) **198**:237–50.
[20] Zelinka I. Real-time deterministic chaos control by means of selected evolutionary techniques. *Eng Appl Artif Intell* (2009) **22**:283–97.
[21] Zgraja J. The optimisation of induction heating system based on multiquadric function approximation. *COMPEL - Int J Comput Math Electr Electron Eng* (2005) **24**:305–13.
[22] Forzan M, Lupi S, Toffano E. Compensation of induction heating load edge-effect by space control. *COMPEL - Int J Comput Math Electr Electron Eng* (2011) **30**:1558–69.
[23] FLUX. (CEDRAT): www.cedrat.com/software/flux/flux.html [last visit January 2015].
[24] Di Barba P, Savini A, Wiak S. *Field models in electricity and magnetism*. [Dordrecht]: Springer; (2008).
[25] Aliferov A, Dughiero F, Forzan M. Coupled Magneto-Thermal FEM Model of Direct Heating of Ferromagnetic Bended Tubes. *Magn IEEE Trans On* (2010) **46**:3217–20.
[26] Karban P, Kotlan V, Dolezel I. Numerical Model of Induction Shrink Fits in Monolithic Formulation. *Magn IEEE Trans On* (2012) **48**:315–8..
[27] Carslaw H. *Conduction of heat in solids*. 2nd ed. Oxford [Oxfordshire] New York: Clarendon Press; Oxford University Press; (1986).
[28] Di Barba P, Dughiero F, Sieni E, Candeo A. Coupled Field Synthesis in Magnetic Fluid Hyperthermia. *Magn IEEE Trans On* (2011) **47**:914–7.
[29] Dughiero F, Forzan M, Pozza C, Sieni E. A Translational Coupled Electromagnetic and Thermal Innovative Model for Induction Welding of Tubes. *IEEE Trans Magn* (2012) **48**:483–6.
[30] P. Di Barba,F. Dughiero, M. Forzan, E. Sieni. Sensitivity-based optimal shape design of induction-heating devices. *IET Sci Meas Technol* (in press).
[31] Deb K, Pratap A, Agarwal S, Meyarivan T. A fast and elitist multiobjective genetic algorithm: NSGA-II. *Evol Comput IEEE Trans On* (2002) **6**:182–97.
[32] Deb K. *Multi-objective optimization using evolutionary algorithms*. 1st ed. Chichester ; New York: John Wiley & Sons; (2001).

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis (Eds)

MULTIPHYSICS FINITE – ELEMENT MODELLING OF AN ALL – VANADIUM REDOX FLOW BATTERY FOR STATIONARY ENERGY STORAGE

F. MORO, A. BERTUCCO, V. FIORENZATO, M. GIOMO, AND M. GUARNIERI

Dipartimento di Ingegneria Industriale (DII)
Università degli Studi di Padova
Via Gradenigo 6/A, 35131 Padova, Italy
e-mail: federico.moro@unipd.it

Key words: Coupled Problems, Multiphysics Problems, Redox Flow Battery, Vanadium, Energy Storage.

This work was supported by the University of Padova under the MAESTRA 2011 strategic project.

Abstract. All-Vanadium Redox Flow Batteries (VRFBs) are emerging as a novel technology for stationary energy storage. Numerical models are useful for exploring the potential performance of such devices, optimizing the structure and operating condition of cell stacks, and studying its interfacing to the electrical grid. A one-dimensional steady-state multiphysics model of a single VRFB, including mass, charge and momentum transport and conservation, and coupled to a kinetic model for electrochemical reactions, is first presented. This model is then extended, including reservoir equations, in order to simulate the VRFB charge and discharge dynamics. These multiphysics models are discretized by the finite element method in a commercial software package (COMSOL). Numerical results of both static and dynamic 1D models are compared to those from 2D models, with the same parameters, showing good agreement. This motivates the use of reduced models for a more efficient system simulation.

1 INTRODUCTION

The growing penetration of renewable sources in the electric grid has boosted the development of new technologies for stationary energy storage. All-Vanadium Redox Flow Batteries exhibit a very high potential for both medium and large scale applications. This is due to power/energy independent sizing, high round-trip efficiency, room temperature operation, and long charge/discharge cycle life. A number of challenges need to be tackled in order to gain full commercial success, regarding single cell and stack design, highly efficient energy-conversion materials, and optimal power management and control operations [1].

Numerical models are powerful tools in designing control and monitoring systems, which are needed for interfacing the stack to the electric grid. Both 1D and 2D finite element models have been developed in order to reduce computing cost in comparison with 3D models and allow for a real-time simulation of VRFB operations. A considerably complete and

comprehensive 2D distributed model, able of capturing time variations of current density and species concentrations inside the cell, was first developed by Shah et al. [2]. Vanadium crossover and related losses were accounted for in [3]. A 2D steady-state model, showing good agreement with experiments despite its simplicity, was later proposed by You et al. [4]. To simulate a VRFB stack rather than a single cell, Vynnycky proposed an asymptotic reduction approach to obtain a 1D multiphysics model for a single electrochemical cell. An enhanced version of this model, with concentration dependent kinetics, was presented in [5].

Starting from this work, a novel 1D steady-state finite-element model of an all-vanadium redox flow battery is here proposed. A typical VRFB cell (Fig.1) consists of: the positive and the negative porous layers where chemical half-reactions occur, the separation polymer membrane for proton conduction, and the current collectors which convey the electrons to/from the external circuit. The FEM model incorporates the following multiphysics equations: momentum, charge and species conservation, mass transport (Nernst-Plank equation), electric conduction, charge generation (Butler-Volmer equation) inside the porous electrodes, and the proton conduction inside the membrane. In particular, the effect of mass transport on the charge generation is taken into account by considering the electrode surface concentrations, which differ from bulk concentrations. Differently from [5], the 1D model is implemented within a commercial software package (COMSOL) and takes into account the behaviour of the current collectors in order to achieve a more accurate representation of the device. The independent variables of the model are the electrode and electrolyte potentials and the ionic species concentrations in the electrodes. The cell voltage is obtained from the electrode potential as a function of the State of Charge (SoC).

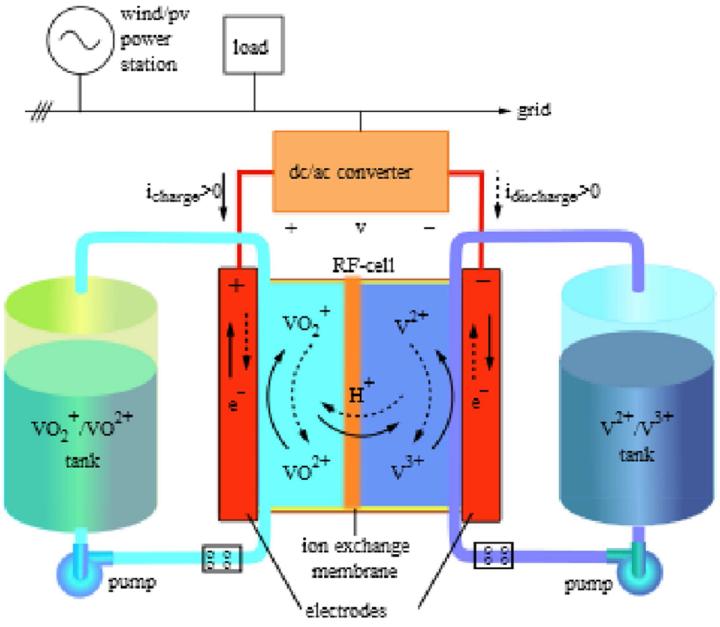


Figure 1: Schematic diagram of a VRFB (collectors + porous electrodes + polymeric membrane) interfaced to the electrical grid by a DC/AC converter. Electrolytes, stored in separated tanks, are pumped to the VRFB.

2 MULTIPHYSICS MODEL

Fig. 1 shows a schematic diagram of a VRFB cell interfaced to the electrical grid. Energy conversion occurring inside VRFBs is based on the following electrochemical reduction and oxidation reactions at the electrode-electrolyte interface:

$$VO^{2+} + H_2O \rightleftarrows VO_2^+ + 2H^+ + e^- \quad \text{(positive half-cell)} \tag{1}$$

$$V^{3+} + e^- \rightleftarrows V^{2+} \quad \text{(negative half-cell)}$$

The electrolytes are vanadium oxides dissolved in a sulphuric acid water solution (1–5 M concentration). Arrows in (1) indicate the charge/discharge operations, e.g. the rightward arrows indicate the charging phase when VO_2^+ and V^{2+} are generated at the positive and negative electrode, respectively. During the charging operation the reduction half-reaction at the negative electrode extracts electrons from the collector and ions from the electrolyte, while the oxidation half-reaction at the other electrode provides electrons to the collector. Protons H^+ migrate from one electrode to the other (from anode to cathode) through a polymer electrolyte (membrane) that is impermeable to electrons, which are thus forced through the external circuit producing an electrical energy exchange.

2.1 Porous electrode equations

The electric charge is generated or consumed at the electrolyte-solid phase interfaces by chemical half-reactions. The reaction kinetics are modeled by using general expressions, derived from the classical Bulter-Volmer equation, which does not take into account the difference between surface and bulk concentrations. Following [5], charge generation rates at positive (1) and negative (2) electrodes are:

$$\begin{aligned} J_{g1} &= J_1^0 \left[\frac{c_{V4,s}}{c_{V4}} \exp\left(\frac{\alpha_{1a} F}{R\,T} \eta_1 \right) - \frac{c_{V5,s}}{c_{V5}} \exp\left(-\frac{\alpha_{1c} F}{R\,T} \eta_1 \right) \right] \\ J_{g2} &= J_2^0 \left[\frac{c_{V2,s}}{c_{V2}} \exp\left(\frac{\alpha_{2a} F}{R\,T} \eta_2 \right) - \frac{c_{V3,s}}{c_{V3}} \exp\left(-\frac{\alpha_{2c} F}{R\,T} \eta_2 \right) \right] \end{aligned} \tag{1}$$

where J_1^0, J_2^0 are (concentration dependent) exchange current densities, $c_{i,s}$ are electrolyte surface concentrations at the interface (which generally differ from bulk ones, c_i), $\alpha_{1,a}, \alpha_{2,a}$ and $\alpha_{1,c}, \alpha_{2,c}$ are the anodic and cathodic charge transfer coefficients, and R, T, F are constants (ideal gas and Faraday constants and temperature). Overpotentials in (1) are:

$$\begin{aligned} \eta_1 &= \phi_s - \phi_e - E_1 \\ \eta_2 &= \phi_s - \phi_e - E_2 \end{aligned} \tag{2}$$

where ϕ_s, ϕ_e are the electric potential in the solid phase and the electrolyte. The open circuit voltage (OCV) of each half-cell is calculated according to the Nernst equation as follows:

$$\begin{aligned} E_1 &= E_{0,1} + \frac{R\,T}{F} \log\left(\frac{a_{V5} a_{H^+}^2}{a_{V4}} \right) \\ E_2 &= E_{0,2} + \frac{R\,T}{F} \log\left(\frac{a_{V3}}{a_{V2}} \right) \end{aligned} \tag{3}$$

where $E_{0,1}, E_{0,2}$ are the equilibrium cell potentials in standard conditions.

The electrochemical energy conversion occurs inside porous electrodes, where reactants and products are delivered. Ion species migration in porous electrodes is governed with a mass concentration equation, which is steady-state conditions becomes:

$$\nabla \cdot \boldsymbol{N}_i = S_i \tag{4}$$

where $\boldsymbol{N}_i$ is the molar flux of the i^{th} ion species in the porous medium and S_i is the source term of electrochemical half-reactions. For the positive electrode $i \in \{V_4, V_5, H_1^+, SO_4^-\}$ and for the negative electrode $i \in \{V_2, V_3, H_2^+, SO_4^-\}$. Source terms for these species are: $S_{V4} = -A\,J_{g1}/F$, $S_{V5} = A\,J_{g1}/F$, $S_{H_1^+} = 2A\,J_{g1}/F$, $S_{V2} = -A\,J_{g2}/F$, $S_{V3} = A\,J_{g2}/F$, $S_{H_2^+} = 0$.

The flow density is then composed of three source terms: molecular diffusion, migration due to electric field, and advection. By assuming dilute electrolyte concentrations, this can be modeled by using the Nernst-Plank equation:

$$\boldsymbol{N}_i = -D_i^{eff} \nabla\, c_i - \frac{z_i c_i}{RT} D_i^{eff} F\, \nabla \phi_e + \boldsymbol{v}\, c_i \tag{5}$$

where z_i, c_i are the valence and molar concentration of the i^{th} species, and $\boldsymbol{v}$ is the velocity.

The effective diffusivity is given by Bruggeman correction to molecular diffusivity, i.e. $D_i^{eff} = \epsilon^{3/2} D_i$, where ϵ is the electrode porosity. The electrolyte velocity in porous media is governed by Darcy's law, in which a Kozeny-Carmen law is used for the hydraulic conductivity in the porous felt

$$\boldsymbol{v} = -\frac{d_f^2}{K\,\mu} \frac{\epsilon^3}{1-\epsilon^2} \nabla p \tag{6}$$

where d_f is the porous felt fiber diameter, K is the Kozeny-Carman constant, μ is the dynamic viscosity of the fluid, and p is the pressure. The velocity field can be derived from (6) by using the continuity condition for an incompressible liquid, i.e. $\nabla \cdot \boldsymbol{v} = 0$, and the assumption of dilute solutions. The continuity condition is useful to obtain advection-diffusion equations by combining (4) and (5). Neglecting the electro-migration term produced by the potential gradient, the following equation is obtained:

$$-D_i^{eff} \nabla^2\, c_i + \boldsymbol{v}\, \nabla c_i = S_i \tag{7}$$

which can be simplified, under the 1D assumption proposed in [5], as:

$$-D_i^{eff} \frac{\partial^2 c_i}{\partial x^2} + v_y \frac{c_i - c_i^{in}}{L} = S_i \tag{8}$$

where c_i becomes the average bulk concentrations (depending only on x), $v_y = \frac{Q}{w h_e}$ is the (equivalent) electrolyte velocity along the flow direction y, Q is the volumetric flow rate, w is the electrode width, and h_e is the electrode thickness.

The charge transfer from the electrolyte to the carbon felt occurs at the interface between solid and liquid phases. Charge balance equations are written for an equivalent volume element, considering thus an averaged volumetric current density:

$$\nabla \cdot \boldsymbol{J}_e = -\nabla \cdot \boldsymbol{J}_s = J_{vol} \tag{9}$$

where $\boldsymbol{J_e}$ and $\boldsymbol{J_s}$ are the current density in the electrolyte and in the solid phase, respectively. The volumetric current density is $J_{vol} = -A\,J_{g1}$ (in the positive electrode) or $J_{vol} = A\,J_{g2}$ (in the negative electrode), where A is the specific surface area (m^2/m^3) which denotes the electrochemical active area for unit volume of electrode. The current density inside the electrolyte is only given by the ion migration so that:

$$\boldsymbol{J_e} = \sum_i z_i\,F\,\boldsymbol{N}_i \tag{10}$$

which, by letting expression (5) for the species molar flux, gives:

$$\boldsymbol{J_e} = -K^{eff}\,\nabla\phi_e - F\sum_i z_i\,D_i^{eff}\nabla c_i \tag{11}$$

where $K^{eff} = \frac{F^2}{RT}\sum_i z_i^2\,D_i^{eff}c_i$ is the effective ionic conductivity. In the solid phase, the current density is only given by the electronic conduction, and, by applying the Bruggeman correction on the felt conductivity $\sigma_s^{eff} = (1-\epsilon)^{3/2}\,\sigma_s$, this becomes:

$$\boldsymbol{J_s} = -\sigma_s^{eff}\nabla\phi_s \tag{12}$$

Electrical conduction equations, needed to obtain the electric potential distributions, are finally obtained by letting (11) and (12) in charge balance equations.

The independent variables of the model in the porous electrodes are the electrolyte and the solid phase potential ϕ_e, ϕ_s, and the species concentrations $c_{V2}, c_{V3}, c_{V4}, c_{V5}, c_{H^+}$. The sulphate ion concentration $c_{SO_4^-}$ is finally obtained from the electro-neutrality condition:

$$\sum_i z_i c_i = 0 \tag{13}$$

Homogeneous Neumann boundary conditions are applied for ϕ_e, while homogeneous Dirichlet boundary conditions are applied at the positive electrode for ϕ_s. The current density J_{coll}, which can be positive (charge) or negative (discharge), is applied at the positive electrode. Inlet concentrations for all ionic species are related to the SoC of the electrolyte:

$$\begin{aligned}
c_{H^+,1}^{in} &= c_{H^+,1}^0 + c_{V,1}\,SoC \\
c_{H^+,1}^{in} &= c_{H^+,1}^0 + c_{V,1}\,SoC \\
c_{V_4}^{in} &= c_{V_4}^0 + c_{V,1}(1-SoC) \\
c_{V_5}^{in} &= c_{V_5}^0 + c_{V,2}\,SoC \\
c_{V_2}^{in} &= c_{V_2}^0 + c_{V,2}\,SoC \\
c_{V_3}^{in} &= c_{V_3}^0 + c_{V,2}\,(1-SoC)
\end{aligned} \tag{14}$$

where $SoC = \frac{c_{V5}}{c_{V,1}} = \frac{c_{V2}}{c_{V,2}}$ is the State of Charge ($SoC = 0$ when the redox flow cell is fully charged, $SoC = 1$ when it is fully discharged) and $c_{V,1} = c_{V4} + c_{V5}$, $c_{V,2} = c_{V2} + c_{V3}$ are the total vanadium concentrations at the positive and negative electrodes, respectively.

2.2 Membrane and current collector equations

Electric charges generated and consumed at porous electrodes are drawn by the current collectors according to Ohm's law for electronic conductors, i.e. $\boldsymbol{J_s} = -\sigma_c\,\nabla\phi_s$, where σ_c is

the (uniform) electronic conductivity. This relationship, combined with the charge conservation equation, provides:

$$-\nabla \cdot \sigma_c \, \phi_s = 0 \tag{15}$$

The polymer membrane provides the protonic exchange and only a ionic current occurs in it, so that electrons are forced to the external circuit. The electrical behavior of the membrane, which can be regarded as a solid electrolyte, is characterized by the conduction equation:

$$-\nabla \cdot \sigma_m^{eff} \, \phi_s = 0 \tag{16}$$

where $\sigma_m^{eff} = \frac{z_{H^+}^2 \, F^2 \, c_f \, D_{H^+}}{R \, T}$ is the effective conductivity of the membrane, with z_{H^+} proton valence, c_f concentration of fixed charge sites, and D_{H^+} the proton diffusion coefficient.

2.2 Dynamic model equations

During VRFB operations the species concentrations in the reservoirs vary due to the processes occurring in the half-cells. Having assumed the conservation of volumes [2], the volumetric flow rate at the outlet of the electrodes is $Q = v_y \, w \, L$. The outlet concentrations are calculated by averaging along the outlet surfaces. The changes in species concentrations at the inlet of the electrodes due to recirculation can be calculated from the following material balance, which assumes perfect mixing and negligible reaction in the reservoir of volume V:

$$\frac{dc_i^{in}}{dt} = \frac{Q}{V}\left(c_i^{out} - c_i^{in}\right), \qquad c_i^{in}(0) = c_i^0 \tag{17}$$

The reservoir volume is obtained by subtracting the electrode volume ($\varepsilon \, L w h_e$) from the volume of electrolyte on each side of cell, V_T . The electrolyte volume can be estimated from Faraday law once the electrical current I and the charge time Δt_{ch} have been fixed, as:

$$V = \frac{I \, \Delta t_{ch}}{F\left(c_{V5}^f - c_{V5}^0\right)} \tag{18}$$

where c_{V5}^0 and c_{V5}^f are the initial and the final (estimated) concentrations of VO_2^+. The volume of electrolyte in the pipes and pump is assumed to be negligible. To account for the evolution in species concentration, the conservation equation (7) is modified as follows:

$$\frac{\partial \, \epsilon \, c_i}{\partial \, t} = S_i + D_i^{eff} \nabla^2 \, c_i - \boldsymbol{v} \, \nabla c_i \tag{19}$$

and a similar expression, corresponding to (8), is obtained under the 1D assumption:

$$\frac{\partial \, \epsilon \, c_i}{\partial \, t} = S_i + D_i^{eff} \frac{\partial^2 c_i}{\partial x^2} - v_y \frac{c_i - c_i^{in}}{L} \tag{20}$$

3 NUMERICAL RESULTS

The 1D redox flow battery model is constructed from the previous non-linear equations: momentum, charge and species conservation, mass transport (Nernst-Plank equation), electric conduction, charge generation (Butler-Volmer equation) inside the porous electrodes, and the proton conduction inside the membrane. The basic assumption of the model regards the derivation of equation (8), and consists in approximating the concentration gradient inside the porous electrode with a finite difference. The partial differential equations (PDEs) for the 1D model are then simply expressed by substituting the gradient operator with ∂_x (derivative along the horizontal $x-$axis in Fig. 2) where electrical conduction and species diffusion occur. The numerical model is implemented in COMSOL®, which provides a robust simulation environment for multiphysics problems. This software, based on the finite element method (FEM) for the discretization of 1D, 2D, 3D PDEs, provides different non-linear solvers.

In order to assess the validity and the accuracy of the dimensional reduction numerical results of the 1D model are compared with those of a 2D model based on the same physical parameters. The meaning of the geometrical parameters of the 2D model is reported in the schematic of Fig. 2. The 1D model geometry consists in the union of intervals $[X_i, X_{i+1}]$.

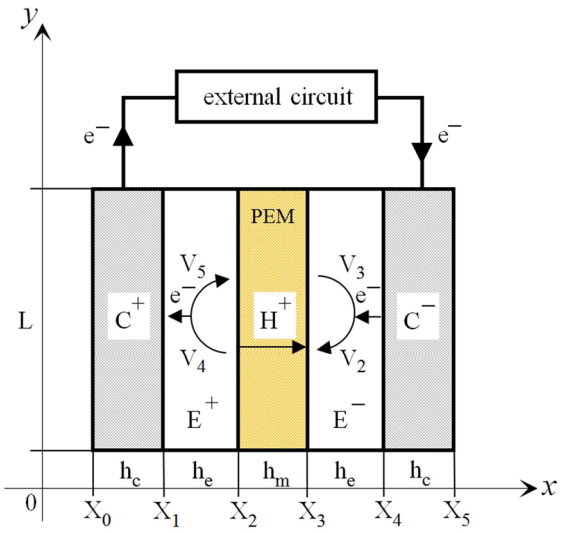


Figure 2: Geometrical parameters in the schematic diagram of 2D VRFB model (C^+, C^-: collectors; E^+, E^-: porous electrodes; PEM: polymer membrane) implemented in COMSOL Multiphysics®.

Due to the generality of PDEs defined in the fully coupled model, "General Form PDE" study nodes were used to implement both conduction and advection-diffusion equations in COMSOL. In particular, (8) was implemented as a diffusion-adsorption equation in order to account for the concentration gradient approximation. The degrees of freedom (DoF) of the discrete algebraic system are concentrations $c_{V4}, c_{V5}, c_{H_1^+}$ (positive electrode) $c_{V2}, c_{V3}, c_{H_2^+}$ (negative electrode), the electrolyte potential ϕ_e (membrane and electrodes), and solid phase potential ϕ_s (electrodes and collectors). These equations, discretized with quadratic finite elements (for both 1D and 2D) defined on structured meshes, were solved by the Newton-Raphson method with a fixed tolerance of 10^{-6}.

The following computing performances were observed in both FEM steady-state models (1D vs. 2D) run on a Intel® Core™@2.70 GHz Processor, 64 bit system, 4 GB RAM: 7.235 vs. 127.361 (number of DoF); 26 s vs. 899 s (solution time); 890 MB vs. 1.89 GB (physical memory); 1.094 MB vs. 2.16 GB (virtual memory). Steady-state solutions of the fully coupled problem were obtained by a parametric sweep (for SoC values ranging from 0.1 to 0.9), after imposing the electric current density J_{coll} at the positive electrode and the inlet species concentrations (14). Geometrical and physical parameters in Tables I and II, used for both 1D and 2D models, were taken from [5] except for the thickness of collectors, which is here presumed. The electrical behavior of the redox flow battery was investigated by computing the electrical cell voltage between the positive and negative current collectors.

Fig. 3 shows that charge and discharge curves of 1D and 2D models are in very good agreement, also with experimental data presented in [5]. The electrolyte and the electrode current density (for the 2D model, in particular, the x-component is considered) in Fig. 4 show a similar behavior. Product and reactant species concentrations at positive and negative electrodes are compared by considering, for the 2D model, the concentration profiles at the electrode outlets ($y = L$). Also in the case of non-electrical variables it can be observed that the reduced model provides high accuracy, even though its complexity is much smaller. The concentration variation inside the device is captured by the 2D model, i.e. Fig. 7 and 8 show that for reactants the concentration is higher near the inlet, whereas for products the concentration is higher near the outlet.

Dynamic 1D and 2D models are constructed by adding the reservoir equation and by enforcing time-varying boundary conditions at the electrode inlets. Time profiles of model variables were obtained by an adaptive time-stepping backward Euler solver. Computing time and memory requirements, with the same computer as above, were 710 s and 1.86 GB for the 2D model, and 232 s and 1.91 GB for the 1D model. Fig. 9 and 10 show that the reduced model again provides an accurate information about multiphysics problem variables.

4 CONCLUSIONS

A one-dimensional steady-state multiphysics model of a single VRFB, encompassing mass, charge, and momentum transport and conservation has been proposed. Compared to previous model in literature current collector have been included aiming at simulating real VRFB stacks. Comparisons with 2D steady-state multiphysics model show that the 1D assumption is reliable to properly simulate both electrical and chemical quantities. Starting from these models, reservoir equations and mass balance equations have been included in order to simulate charge-discharge operations and, in particular, to estimate the battery run-time.

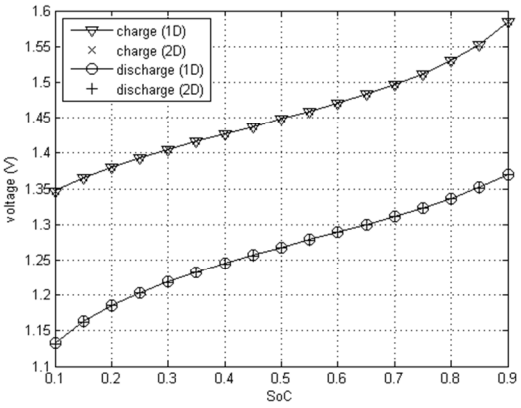


Figure 3: VRFB cell voltage vs. State of Charge.

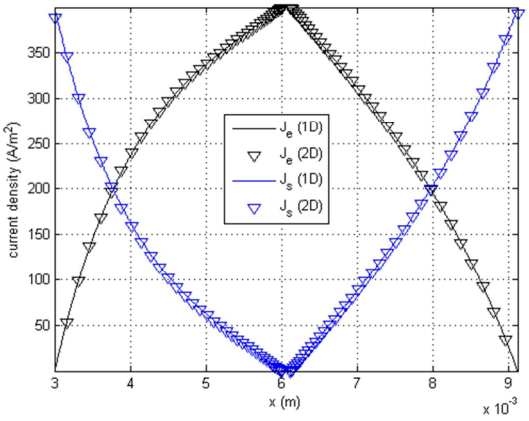


Figure 4: Electrolyte (J_e) and electrode (J_S) current density in electrode + membrane region.

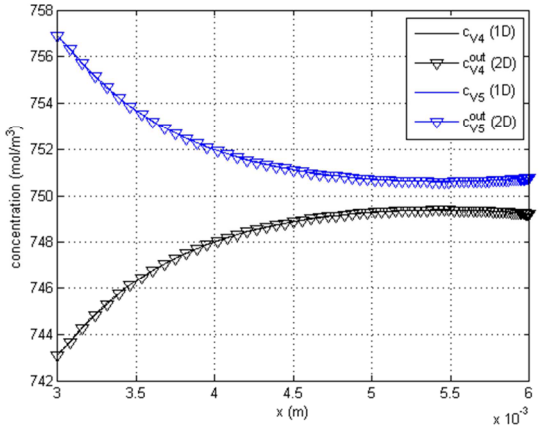


Figure 5: Positive electrode species concentrations c_{V4}, c_{V5} during charge phase ($SoC = 0.5$).

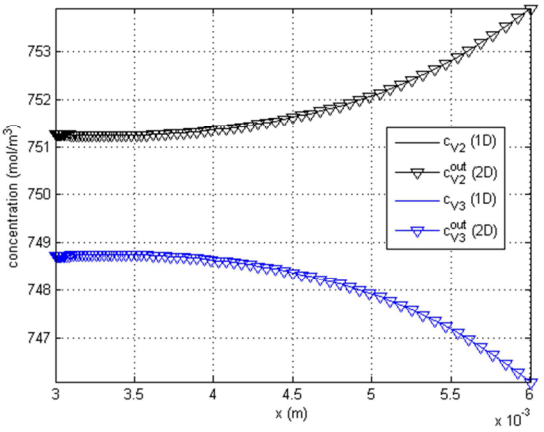


Figure 6: Negative electrode species concentrations c_{V2}, c_{V3} during charge phase ($SoC = 0.5$).

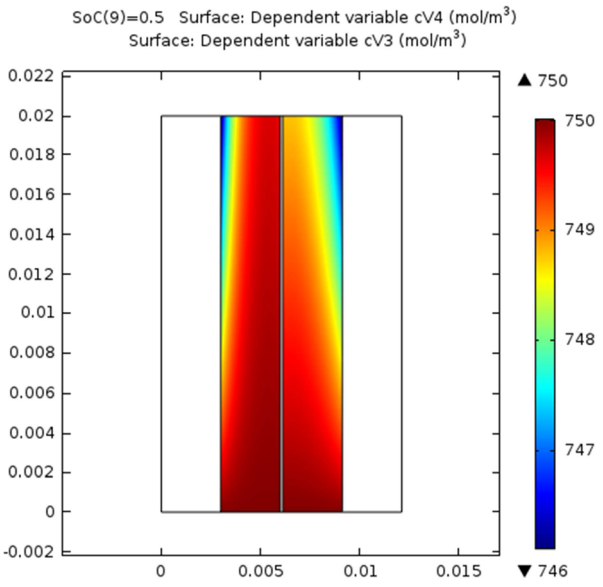


Figure 7: Reactant concentration c_{V4}, c_{V3} distributions during the charge phase (2D model, $SoC = 0.5$).

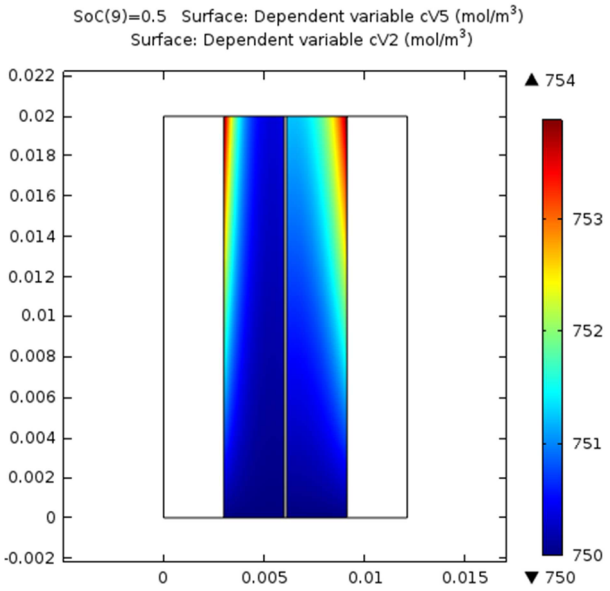


Figure 8: Product concentration c_{V5}, c_{V2} distributions during the charge phase (2D model, $SoC = 0.5$).

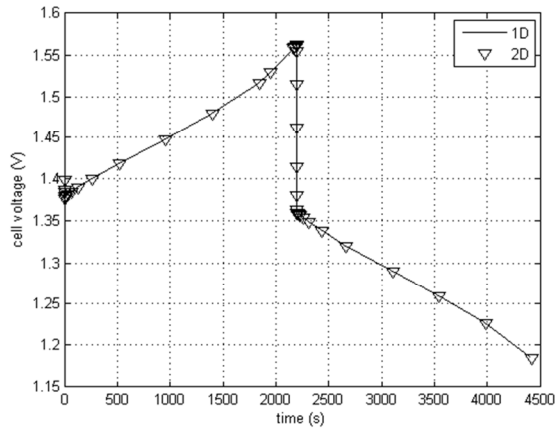


Figure 9: Cell voltage time profile during charge ($\Delta t_{ch} = 2200$ s) and discharge phases.

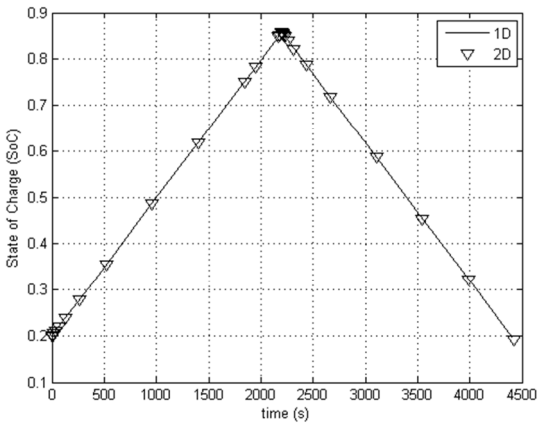


Figure 10: State of charge time profile during charge ($\Delta t_{ch} = 2200$ s) and discharge phases.

Table 1: Geometrical parameters of the 1D–2D VRFB model

h_m	$1.25 \cdot 10^{-4}$ m	PEM thickness
h_e	$3.00 \cdot 10^{-3}$ m	electrode thickness
h_c	$1.25 \cdot 10^{-4}$ m	collector thickness
L	$2.00 \cdot 10^{-2}$ m	cell length
w	$2.50 \cdot 10^{-2}$ m	cell width
V	$3 \cdot 10^{-6}$ m^3	reservoir volumes

Table 2: Physical parameters of the 1D–2D VRFB model

ϵ	0.929	electrode porosity
D_H	$9.3 \cdot 10^{-9}\epsilon^{1.5}$ m/s	H^+ diffusivity
D_{V4}	$3.9 \cdot 10^{-10}\epsilon^{1.5}$ m/s	V4 diffusivity
D_{V5}	$3.9 \cdot 10^{-10}\epsilon^{1.5}$ m/s	V5 diffusivity
D_{V2}	$2.4 \cdot 10^{-10}\epsilon^{1.5}$ m/s	V2 diffusivity
D_{V3}	$2.4 \cdot 10^{-10}\epsilon^{1.5}$ m/s	V3 diffusivity
D_{SO4}	$1.1 \cdot 10^{-9}\epsilon^{1.5}$ m/s	SO4 diffusivity
Q	60 ml/min	molar flow rate
v_y	$Q/h_e w\epsilon$	electrolyte velocity
T	300 K	cell temperature
σ_s	$1000 \cdot (1-\epsilon)^{1.5}$ S/m	solid phase conductivity
$E_{0,1}$	1.004 V	std potential (+)
$E_{0,2}$	-0.255 V	std potential (–)
K_1	$6.8{\cdot}10^{-7}$ m/s	rate constant (+)
K_2	$1.7{\cdot}10^{-7}$ m/s	rate constant (–)
A	$1.62{\cdot}10^4$ m^2/ m^3	specific area
α_1	0.5	transfer coefficient
α_2	0.5	transfer coefficient
J_{coll}	$\pm$ 400 A/m^2	collector current density (+: charge)
σ_c	1000 S/m	collector conductivity
$c_{V,1}$	1500 mol/m^3	tot concentration (+)
$c_{V,2}$	1500 mol/m^3	tot concentration (–)
$c^0_{H^+,1}$	6000 mol/m^3	H^+ concentration (+)
$c^0_{H^+,2}$	4500 mol/m^3	H^+ concentration (–)
c_f	1200 mol/m^3	fixed charge concentration (PEM)
$D_{H,m}$	$3.5 \cdot 10^{-10}$ m/s	H^+ diffusivity (PEM)
z_H	1	H^+ valence
z_{V2}	2	V2 valence
z_{V3}	3	V3 valence
z_{V4}	2	V4 valence
z_{V5}	1	V5 valence
$K_{m,1}$	$1.6 \cdot 10^{-4} ms^{-1}\left(v_y/1\,ms^{-1}\right)^{0.4}$	mass transfer coefficient (+)
$K_{m,2}$	$1.6 \cdot 10^{-4} ms^{-1}\left(v_y/1\,ms^{-1}\right)^{0.4}$	mass transfer coefficient (–)

REFERENCES

[1] Alotto, P., Guarnieri, M., and Moro, F., Redox flow batteries for the storage of renewable energy: A review, *Renewable and Sustainable Energy Reviews* (2014) **29**:325-335.
[2] Shah, A.A., Watt-Smith, A.A., and Walsh, F.C., A dynamic performance model for redox-flow batteries involving soluble species, *Electrochemica Acta* (2008) **53**:8087-8100.
[3] Knehr, K. W., et al., A Transient Vanadium Flow Battery Model Incorporating Vanadium Crossover and Water Transport Through the Membrane, *Journal of the Electrochemical Society* (2012) **159**:A1446-A1459.
[4] You, D., Zhang, H., and Chen, J., A simple model for the vanadium redox battery, *Electrochemica Acta* (2009) **54**:6827-6836.
[5] Chen, C.L., Yeoh, H. K., and Chakrabarti, M. H., An enhancement to Vynnycky's model for the all-vanadium redox flow battery, *Electrochemica Acta* (2014) **120**:167-179.

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis (Eds)

MULTIPHYSICS SIMULATION TOOLS FOR DESIGNING MOTORS FOR TRACTION APPLICATIONS IN HYBRID AND ELECTRIC VEHICLES

NIKHIL P. SAWANT[*], CHRISTOPHER P. RILEY[*], ARTURAS VENSKUS[*], DANAIL H. VASSILEV[*], JOHN D. WALE[†], EDDIE WEARING[†], ALEX M. MICHAELIDES[#], PHILIP J. TOPPING[#], HERMINDER MATHARU[#]

*Cobham Technical Services,
Network House, Langford Locks, Kidlington, OX5 1LH, UK
e-mail: vectorfields.info@cobham.com

[†]Ricardo UK Ltd
Midlands Technical Centre
Southam Road, Radford Semele, Leamington Spa, CV31 1FQ, UK
e-mail: electric.drives@ricardo.com

[#]Jaguar Land Rover Limited
Abbey Road, Whitley, Coventry CV3 4LF, UK
e-mail: hmathar2@jaguarlandrover.com

Key words: Multiphysics, motors, electric vehicle, Opera, python, simulation.

Abstract. Motor manufacturers are facing a difficult challenge in designing traction motors for the latest generation of hybrid and all-electric vehicles. The efficiency with which these motors can perform is critical, as it impacts on the vehicle range and battery life. Many of the issues involved in the motor design have a complex nature which requires multiple fields of physics such as electromagnetics (EM), mechanics and thermal analysis. All these physics are usually interdependent and have to be considered collectively in order to obtain optimal performance for a particular scenario.

This paper presents a multiphysics simulation tool that was implemented to address this situation. The Opera FEA software suite [1] was developed to include a multiphysics analysis that can link several EM, thermal and stress analyses. Opera's Machines Environment (parameterised template software for designing motors and generators) has been extended to allow easy setup of coupled multiphysics analyses such as EM to thermal and EM to stress. In order to further facilitate the coupling of different analyses, a link to the Python programming language was embedded in Opera FEA software. The embedded Python facility offers options to perform certain post-processing operations during the solving stage and hence allow data transfer between different stages of the multiphysics analysis. It also extends Opera's capabilities to interact with other FEA software.

1 INTRODUCTION

With growing concerns over future petroleum supplies and air pollution, sustainable transportation has emerged as a vital mission for countries that seek to expand infrastructure

whilst minimising the associated economic and environmental costs. Hybrid electric vehicles (HEVs) and electric vehicles (EVs) represent one of the most promising routes to better energy security and reduced emissions of greenhouse gases. The research in HEVs/EVs has been widely supported by governments worldwide. The U.S. Department of Energy predicts that by 2030, alternative vehicles will comprise 28% of the total U.S. light-duty cars and trucks–a 20% increase from 2005 [2]. Consequently, there is a strong competition to develop efficient and cost-effective electric powertrains in order to meet these demands by focussing on simulation based development rather than prototype testing.

The motor/generator plays a crucial role in driving the vehicle. Traction motors for EVs are different from other motors because they must work reliably in a severe physical environment. Motors must operate consistently under extreme hot and cold temperatures, hard duty cycles and rough road conditions. All of these variables must be addressed in motor design.

Electrical machine designers have to consider the following fields of physics to obtain optimal behaviour of a particular machine:

- Electromagnetics (EM)
- Mechanics (stress)
- Thermal characteristics

All these physical phenomena usually have an effect on one another and have to be considered collectively. So, machine designers have to consider a combination of two or more physics problems while optimizing the overall design. Also, the focus is on getting very accurate results in less time.

In this paper, various tools implemented in Opera for characterizing machine performance within a multiphysics scenario are discussed. The paper then presents the coupling options implemented in Opera's Machines Environment. All the developments discussed in this paper were developed as part of a collaborative development programme named the 'Rapid SR' project [3] between Cobham Technical Services, Jaguar Land Rover and Ricardo, part supported by Innovate UK (previously known as the Technology Strategy Board).

2 MULTIPHYSICS ANALYSIS

In order to evaluate accurate performance of electric motors for traction application, it is necessary to assess the interaction between the various physics arising from the natural phenomena. For example, material properties are usually temperature dependent. A multiphysics analysis tool to address this situation has been implemented that can chain EM, thermal and/or stress analyses. Results from one analysis are automatically transferred as input to any subsequent analysis in the chain. Table 1 shows the exchange of data between different analyses.

For machines where the air gaps are small, it is important to include the structural changes evaluated by the stress analysis to characterize the machine performance. This can be achieved by chaining magnetostatic, stress and magnetostatic solutions, as shown in figure 1. Element force densities evaluated by the Maxwell stress in the magnetostatic analysis are supplied to the stress analysis, along with the rotational forces on the rotor. The last stage of the chain uses the displaced nodes to re-evaluate the electromagnetic behaviour of the machine. This loop can be continued iteratively until convergence is obtained.

Table 1: Data exchange between different physics [4].

		Fields Calculated		
Inputs used		**EM**	**Thermal**	**Stress**
	EM		Temperature	Displacements
	Thermal	Heat density from coils, and/or eddy currents		Displacements
	Stress	Magnetic flux density, Element force densities	Temperature	

For a multi-case analysis, where the electromagnetic simulation consists of different excitations (for example, different currents or frequencies) then the chain is repeated for all the cases. For example, for a solution at two current excitations chained with stress and a second EM, the multiphysics analysis will run a chain of EM, stress and EM for each excitation (six simulations in total).

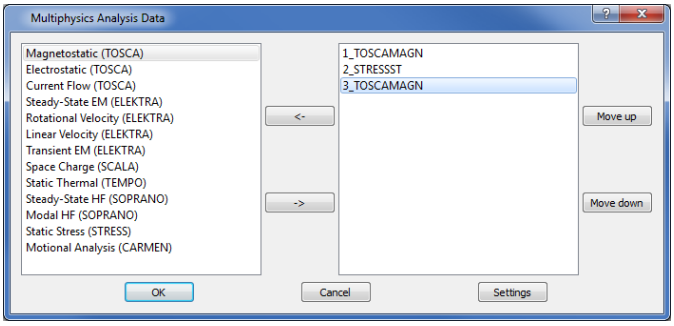


Figure 1: Multiphysics setup showing chained analyses

For a transient analysis, where the electromagnetic simulation is a time-stepping solution, results may be averaged over a cycle and passed to the next analysis in the chain. For example, a transient electromagnetic simulation, which is set up to run for one cycle after reaching steady-state, is chained with a steady state thermal simulation. The program will automatically average the results (electromagnetic losses) from the EM solution over the one cycle and pass that average on to the thermal simulation to evaluate temperature rise.

3 PYTHON

A multiphysics analysis where different analyses are chained involves evaluating certain quantities from a particular solution, which are passed on to the next analysis as input. In some scenarios, this involves accessing the simulation while it is still running. In order to address this, Python (programming language) was embedded in Opera. Python's functionality such as advanced string manipulation and numerical analysis components were the key features for it being a viable solution for addressing the requirements of the multiphysics analysis. The integration between Python and Opera has also enabled easy interaction of Opera with other 3rd party scientific programs. The following sections give an overview of how the Opera's functionality was extended by using Python:

3.1 Opera Database Extraction

The Opera Database Extraction package contains functionality to extract a data set of results from a particular simulation in a sequence (for example, one case is a transient simulation), and to subsequently store a subset of this data in an easily referable manner. The primary purpose of this is to allow user computations to be performed during the solving stage, although it may equally be used in a post-processing stage. For example, it may be used to extract x, y and z components of flux density, and store only the subset of this data relevant to ferrous regions (rotor or stator steel) throughout a transient simulation. Such data may then be used once the transient period has completed to evaluate iron loss.

3.2 Opera Steady State Detection

The package Opera Steady State Detection provides functionality to determine steady state in transient simulations (i.e. the time at which the behaviour of the system becomes time periodic). It includes a sequential analysis sub-package, which provides two detection methods that can be applied to global simulation data such as torque, speed, current etc. Sequential analysis solves hypothesis-testing problems in which the sample size is not fixed a priori but depends upon the data that have been already observed. While rudiments of sequential analysis date back to the works of Huyghens, Bernoulli, DeMoivre, Lagrange and Laplace as briefly noted in [7], the theory of sequential analysis was formulated by A. Wald in his celebrated work [8]. Over the recent years sequential analysis has become an important tool for handling real world problems arising across various branches of science and engineering: environment surveillance and monitoring, navigation and radar signal processing, fault detection, quality control, biomedical signal processing, econometrics and financial markets. The common feature of the above problems is the goal to detect one or several abrupt changes in some characteristic properties of the considered object.

In the first method, called a "variance test" following [9] parameter estimation is performed on collected simulation data. Once the required criteria for steady state detection have been satisfied, further sampling is disabled. The stopping criterion is based on the minimization of a cost function that measures the data heteroscedasticity, the tendency of the standard deviation of the observed quantity to vary.

The second sequential analysis detection method, referred to as the "mean test", is a version of CUSUM algorithm [10] specially designed to detect the transient time at which the mean of the data settles down to zero. It may be used, for example, to analyse the DC offset in a transient simulation. A log-likelihood function for the samples is computed and compared to a threshold value to decide whether the parameters of the underlying probability distribution represent steady or transient state. In transient electromagnetic simulations, the data (e.g. current, voltage) is periodic and the assumption is that the samples follow an arcsine distribution, in which the amplitude and the offset are the control parameters. The nominal values of these parameters define the steady state and need to be specified along with the threshold for the log-likelihood function.

3.3 Opera Loss Calculation

The package Opera Loss Calculation provides functionality to calculate losses over a steady state cycle and to subsequently upload such losses to an Opera database. The package

evaluates losses based on current density (J) and material conductivity (σ) and iron losses. The iron losses are based on the Steinmetz formulae and use the following methodology:

- A full harmonic analysis is performed on the flux density waveform on an element-by-element basis, to an order which is user specified. The number of sampling steps is also user-specified, and should be a function of the highest harmonic order chosen by the user, ensuring that the stored flux density variation with angle (time) contains enough information for correct harmonic amplitude extraction.
- Element loss intensity tables are created which are then automatically transferred to the thermal model.

3.3 Coupling to other software

The close integration between Python and Opera has enabled easy extension of Opera's functionality with 3rd party Python packages and other FEA systems. In order to extend the functionality, any new Python package needs to be installed preferably (but not necessarily) into the Opera Python distribution location and the new functionality can be accessible immediately via a Python "import package_name" statement. In such a manner, a package like *matplotlib* can provide functionality to perform 3D plots where the data can be supplied dynamically, for example during solver solution.

More advanced Python usage can be, for example, chaining Opera computations with other FEA solvers to perform acoustic, fatigue, fracture, dynamic vibration and other required analysis for electrical machines. The open source structural mechanics solver from EDF Code_Aster [6] can be used for this purpose. The following Python script snippet, shown in figure 2, demonstrates how to prepare FE data to Code_Aster using its native data exchange format MED (HDF based) library.

```
# create index class to manage elements of particular geometry
group
geom_label = "STEEL"
multiIndex = odi.opera_multi_index_by_element_class([geom_label])
multiIndex.generateIndices(simu)
index_type = "NODAL"
nodal_indices = multiIndex.getIndex(geom_label, index_type)
# set mesh spatial dimensionality, name and node count
mdim = 3;
mname  = "mesh1";
mncount = nodal_indices.size;
# create numpy array to accommodate nodal coordinates
coord = np.empty(nodal_indices.size*mdim)
for i in range(nodal_indices.size):
    loc_index = i*3
    node_index = int(nodal_indices[i])
    node_coord = simu.getCoordForIndex(index_type, node_index)
    coord[loc_index]   = node_coord[0]  # x
    coord[loc_index+1] = node_coord[1]  # y
    coord[loc_index+2] = node_coord[2]  # z
# convert to MED float data type format
coo = MEDFLOAT(coord)
# create table name and unit headers
namecoo = "x               y               z               "
unicoo  = "m               m               m               "
# create and open MED (HDF) file
fid = MEDfileOpen("plate.med", MED_ACC_CREAT)
# create mesh
MEDmeshCr( fid, mname, mdim, mdim, MED_UNSTRUCTURED_MESH,
    "mesh from Opera", "s", MED_SORT_DTIT,
    MED_CARTESIAN, namecoo, unicoo)
# write nodal coordinates
MEDmeshNodeCoordinateWr(fid,mname,MED_NO_DT,MED_NO_IT,MED_UNDEF_DT,
    MED_FULL_INTERLACE,nodal_indices.size,coo)
# close MED file
MEDfileClose(fid)
```

Figure 2: Code snippet for transferring mesh data.

The Python script above demonstrates only how to export FE nodes into MED file format. Similarly other FE mesh data (nodal connectivity, faces, and edges) can be exported using Python interfaces from the MED library. The same interface is used to extract field values back from an MED file after a Code_Aster analysis is performed.

4 MACHINE DESIGN ENVIRONMENT

The machines environment is a design package, which capitalizes on the accuracy of Finite Element Analysis of the Opera Suite to offer various options to the machine designer, including standard analysis options required for the design process. The machines environment was extended to exploit the developments in Opera to provide standard multiphysics scenarios for designing motors for traction applications. Two such scenarios will be described in detail using a practical example of a switched reluctance motor, as shown in figure 3. Table 2 shows the specifications of the SRM considered.

Table 2 Specification of SRM

Parameters	Values (cm)
Stator outer radius	**3.9**
Rotor outer radius	**2.165**
Air gap	**1.735**
Stack length	**8.16**
Number of turns per phase	**35**
Stator slots	**8**
Rotor poles	**6**

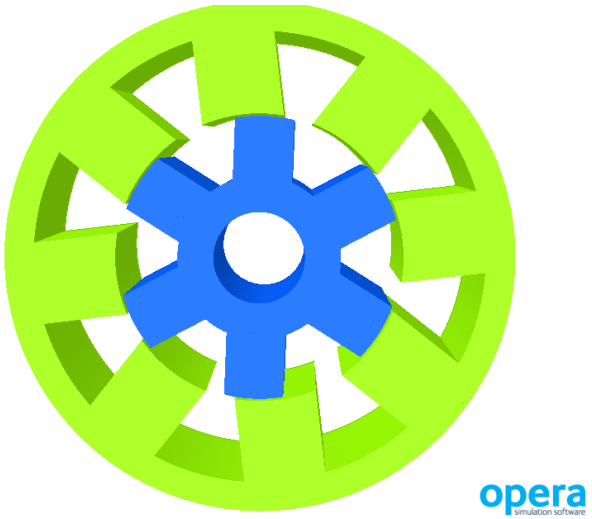


Figure 3: SRM geometry modelled in Opera

4.1 Electromagnetic and Stress Analysis

An analysis required to evaluate how induced stress in various parts of the machine affects its electromagnetic performance was implemented by coupling the electromagnetic and stress analyses. This analysis also includes stress induced by centrifugal loading. The electromagnetic analysis creates a series of static jobs for different rotor positions. The performance of the machine is characterized by evaluating peak torque obtained for the given excitation. For each rotor position, a stress stage was added to evaluate the stress due to electromagnetic and centrifugal loadings.

Convergence criteria with an iteration limit can be set by the user to decide whether further additional EM and stress stages should be added or not. Following are the two criteria which will be observed:

- *Absolute change on the deformation*: For each of the rotor positions, the maximum displacement as a percentage of the air gap radius on the rotor faces is recorded. If this value, for any of the rotor positions, is greater than that specified by the user then the EM and stress stages are repeated up to the iteration limit.
- *Relative change between the iterations*: The change in the percentage displacement between the iterations is also recorded. If this value is greater than that specified by the user then the EM and stress stages are repeated up to the iteration limit.

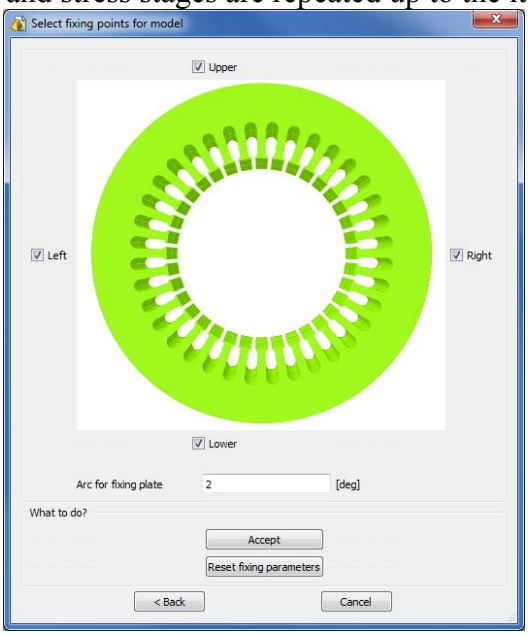


Figure 4: Option to select fixing point for the stator.

In order to accurately evaluate stress in the machine, appropriate boundary conditions need to be specified. Figure 4 shows the standard options that users can choose for the stator. For the rotor, the fixing points will be represented by the extremities of the shaft where a fixed displacement boundary condition will be applied. The length of the shaft can be set independently as compared to the length of the rotor to allow for a more realistic representation of the stresses produced in the device.

Figure 5 shows the electromagnetic characteristics (torque vs. rotor position) of the SRM at the end of simulation after accounting for the stress effects.

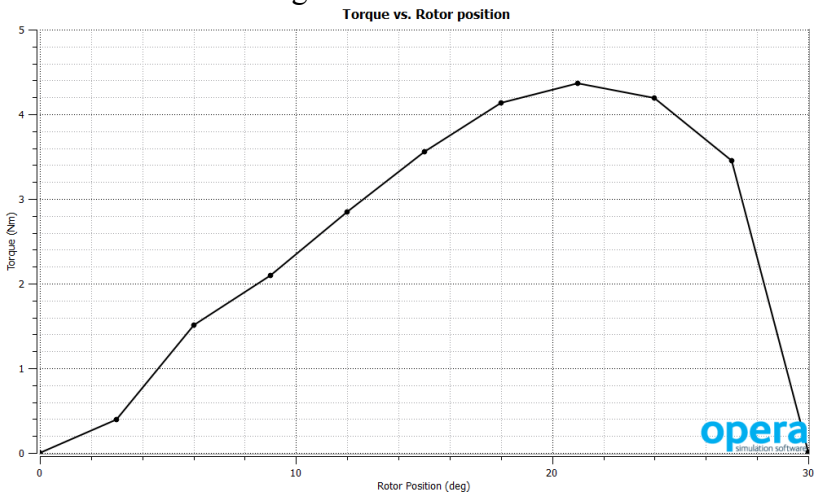


Figure 5: Torque vs rotor position.

Figure 6 shows an exaggerated view of the displacements due to the stress effects, where the winding displayed on the stator tooth is to show which of the phases was excited. The coloured plot shows that the displacement due to centrifugal loading is dominant and hence an even distribution across all the rotor teeth. Figure 7 shows the effect of stress on the air gap of the machine. For both the figures, the mesh shows the actual position of the rotor and the solid view shows the deformed position.

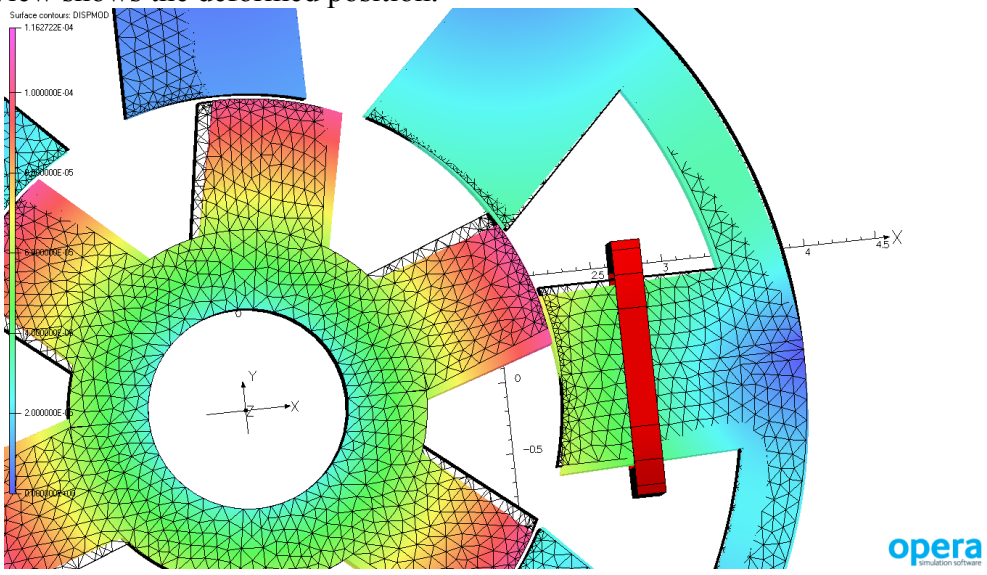


Figure 6: Displacements due to stress effects.

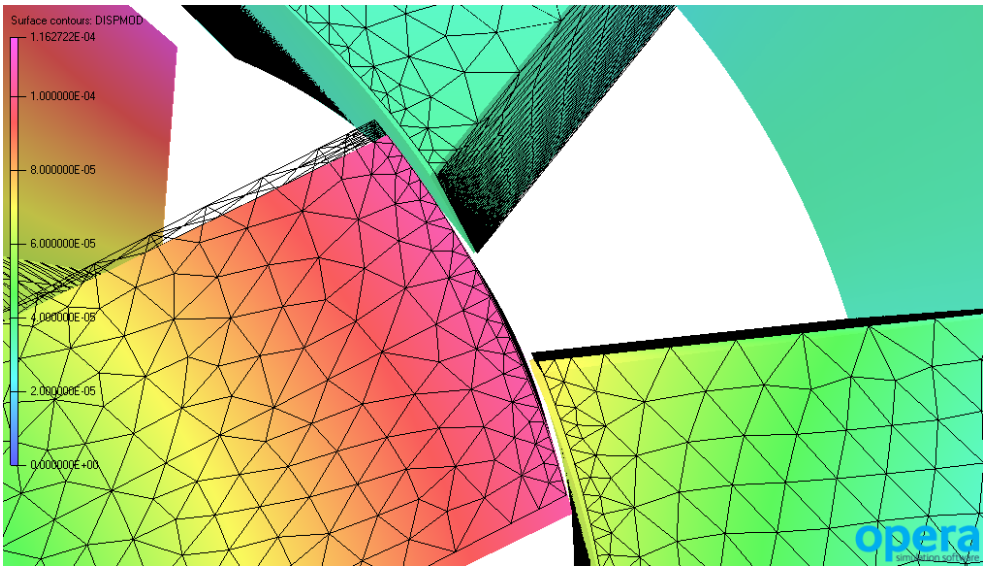


Figure 7: Gap closure due to stress effects

4.2 Electromagnetic and Thermal Analysis

An analysis required to evaluate how temperature rise in a machine affects its electromagnetic performance was implemented by coupling the electromagnetic and thermal analyses. A transient mechanically coupled analysis with rotational motion is run for a machine with a user defined excitation. A default switching strategy is calculated based on the geometric parameters and the number of phases selected, but a custom strategy can also be defined. A hysteretic speed control is also implemented in order to maintain the speed at the defined level. After the transient stage, a static thermal stage is added where the losses calculated over the last cycle of the transient stage are used as input.

This analysis uses the steady state detection package developed using Python in Opera in order to determine the start of the steady state. The analysis is run for one cycle after the steady state is detected. The loss calculation package of Python in Opera is used to perform the calculations over the one steady state cycle and pass the averaged losses over the cycle to the thermal stage.

For the thermal stage, the Stator and Rotor core thermal conductivities can be defined as anisotropic having different coefficients in the in-plane (radial/azimuthal) and axial directions. The conductor thermal conductivity is homogenized as a combination of conductivities for the conducting material and insulation and accounts for the fill factor. The boundary conditions in the form of heat transfer coefficients for the core back and air gap are set along with an ambient temperature.

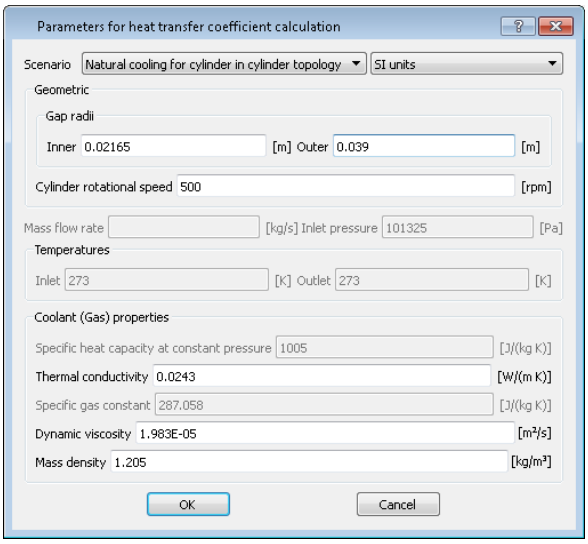


Figure 8: Tool for calculating heat transfer coefficients.

The heat transfer coefficient for the air gap of the machine is evaluated using the Heat Transfer Tool in Opera, developed in collaboration with City University, London [5], the interface to which is shown in figure 8. It calculates an estimate of the heat transfer coefficient for a cylindrical surface of a machine rotor, rotating in a larger cylindrical space, i.e. a rotor within a stator. The formulae used for the calculation have been collated from theory and experiment over different regimes of dimensions, rotation speeds and coolant properties. While this simple theory may only provide limited accuracy, the coefficients given provide good initial estimates and avoid the need for expensive computational fluid dynamics simulation before thermal analysis can proceed.

Figure 9 shows the torque profile of the SRM obtained by running the analysis while maintaining the speed around 500 rpm, as shown in figure 10. Figures 11 and 12 show the temperature rise in the stator and rotor sections respectively. This analysis exploited Opera's functionality of setting a 2d slice [4] of full 3d models for rapid simulation, but, consequently, neglects axial heat flow.

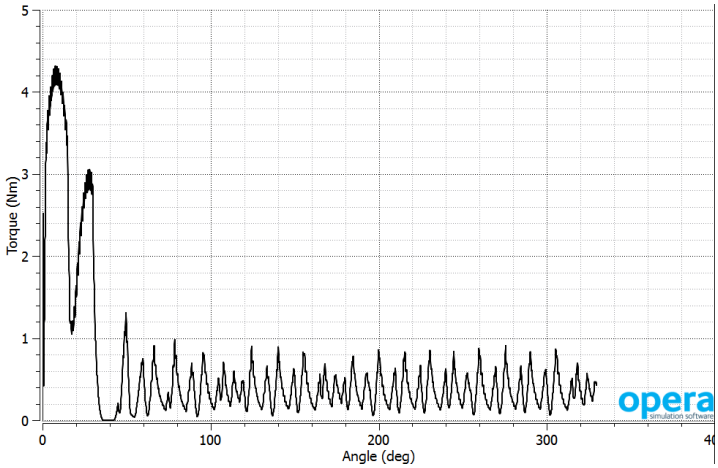


Figure 9: Torque vs. rotor position obtained from CarmenRM analysis

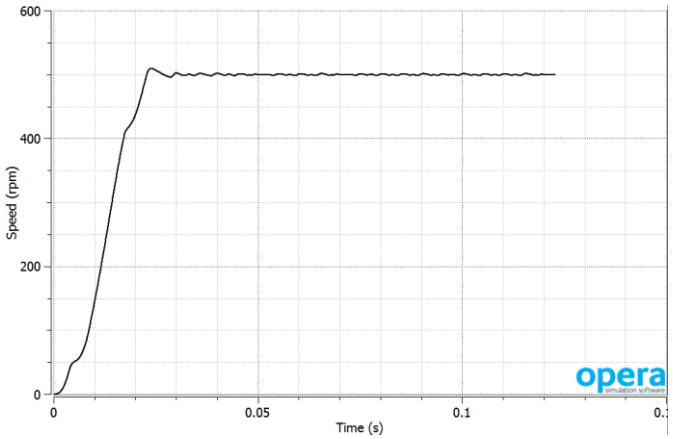


Figure 10: Speed vs time showing the hysteretic control of speed around 500 rpm

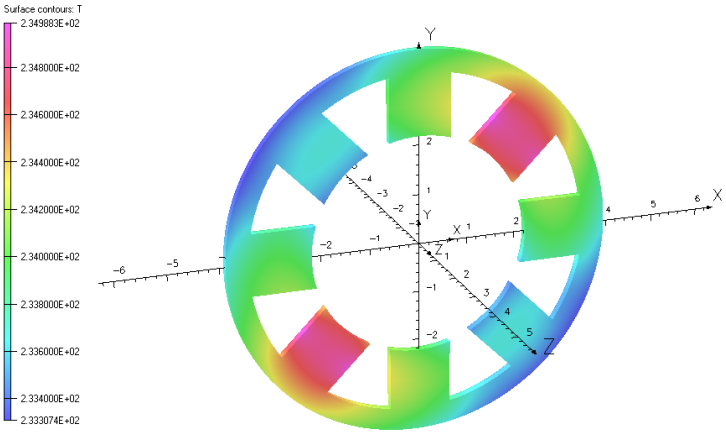


Figure 11: Temperature rise in the stator section

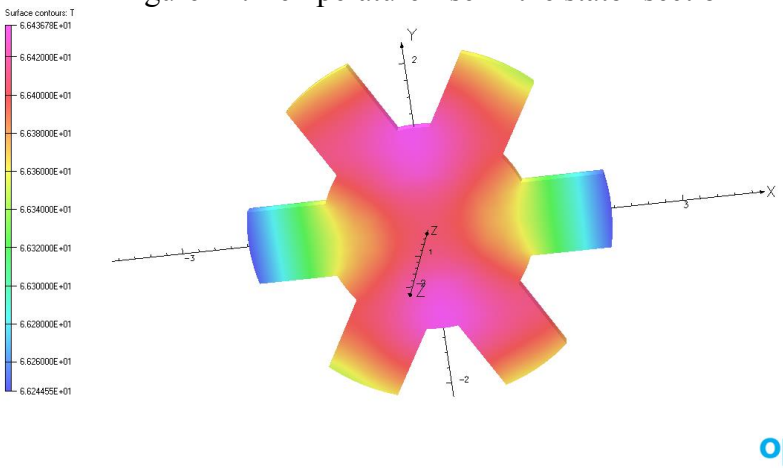


Figure 12: Temperature rise in the rotor section

5 CONCLUSIONS

The impetus for understanding the behavior of electrical machines has increased, following the increased on-going research in HEV/EV for a sustainable future. An in-depth knowledge of the interactions between electromagnetics, thermal and mechanical behavior of machines has become vital in order to understand the material performance, reliability and electromagnetic characteristics in extreme environments and at high speeds. This has driven the rise in interest in multi-physics simulations.

Software tools for characterizing machine performance based on various physics were discussed. Also, tools developed in the machines environment for coupling various fields of physics in order to get the optimized designs were discussed. The tools developed were successfully applied to an 8/6 SRM in order to understand and quantify its advantages. Further work as a part of the Rapid SR project will be carried out in order to verify the simulation results against practical measurements.

ACKNOWLEDGEMENT

The authors gratefully acknowledge the support of Innovate UK in this work.

REFERENCES

[1] Cobham Technical Services, "Opera Electromagnetics Finite Element Software Suite", www.operafea.com
[2] Annual Energy Outlook 2007 with Projections to 2030, report DOE/EIA.
[3] Technology Strategy Board Collaborative R&D Project 400233 http://motor-design-software.com/news/research-programme-targets-nextgeneration-electric-motors-luxury-automobiles
[4] C P Riley, et al., "Simulation based design of reluctance motors for traction applications in hybrid and electric vehicles", PEMD, Manchester 2014
[5] K.R. Pullen and B. Yoheswaran, Private Communiucation.
[6] EDF, "Code_Aster", http://www.code-aster.org
[7] B.K. Ghosh, P.K. Sen, CRC Press, "Handbook of Sequential Analysis" , 1991
[8] A. Wald, J. Wiley & sons, inc., 1947, "Sequential analysis"
[9] Killick, R., Fearnhead, P., Eckley, I.A., "Optimal detection of changepoints with a linear computational cost", Journal of the American Statistical Association 107(500) 1590-1598, 2012
[10] P. Granjon, "The CUSUM algorithm - a small review", https://hal.archives-ouvertes.fr/hal-00914697 , 2014

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis (Eds)

ON THE PSEUDO-INCIDENT WAVE TECHNIQUE FOR INTERACTING INHOMOGENEITIES IN ELECTROMECHANICAL PROBLEMS

XIAODONG WANG*, HUANGCHAO YU, S. ABDEL-GAWAD AND CHEN WANG

Department of Mechanical Engineering, University of Alberta
Edmonton, Alberta, Canada T6G 2G8
*e-mail: xiaodong.wang@ualberta.ca

Key words: Interaction, Elastic Wave, Electromechanical Problems, Inhomogeneities, Numerical Methods.

Abstract. The current paper presents the Pseudo-Incident Wave method for the theoretical treatment of the dynamic interaction between general inhomogeneities in advanced piezoelectric structures. Instead of simulating the response of such complicated systems using purely numerical or analytical methods, the current technique will take the advantages of the accuracy and reliability of analytical solutions and the flexibility of numerical methods. Using this method the original interaction problem is reduced to the solution of coupled single inhomogeneity problems, for which analytical solutions or simpler numerical solutions could be derived. By considering the consistency condition between different inhomogeneities, the steady state dynamic solution of multiple interaction problems can be formulated in terms of coupled single inhomogeneity solutions. The current method is very general and can provide reliable simulation of complicated interaction problems. Numerical examples are presented to illustrate the effectiveness of the Pseudo-Incident Wave method in simulating dynamic interaction between general inhomegeneities under complicated geometries.

1 INTRODUCTION

Piezoelectric materials are widely used in advanced structures to form self-monitoring and self-controlling smart systems and have drawn considerable attention from the research community. Mechanical deformation of such materials can be directly converted into electric signals for monitoring the mechanical deformation. In the reverse process, an applied electric field could induce deformation in the material. Therefore, the designers of such advanced structures will constantly face the challenge of properly modeling the electromechanical coupling between the electric and mechanical fields.

One of the most fundamental issues surrounding the optimization of the effectiveness and reliability of piezoelectric structures with multiple piezoelectric components is the evaluation of the effect of the interaction between different inhomogeneities, which will significantly affect both the local and global response of the coupled system. An accurate assessment of the coupled electromechanical behavior of an integrated piezoelectric composite system would,

therefore, require the determination of the local stress and electric fields involving interacting inhomogeneities.

Significant efforts had been made to the study of the quasistatic electromechanical behavior of piezoelectric composite materials. For example, the problems of a single elliptical (ellipsoidal) inhomogeneity in unbounded piezoelectric materials were solved using the Green's function approach [1, 2], the effective properties of piezoelectric composites was determined using different micromechanical models [3-6], and the fracture and damage of piezoelectric materials was also studied [7]. However, relatively few studies have focused on the interaction between inhomogeneities, especially the dynamic interacting behavior of inhomogeneities. It should be noted, however, that piezoelectric structures are currently being used or intended to use in situations where dynamic loading is involved, such as smart structures under impact loading and the acoustic control of smart skin systems. Even for static cases, it was observed that for composite material systems the mechanical properties are more sensitive to the local response of individual inhomogeneities [8, 9], which is closely related to the interaction among inhomogeneities.

The mechanical and electrical properties of advanced piezoelectric composite structures are greatly affected by the attached piezoelectric sensors/actuators, debondings, fibres and/or embedded cracks. The interaction between these inhomogeneities will cause the redistribution of the local stress and electric fields, which results in mechanical shielding or amplification effects and affects the overall failure mechanism, and alters the electrical behavior of the structures. Because of the complexity of the problem, when dynamic loads are applied, the simulation of the dynamic response of such coupled systems possesses a significant challenge. Typical numerical methods, such as finite element method or boundary element method, can be used to conduct dynamic simulation of these problems under certain conditions but have their own limitations when multiple interactions are involved, because of the computing resource needed to obtain reliable results. Analytical study of interacting inhomogeneities under dynamic loads is very attractive because of its high reliability and accuracy, but is limited to only simple cases of single inhomogeneity of certain types.

It is therefore the objective of the present paper to provide a comprehensive treatment of the steady-state dynamic behavior of interacting inhomogeneities in piezoelectric composites. The original problem is decomposed into single inhomogeneity subproblems. The solutions of these subproblems are then implemented into a pseudo-incident wave method to account for the interaction between different inhomogeneities. Numerical examples are provided to show the effect of the interaction between inhomogeneities, the material mismatch and the loading frequency upon the dynamical field.

2 PROBLEM FORMULATION

Consider an infinitely extended piezoelectric medium containing $\overline{M}$ arbitrarily located circular piezoelectric inhomogeneities of radius R_m ($m=1,2, \ldots, \overline{M}$), as shown in Fig.1. Inhomogeneity m and the matrix are bonded through a thin layer of thickness h_m. A global Cartesian (x, y) and local polar (r_m, θ_m) coordinate systems are used to characterize the inhomogeneities. The position of the center of inhomogeneity m is denoted (x_m, y_m) in the global coordinate system. The piezoelectric materials are assumed to be transversely isotropic

with the axis of symmetry being perpendicular to the $x-y$ plane. The shear moduli, the piezoelectric constants, the dielectric constants and the mass densities of the matrix and the inhomogeneities are denoted as $c_{44}^{M}, e_{15}^{M}, \kappa_{11}^{M}, \rho^{M}$ and $c_{44}^{F}, e_{15}^{F}, \kappa_{11}^{F}, \rho^{F}$, respectively. The corresponding material constants of the interphase of inhomogeneity m are c_{44}^{m}, e_{15}^{m} and κ_{11}^{m}.

When subjected to a steady state load of frequency ω, the resulting fields will generally involve an exponential harmonic factor $\exp(-i\omega t)$. For the sake of convenience, this factor will be suppressed and only the amplitude of different field variables will be considered.

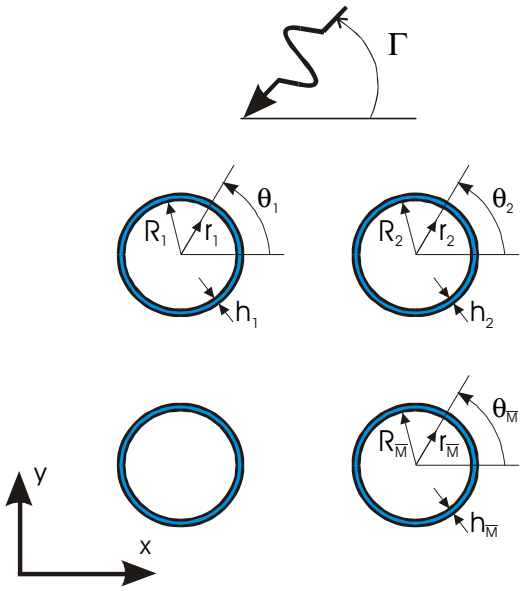


Figure 1: Interacting Piezoelectric Inhomogeneities

The steady state behavior of a homogeneous piezoelectric material under antiplane mechanical and inplane electric loading is fully described by the following governing equations:

$$\nabla^2 w + k^2 w = 0, \qquad \nabla^2 f = 0 \tag{1}$$

where the Laplacian operator ∇^2 stands for $\partial^2/\partial r^2 + \frac{1}{r}\partial/\partial r + \frac{1}{r^2}\partial^2/\partial\theta^2$, w is the antiplane displacement and k is the wave number defined by

$$k^2 = \frac{\rho\omega^2}{c_*} \quad \text{with} \quad c_* = c_{44} + \frac{e_{15}^2}{\kappa_{11}} \tag{2}$$

The electric potential ϕ is given by

$$\phi = \frac{e_{15}}{\kappa_{11}} w + f \tag{3}$$

and the corresponding non-vanishing stress and electric displacement components can be expressed as:

$$\tau_{rz} = c_* \frac{\partial w}{\partial r} + e_{15} \frac{\partial f}{\partial r} \quad \tau_{\theta z} = c_* \frac{1}{r} \frac{\partial w}{\partial \theta} + e_{15} \frac{1}{r} \frac{\partial f}{\partial \theta} \tag{4}$$

$$D_r = -\kappa_{11} \frac{\partial f}{\partial r}, \quad D_\theta = -\kappa_{11} \frac{1}{r} \frac{\partial f}{\partial \theta} \tag{5}$$

In these equations, $c_{44}, e_{15}, \kappa_{11}$ and ρ are the elastic modulus, the piezoelectric constants, the dielectric constant and the mass density, which should be replaced by the corresponding material constants of the matrix, the inhomogeneities and the interphases, respectively, when these media are considered.

3 INTERFACE MODEL

Interphases between matrix and fibers play a dominant role in characterizing the behavior of composites [10, 11]. In the current study, it is assumed that the thickness of interfacial layers is much smaller than the radius of fibers and the wavelength and that the inertial effect of the interfacial layers can be ignored. Accordingly, the radial shear stress τ_{rz}, strain γ_{rz}, electric displacement D_r and electric field intensity E_r are assumed to be uniform across the thickness of interphases, and the constitutive relation of interphase *m* can be expressed as

$$\begin{aligned} \tau_{rz}(\theta_m) = {} & \frac{c_*^m}{h_m}[w(R_m + h_m, \theta_m)\big|_{matrix} - w(R_m, \theta_m)\big|_{in\,\mathrm{hom}\,ogeneity}] \\ & + \frac{e_{15}^m}{h_m}[f(R_m + h_m, \theta_m)\big|_{matrix} - f(R_m, \theta_m)\big|_{in\,\mathrm{hom}\,ogeneity}] \end{aligned} \tag{6}$$

$$D_r(\theta_m) = -\frac{\kappa_{11}^m}{h_m}[f(R_m + h_m, \theta_m)\big|_{matrix} - f(R_m, \theta_m)\big|_{in\,\mathrm{hom}\,ogeneity}] \tag{7}$$

4 SINGLE INHOMOGENEITY PROBLEM

For a single inhomogeneity *m* subjected to an incident wave. The general solution of the displacement and electric fields can be expressed in terms of Fourier expansions in a local polar coordinate system (r_m, θ_m), as

$$w^m(r_m, \theta_m) = \begin{cases} \sum_{n=0}^{\infty} H_n^{(1)}(k_M r_m)[a_n^m e^{in\theta_m} + b_n^m e^{-in\theta_m}] + w^{inc} & in\ the\ matrix \\ \sum_{n=0}^{\infty} J_n(k_F r_m)[c_n^m e^{in\theta_m} + d_n^m e^{-in\theta_m}] & in\ the\ in\,\mathrm{hom}\,ogeneity \end{cases} \tag{8}$$

$$f^m(r_m,\theta_m)=\begin{cases}\sum_{n=0}^{\infty} r_m^{-n}[f_n^m e^{in\theta_m}+g_n^m e^{-in\theta_m}]+f^{inc} & \textit{in the matrix}\\ \sum_{n=0}^{\infty} r_m^{n}[p_n^m e^{in\theta_m}+q_n^m e^{-in\theta_m}] & \textit{in the inhomogeneity}\end{cases} \tag{9}$$

where $H_n^{(1)}$ and J_n are Hankel function and Bessel function of the first kind, respectively. w^{inc} and f^{inc} represent the incident wave. The solution of $a_n^m, b_n^m, c_n^m, d_n^m$ and f_n^m, g_n^m, p_n^m and q_n^m can be obtained by making use of the interphase model, eqns. (6) and (7).

Truncating the Fourier expansion in (8) and (9) into *Nth* term, the solution of the governing parameters $a_n^m, b_n^m, c_n^m, d_n^m$ and f_n^m, g_n^m, p_n^m and q_n^m can then be expressed in terms of the incident wave at the following integral points

$$\theta_m^l=\frac{2\pi(l-1)}{L-1},\quad l=1,2,...,L \tag{10}$$

as

$$\{C\}^m=[A]^m\{F\}^m \tag{11}$$

where $[A]^m$ is a known matrix, $\{C\}^m$ represents the coefficients of Fourier expansion of inhomogeneity m, given by

$$\{C\}^m=\{a_0^m, f_0^m, c_0^m, q_0^m, a_1^m, b_1^m, f_1^m, g_1^m, c_1^m, d_1^m, p_1^m, q_1^m, ..., c_N^m, d_N^m, p_N^m, q_N^m\}^T \tag{12}$$

and $\{F\}$represents the general loads acting on the integral points given by (10), i.e.

$$\{F\}^m=\{\tau_1, D_1, \frac{\kappa_{11}^m}{h^m}w_1, f_1, ..., \tau_L, D_L, \frac{\kappa_{11}^m}{h^m}w_L, f_L\}^T \tag{13}$$

5 INTERACTING INHOMOGENEITIES

For the cases where multiple inhomogeneities are involved, the interaction between these inhomogeneities may significantly affect the electromechanical behavior of the composites.

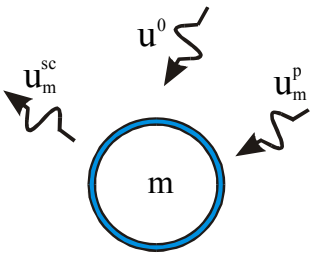


Figure 2: Pesudo-incident wave

5.1 Pseudo-incident wave method

Let us now focus our attention on a specific inhomogeneity m. The total incident wave for inhomogeneity m can be expressed as

$$u_m^I = u^0 + u_m^p \tag{14}$$

where u^0 is the initial incident wave and u_m^p is the unknown pseudo-incident wave from other inhomogeneities with u representing both the mechanical and the electric fields. As a result of this incident wave, inhomogeneity m will result in a scattering wave u_m^{sc}, as shown in Fig.2. The total field in the matrix can then be expressed as

$$u^{total} = u_m^I + u_m^{sc} = u^0 + u_m^p + u_m^{sc} \tag{15}$$

can also be obtained by summing up the initial field and the contributions from all inhomogeneities, such that

$$u^{total} = u^0 + \sum_{j=1}^{\overline{M}} u_j^{sc} \tag{16}$$

The equivalence between eqns. (15) and (16) indicates that

$$u_m^p = \sum_{j\neq m}^{\overline{M}} u_j^{sc}, \quad j = 1,2,...,\overline{M} \tag{17}$$

Equation (17) represents the relation between different inhomogeneities.

5.2 Solution of interacting inhomogeneities problems

According to eqns. (14) and (17), the general load $\{F\}^m$ used in eqn. (11) can be expressed as

$$\{F\}^m = \{F^0\}^m + \sum_{j\neq m}^{\overline{M}} \{F^{sc}\}_j^m \tag{18}$$

where $\{F^0\}^m$ is due to the original incident wave and $\{F^{sc}\}_j^m$ is the general load due to the scattered wave of inhomogeneity j. $\{F^{sc}\}_j^m$ can be obtained by using the general solution given by (8) and (9), and the constitutive relation (4) and (5) as

$$\{F^{sc}\}_j^m = [T]_j^m \{C\}^j \tag{19}$$

where $\{C\}^j$ represents the Fourier expansion coefficients of inhomogeneity j and $[T]_j^m$ is a known matrix obtained by rearranging (8) and (9). Substituting eqns. (18) and (19) into (11) results in

$$\{C\}^m - \sum_{j\neq m}^{\overline{M}} [A]^m [T]_j^m \{C\}^j = [A]^m \{F^0\}^m, \quad m = 1,2,...,\overline{M} \tag{20}$$

from which $\{C\}^m (m = 1,2,...,\overline{M})$ can be determined by solving a system of linear algebraic equations.

According to eqn. (16), the resulting mechanical and electric fields in the matrix given by eqns. (8) and (9) can be expressed in terms of the coefficients of Fourier expansions as

$$w(x,y) = w^0(x,y) + \sum_{m=1}^{\bar{M}} \sum_{n=0}^{N} H_n^{(1)}(k_M \bar{r}_m)[a_n^m e^{in\bar{\theta}_m} + b_n^m e^{-in\bar{\theta}_m}] \tag{21}$$

and

$$f(x,y) = f^0(x,y) + \sum_{m=1}^{\bar{M}} \sum_{n=0}^{N} \bar{r}_m^{-n}[f_n^m e^{in\bar{\theta}_m} + g_n^m e^{-in\bar{\theta}_m}] \tag{22}$$

where

$$\bar{r}_m = \sqrt{(x-x_m)^2 + (y-y_m)^2}, \ \bar{\theta}_m = \mathrm{sgn}(y-y_m)\cos^{-1}\frac{x-x_m}{\bar{r}_m} \tag{23}$$

From Eqns. (21) and (22), the stress and the electric displacement in the matrix can be obtained using the constitutive relation given by (4) and (5).

6 RESULTS AND DISCUSSIONS

The theoretical analysis described in previous section is used to investigate the coupled electromechanical response of piezoelectric composites to an incident harmonic wave, as shown in Fig. 1. The incident antiplane displacement and the electric potential can be expressed as

$$w^{(in)} = \frac{i\tau}{c_*^M k_M} e^{-ik_M(x\cos\Gamma + y\sin\Gamma)}, \ \phi^{(in)} = \frac{e_{15}^M}{\kappa_{11}^M} w^{(in)}$$

where τ is the maximum value of the shear stress carried by the incident wave, and Γ is the incident angle. In the following examples, uniformly distributed identical piezoelectric fibers will be examined. The present formulations predict the dependence of the electric and mechanical fields upon the geometry of the fibers, the material combination, the interfacial property, the frequency and angle of the incident wave. It should be recognized that only the amplitudes of the complex shear stress and electric potential are considered in the following figures.

6.1 Local Shear Stress

It is well known that local stress field in composites will be disturbed by the existence of fibers due to material mismatch [12]. The current study indicates that even when there no elastic mismatch exists, the electromechanical coupling between electric and mechanical fields in piezoelectric inhomoheneities may also result in significant change in the local stress level. Figure 3 shows the distribution of the scattering shear stress ($\tau^*=\tau_{rz}/\tau$) along the circumference of a single fiber embedded in an insulating medium subjected to an incident wave with $\Gamma = 90^0$. The shear wave speeds (and shear moduli) of the fiber and the matrix are assumed to be same. In this figure, the normalized piezoelectric constant $\lambda = \{e_{15}^F\}^2 \Big/ c_*^F \kappa_{11}^F$ is taken to be $\lambda = 1$ and $k^* = Rk_F$ represents the normalized frequency of the incident wave. Unlike non-piezoelectric materials for which no scattering will be generated for the current material combination, the piezoelectric fiber results in shear stresses over 0.25τ at $\theta = 30^0$ and 150^0. Significant stress distribution is also observed in Fig. 4, which shows the corresponding interfacial shear stress distribution along the surface of inhomogeneity one with

the presence of inhomogeneity two. The distance between them is $0.2R$ with R being the radius of the inhomogeneities.

6.2 Electric field

Electric potential induced in a piezoelectric composite could be used to monitor the deformation of the material and the property of the incident wave. Figure 5 shows the distribution of the electric potential ($\phi^* = e_{15}^F \phi / R\tau$) along the boundary of a single fiber examined in Fig. 3. It is very interesting to note that the peak value of ϕ^* has a unique relation with the incident angle ($\Gamma = 90^0$) for different loading frequencies. Figure 6 shows the corresponding results for two interacting inhomogeneities, similar to the case discussed in Fig. 4. In this case, the relation between the incident angle and the position of the peak value of ϕ^* shows a strong frequency-dependence.

6.3 Multiple interaction of inhomogeneities

Interaction between multiple inhomogeneities was studied using the current method for mechanical problems. Figure 7 shows the displacement field (real part) of two interacting inhomogeneities with $Rk^M = 7.5, \Gamma = 90^0$ and the distance between the inhomogeneities being *R*. The distribution of displacement (amplitude) around the boundary of the left inhomogeneity for different distances between the inhomogeneities is given in figure 8. The asymmetry of the results indicates the effect of the interaction between the inhomogeneities. Figure 9 shows the displacement field caused by four interacting inhomogeneities with $Rk^M = 1.0, \Gamma = 90^0$, the distance between adjacent inhomogeneities being *0.5R* and the ratio of the wave speed (matrx/inhomogeneity) being *0.707*.

Interaction between multiple inhomogeneities, as shown in figure 10, is also considered, where the distance between the adjacent inhomogeneities is *R*. Figures 11 and 12 show the displacement field generated by the multiple interaction for $Rk^M = 0.59, \Gamma = 90^0$ and $Rk^M = 1.37, \Gamma = 90^0$, respectively. The incident wave is blocked by the inhomogeneity array for both frequencies, which are in the bandgap of the inhomogeneity array. The successful treatment of the interaction between large numbers of inhomogeneities clearly shows the advantage of the current method.

ACKNOWLEDGEMENT

This research was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) .

REFERENCES

[1] Deeg, W. F. *The Analysis of dislocation, crack and inclusion problems in piezoelectric solids*, Ph. D. Dissertation, Stanford University. (1980).

[2] Benveniste, Y. The determination of the elastic and electric fields in a piezoelectric inhomogeneity. *J. Appl. Phys*. (1992) **72**: 1086-1095.

[3] Dunn, M. and Taya, M. Micromechanical predictions of the effective electro elastic moduli of piezoelectric composites. *Int. J. Solid Struct*. (1993) **30**: 161-175.
[4] Pak, Y. E. Crack extension force in a piezoelectric media with defects. *Int. J. Solids Struct*. (1990) **28**: 491-505.
[5] Suo, Z., Kuo, C. M., Barnett, D. M. and Wills, J. R. Fracture mechanics for piezoelectric ceramics. *J. Mech. Phys. Solids*. (1992) **40**: 739-765.
[6] Jain, A. K. and Sirkis, J. S. Continuum damage mechanics in piezoelectric ceramics. *Adaptive Structures and Composite Materials: Analysis and Application*. ASME (1994) AD **45**/MD **54**: 47-58.
[7] He, M. Y., Suo, Z., McMeeking, R. M., Evans, A. G. and Lynch, C.S. The mechanics of some degradation mechanisms in ferroelectric ceramic actuators, *SPIE*, (1994) **2189**: 344-356.
[8] Kachanov, M., Sevostianov, I. Non-interaction approximation in the problem of effective properties. In: Kachanov, M., Sevostianov, I. (Eds.), *Effective Properties of Heterogeneous Materials.* Springer, Dordrecht. (2014) 1–96.
[9] Rodin G. J. and Weng, G. J. On reflected interactions in elastic solids containing inhomogeneities. *Journal of the Mechanics and Physics of Solids* (2014) **68**: 197–209.
[10] Matikas, T. E. Ultrasonic reflectively technique for the characterization of fiber-matrix interface in metal matrix composites. *J. Appl. Phys*. (1993) **74**: 228-236.
[11] Veazie, D. R. and Qu, J. Effects of interphase on the transverse stress-strain behavior in unidirectional fiber reinforced metal matrix composites. *Composites Engineering*. (1995) **5**: 597-610.
[12] Wang, X. D. and Meguid, S. A. Diffraction of SH-wave by interacting matrix crack and an inhomogeneity. *ASME J. Appl. Mech*. (1997) **64**: 568-575.

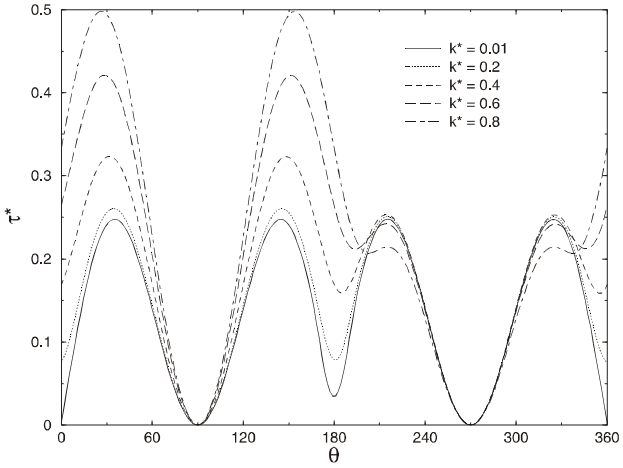


Figure 3: Shear stress of a single inhomogeneity

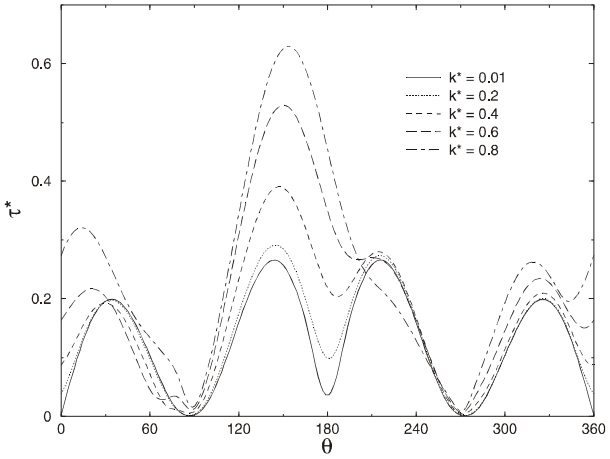


Figure 4: Shear stress of interacting inhomogeneities

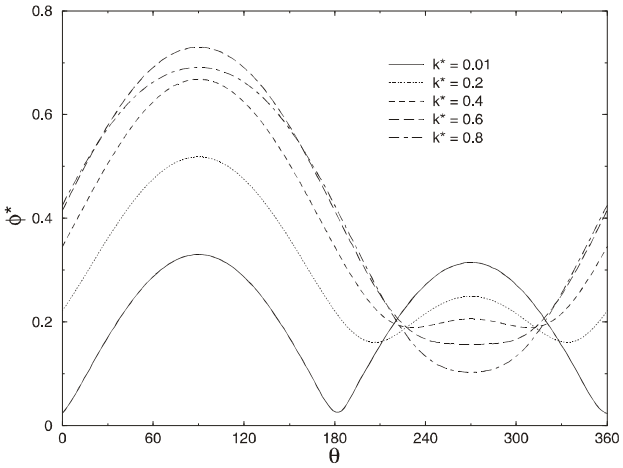


Figure 5: Potential of a single inhomogeneity

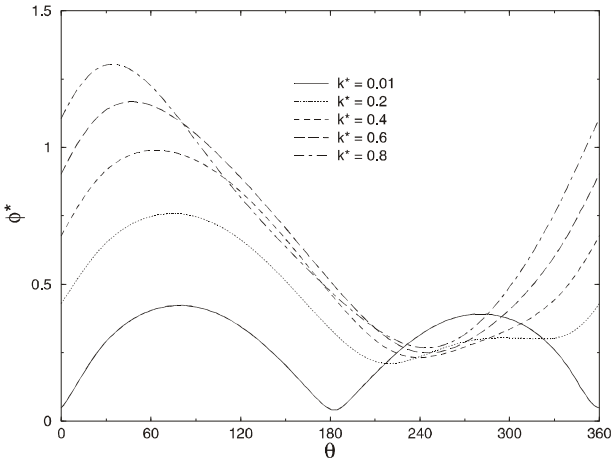


Figure 6: Potential of interacting inhomogeneities

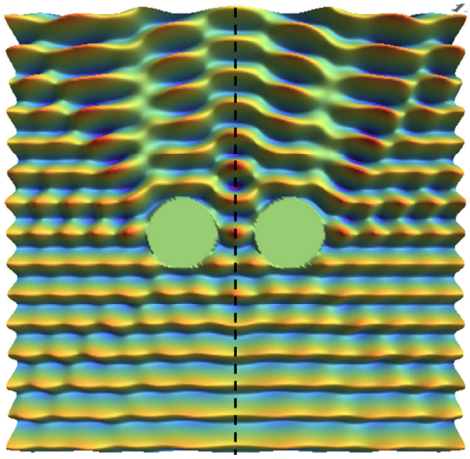


Figure 7: Wave field of two inhomogeneities

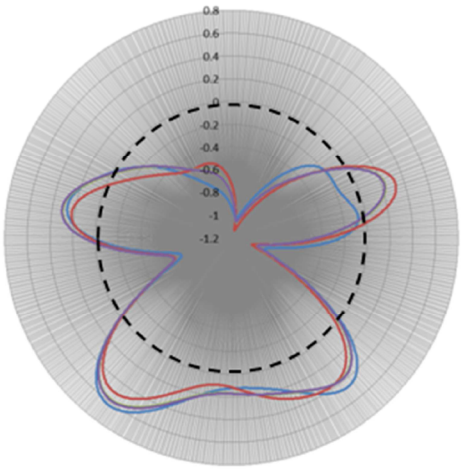


Figure 8: Displacement around the boundary

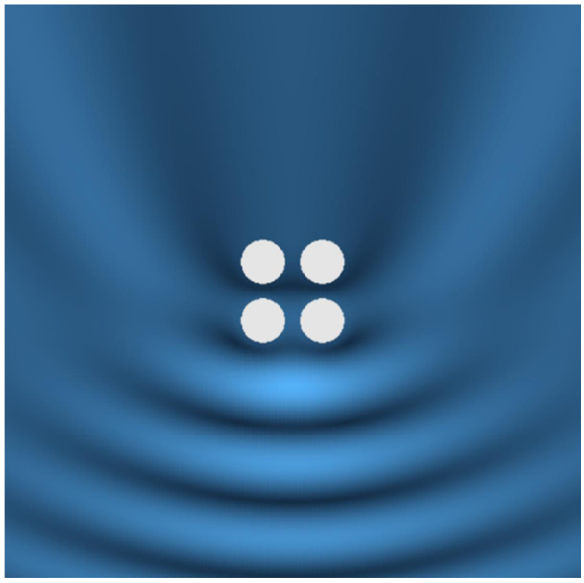


Figure 9: Displacement field due to four inhomogeneities

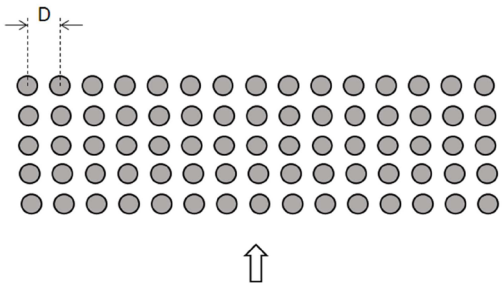


Figure 10: Array of multiple inhomogeneities

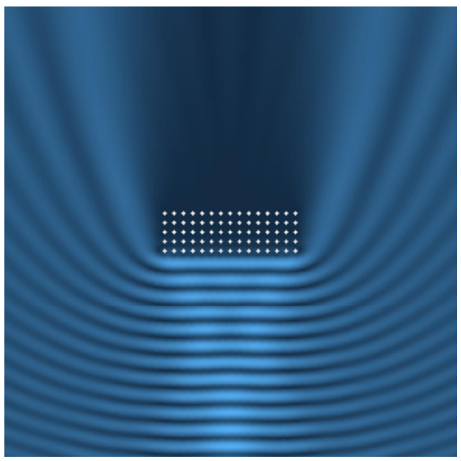


Figure 11: Displacement field for a normal incident wave, $k_M R=0.59$

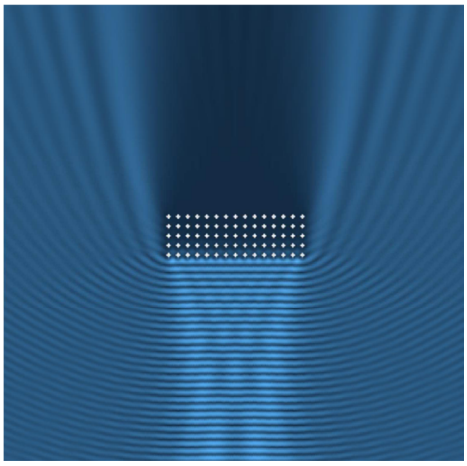


Figure 12: Displacement field for a normal incident wave, $k_M R=1.37$

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis(Eds)

CONVERGENCE SPEED OF COUPLING ITERATIONS FOR THE UNSTEADY TRANSMISSION PROBLEM

AZAHAR MONGE* AND PHILIPP BIRKEN*

*Centre for Mathematical Sciences
Lund University
Box 118, 22100, Lund, Sweden
e-mail: azahar.monge@na.lu.se, web page: http://www.maths.lu.se/staff/azahar-monge

Key words: Thermal Fluid Structure Interaction, Coupled Problems, Transmission Problem, Fixed Point Iteration, Dirichlet-Neumann Iteration

Abstract. We present an estimate for the convergence rate of the Dirichlet-Neumann iteration for the discretized unsteady transmission problem. Specifically, we consider the coupling of two heat equations on two identical squared domains. The Laplacian is discretized by second order central finite differences and the implicit Euler method is used for the time discretization. For the semidiscrete case, Henshaw and Chad provided in 2009 a method to analyse stability and convergence speed based on applying the continuous Fourier transform to the semi-discretized equations. Numerical results for the fully discrete case show differences, which is why we propose a complementary analysis based on approximating the spectral radius of the iteration matrix. Numerical results are presented to illustrate the analysis.

1 INTRODUCTION

Thermal fluid structure interaction occurs when a heat flux from a fluid leads to temperature changes in a structure or vice versa. Examples for this are cooling of gas-turbine blades, cooling of rocket thrust chambers [9], thermal anti-icing systems of airplanes [3], supersonic reentry of vehicles from space [8, 10] or gas quenching [6, 14].

Unsteady thermal fluid structure interaction is modelled using two partial differential equations describing a fluid and a structure on different domains. The equations are coupled at an interface to model the heat transfer between fluid and structure. For the solution of a coupled problem, two general approaches can be distinguished. In a partitioned approach [4], different codes for the sub-problems are used and the coupling is done by a master program which calls interface functions of the other codes. This allows to use existing software for each sub-problem, in contrast to a monolithic approach, where a new code is tailored for the coupled equations. The standard partitioned algorithm is

the Dirichlet-Neumann iteration, where the PDEs are solved separately using Dirichlet, respectively Neumann boundary conditions with data given from the solution of the other problem. This iteration has been analyzed and a convergence condition is given by [12], but convergence rates have not been computed.

Henshaw and Chad provided in [7] a method to analyse stability and convergence speed of the Dirichlet-Neumann iteration for the thermal transmission problem based on applying the continuous Fourier transform to the semi-discretized equations. Their result depends on the thermal conductivities and diffusivities of the materials. However, in the fully-discrete case we observe that the iteration converges much faster [2]. Therefore, we propose a complementary stability and convergence study for this case.

In this paper we consider the transmission problem because it is a basic building block in fluid structure interaction. In particular, we consider the coupling of two heat equations on two identical squared domains. The Laplacian is discretized by second order central finite differences and the implicit Euler method is used for the time discretization.

To study the convergence behaviour of this problem, we reformulate the fully-discrete iteration as a system of algebraic equations. For each domain, we have a matrix describing the discretization of the Laplacian and a matrix describing the coupling conditions. This leads us to a linear coupled system of equations with sparse block tridiagonal matrices. Using the Schur complement [15], the exact iteration matrix can be written down. However, the spectral radius of that is too difficult to compute. We therefore present an estimate based on approximating the iteration matrix by its block diagonal because the iteration matrix is a strictly diagonally dominant matrix.

An outline of the paper now follows. In Section 2, we define the problem to be solved in terms of the partial differential equations, boundary conditions and interface conditions. We also give a description of the discretization. In Section 3, we explain how to solve the model problem using a fixed point iteration. Our stability analysis for the fully-discretized case of the model problem using Dirichlet-Neumann interface conditions is presented in Section 4. In Section 5, we present numerical results that show the theoretical stability analysis. Conclusions are given in the final section.

2 MODEL PROBLEM AND DISCRETIZATION

The unsteady transmission problem is as follows, where we consider a domain Ω which is cut into two subdomains $\Omega = \Omega_1 \cup \Omega_2$ with transmission conditions at the interface $\Gamma = \Omega_1 \cap \Omega_2$:

$$\begin{aligned}
\frac{\partial u_m(\mathbf{x},t)}{\partial t} + D_m \Delta u_m(\mathbf{x},t) = f(\mathbf{x}), \quad t \in [t_0, t_f] \quad \mathbf{x} = (x,y) \in \Omega_m \subset \mathbb{R}^2, \ m = 1,2 \\
u_m(\mathbf{x},t) = 0, \quad t \in [t_0, t_f], \quad \mathbf{x} \in \partial\Omega_m \backslash \Gamma \\
u_1(\mathbf{x},t) = u_2(\mathbf{x},t), \quad \mathbf{x} \in \Gamma \\
K_1 \partial_x u_1(\mathbf{x},t) \cdot \mathbf{n} = K_2 \partial_x u_2(\mathbf{x},t) \cdot \mathbf{n}, \quad \mathbf{x} \in \Gamma \\
u_m(\mathbf{x},0) = g_m(\mathbf{x}), \quad \mathbf{x} \in \Omega_m
\end{aligned} \tag{1}$$

The constants K_1 and K_2 describe the thermal conductivities of the materials on Ω_1 and Ω_2 respectively. Analogously, D_1 and D_2 represent the thermal diffusivities of the materials.

For this study, we use $\Omega_1 = [0,1] \times [0,1]$, $\Omega_2 = [1,2] \times [0,1]$ and

$$\begin{aligned} f(x,y) = \sin \pi y^2 (\pi \cos \frac{\pi}{2} x^2 - \pi^2 x^2 \sin \frac{\pi}{2} x^2) \\ + \sin \frac{\pi}{2} x^2 (2\pi \cos \pi y^2 - 4\pi^2 y^2 \sin \pi y^2). \end{aligned} \tag{2}$$

This was chosen such that the exact solution is

$$u(x,y) = \sin \pi y^2 \sin \frac{\pi}{2} x^2, \tag{3}$$

which satisfies the boundary conditions.

We discretize this problem using second order central differences for the Laplacian with a constant mesh width of $\Delta x = \Delta y$. For the time discretization we use the implicit Euler method. All linear systems are solved using CG.

2.1 Space Discretization

Let's first discretize the equation $D_1 \Delta u_1(x,y) = f(x,y)$ on Ω_1 with Dirichlet boundary conditions at the interface Γ. Consider $u_1(x_i, y_j) \approx u_{1,i,j}$ for $i = 0, ..N_x$, $j = 0, .., N_x + 1$ and $\Delta x = \Delta y = 1/(N_x + 1)$. We use the following second order central difference to approximate the second order spatial derivatives,

$$u_{1,xx}(x_i, y_j) \approx \frac{1}{\Delta x^2}(u_{1,i+1,j} - 2u_{1,i,j} + u_{1,i-1,j}), \tag{4}$$

$$u_{1,yy}(x_i, y_j) \approx \frac{1}{\Delta x^2}(u_{1,i,j+1} - 2u_{1,i,j} + u_{1,i,j-1}). \tag{5}$$

These approximations lead us to the following linear system,

$$\mathbf{A}_1 \mathbf{u}_1 + \mathbf{f}_1 = 0, \tag{6}$$

where $\mathbf{A}_1 \in \mathbb{R}^{N_x \cdot N_x \times N_x \cdot N_x}$ and $\mathbf{u}_1$, $\mathbf{f}_1 \in \mathbb{R}^{N_x \cdot N_x \times 1}$. $\mathbf{A}_1$ is given by

$$\mathbf{A}_1 := -\frac{D_1}{\Delta x^2} \begin{pmatrix} \tilde{\mathbf{A}} & \mathbf{I} & & \mathbf{0} \\ \mathbf{I} & \tilde{\mathbf{A}} & \ddots & \\ & \ddots & \ddots & \mathbf{I} \\ \mathbf{0} & & \mathbf{I} & \tilde{\mathbf{A}} \end{pmatrix} \quad \text{where} \quad \tilde{\mathbf{A}} := \begin{pmatrix} -4 & 1 & & 0 \\ 1 & -4 & \ddots & \\ & \ddots & \ddots & 1 \\ 0 & & 1 & -4 \end{pmatrix}_{N_x \times N_x} \tag{7}$$

and $\mathbf{I}$ is the identity matrix. Note that each block of the matrix $\mathbf{A}_1$ has size $N_x \times N_x$.

Finally, on the right boundary we have to add a matrix $\mathbf{P}$ describing the Dirichlet coupling conditions at the interface. The full discretization is described as follows,

$$\mathbf{A}_1 \mathbf{u}_1 + \mathbf{P} \mathbf{u}_2 + \mathbf{f}_1 = 0,$$

where $\mathbf{P} \in \mathbb{R}^{N_x \cdot N_x \times (N_x+1)\cdot N_x}$ and is defined by

$$\mathbf{P} = \frac{D_1}{\Delta x^2} \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{I} & \mathbf{0} \end{pmatrix}. \tag{8}$$

The bottom left block of the matrix $\mathbf{P}$ has size $N_x \times N_x$.

On the other hand, we also need to discretize the equation $D_2 \Delta u_2(x,y) = f(x,y)$ on Ω_2 with Neumann boundary conditions at the interface Γ and Dirichlet boundary conditions on $\partial\Omega_2 \backslash \Gamma$. The discretization for the interior points of the grid is the same as for Ω_1. Therefore, we only need to develop the discretization of this equation on the left boundary.

We discretize the Neumann boundary conditions at the interface using forward finite differences:

$$K_1 \partial u_1(x_{N_x+1}, y_j) \cdot \mathbf{n} \approx \frac{K_1(u_{1,N_x,j} - u_{1,N_x+1,j})}{\Delta x} \quad \text{and} \tag{9}$$

$$K_2 \partial u_2(x_0, y_j) \cdot \mathbf{n} \approx \frac{K_2(u_{2,0,j} - u_{2,1,j})}{\Delta x} \quad \text{for} \quad j = 0, .., N_x + 1. \tag{10}$$

In general, the entire discretization looks as follows,

$$\mathbf{A}_2 \mathbf{u}_2 + \mathbf{D}\mathbf{u}_1 + \mathbf{f}_2 = 0, \tag{11}$$

where $\mathbf{A}_2 \in \mathbb{R}^{(N_x+1)\cdot N_x \times (N_x+1)\cdot N_x}$, $\mathbf{D} \in \mathbb{R}^{(N_x+1)\cdot N_x \times N_x \cdot N_x}$ is the discrete normal derivative at the interface and $\mathbf{u}_1$, $\mathbf{u}_2$, $\mathbf{f}_2 \in \mathbb{R}^{(N_x+1)\cdot N_x \times 1}$. $\mathbf{A}_2$ is given by

$$\mathbf{A}_2 := -\frac{D_2}{\Delta x^2} \begin{pmatrix} \hat{\mathbf{A}} & \mathbf{I} & & \mathbf{0} \\ \mathbf{I} & \tilde{\mathbf{A}} & \ddots & \\ & \ddots & \ddots & \mathbf{I} \\ \mathbf{0} & & \mathbf{I} & \tilde{\mathbf{A}} \end{pmatrix} \quad \text{where} \quad \hat{\mathbf{A}} := \begin{pmatrix} -3 & 1 & & 0 \\ 1 & -3 & \ddots & \\ & \ddots & \ddots & 1 \\ 0 & & 1 & -3 \end{pmatrix}_{N_x \times N_x}. \tag{12}$$

Note that each block of the matrix $\mathbf{A}_2$ has size $N_x \times N_x$. $\mathbf{D}$ is described as follows,

$$\mathbf{D} = \frac{D_2}{\Delta x^2} \begin{pmatrix} \mathbf{0} & \tilde{\mathbf{D}} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \quad \text{where} \quad \tilde{\mathbf{D}} = \frac{K_1}{K_2} \begin{pmatrix} \mathbf{I} & -\mathbf{I} \end{pmatrix}_{N_x \times 2N_x}. \tag{13}$$

2.2 Time Discretization

As we said before, we use the implicit Euler method for the time discretization of (1). In order to simplify the analysis, we assume that $\mathbf{f}_1 = \mathbf{f}_2 \equiv \mathbf{0}$ without loss of generality. Therefore, one can rewrite the time dependent transmission problem as the coupled equations

$$\begin{aligned} \dot{\mathbf{u}}_1 &= \mathbf{A}_1 \mathbf{u}_1 + \mathbf{P}\mathbf{u}_2, \\ \dot{\mathbf{u}}_2 &= \mathbf{A}_2 \mathbf{u}_2 + \mathbf{D}\mathbf{u}_1, \end{aligned} \tag{14}$$

where $\mathbf{A}_1$ corresponds to the space discretization on Ω_1, $\mathbf{A}_2$ to the space discretization on Ω_2, $\mathbf{P}$ to the Dirichlet boundary conditions mapping from Ω_2 to Ω_1 and $\mathbf{D}$ to the Neumann boundary conditions mapping from Ω_1 to Ω_2.

Applying the implicit Euler method to the system (14), we get

$$\mathbf{u}_1^{n+1} = \mathbf{u}_1^n - \Delta t(\mathbf{A}_1\mathbf{u}_1^{n+1} + \mathbf{P}\mathbf{u}_2^{n+1}), \tag{15}$$
$$\mathbf{u}_2^{n+1} = \mathbf{u}_2^n - \Delta t(\mathbf{A}_2\mathbf{u}_2^{n+1} + \mathbf{D}\mathbf{u}_1^{n+1}), \tag{16}$$

with the time step size Δt fixed.

3 FIXED POINT ITERATION

We now employ a standard Dirichlet-Neumann iteration to solve the discrete system (15)-(16). This corresponds to alternately solving the problems (15) and (16), where problem (15) corresponds to a discretization of the transmission problem (1) on Ω_1 only with Dirichlet data on Γ given by $\mathbf{u}_2$ on the coupling interface and problem (16) corresponds to a discretization of (1) on Ω_2 only with Neumann data on Γ given by the discrete normal derivative of $\mathbf{u}_1$ on Γ. This iteration is also known as Gauss-Seidel process [12].

Applying this to (15)-(16), one gets for the k-th iteration

$$\mathbf{u}_1^{n+1,k+1} = \mathbf{u}_1^n - \Delta t(\mathbf{A}_1\mathbf{u}_1^{n+1,k+1} + \mathbf{P}\mathbf{u}_2^{n+1,k}), \tag{17}$$
$$\mathbf{u}_2^{n+1,k+1} = \mathbf{u}_2^n - \Delta t(\mathbf{A}_2\mathbf{u}_2^{n+1,k+1} + \mathbf{D}\mathbf{u}_1^{n+1,k+1}), \tag{18}$$

with some initial condition, here $\mathbf{u}_2^{n+1,0} = \mathbf{u}_2^n$.

The iteration is terminated according to the standard criterion $\|\mathbf{u}^{k+1} - \mathbf{u}^k\| \leq \tau$ [1].

4 ANALYSIS

In this section we present the convergence analysis for the semi discrete and the fully discrete case of the model problem.

4.1 Semi Discrete Case

Here, one applies the implicit Euler method for the time discretization on both equations in (1) but keeps the space continuous. Then, Henshaw and Chand applied in [7] the Fourier transform in space in order to transform the second order derivatives into algebraic expressions. This converts the partial differential equations into a system of purely algebraic equations. Once we have a coupled system of algebraic equations, we can insert one into the other one and obtain the amplification factor β which is approximated by

$$\beta \approx \frac{K_1}{K_2}\sqrt{\frac{D_2}{D_1}}. \tag{19}$$

4.2 Discrete Case

To analyze the iteration (17)-(18), we isolate the term $\mathbf{u}_i^{n+1,k+1}$ for $i = 1, 2$ in both equations:

$$\mathbf{u}_1^{n+1,k+1} = (\mathbf{I} + \Delta t \mathbf{A}_1)^{-1}(\mathbf{u}_1^n - \Delta t \mathbf{P}\mathbf{u}_2^{n+1,k}), \tag{20}$$
$$\mathbf{u}_2^{n+1,k+1} = (\mathbf{I} + \Delta t \mathbf{A}_2)^{-1}(\mathbf{u}_2^n - \Delta t \mathbf{D}\mathbf{u}_1^{n+1,k+1}). \tag{21}$$

Now we insert (20) into (21)

$$\begin{aligned}\mathbf{u}_2^{n+1,k+1} &= (\mathbf{I} + \Delta t \mathbf{A}_2)^{-1}(\mathbf{u}_2^n - \Delta t \mathbf{D}(\mathbf{I} + \Delta t \mathbf{A}_1)^{-1}(\mathbf{u}_1^n - \Delta t \mathbf{P}\mathbf{u}_2^{n+1,k})) \\ &= (\mathbf{I} + \Delta t \mathbf{A}_2)^{-1}\mathbf{u}_2^n - (\mathbf{I} + \Delta t \mathbf{A}_2)^{-1}\Delta t \mathbf{D}(\mathbf{I} + \Delta t \mathbf{A}_1)^{-1}\mathbf{u}_1^n \\ &\quad + (\mathbf{I} + \Delta t \mathbf{A}_2)^{-1}\Delta t \mathbf{D}(\mathbf{I} + \Delta t \mathbf{A}_1)^{-1}\Delta t \mathbf{P}\mathbf{u}_2^{n+1,k}.\end{aligned} \tag{22}$$

To prove convergence, one needs to show that the norm of the iteration matrix $\mathbf{M} = (\mathbf{I}+\Delta t\mathbf{A}_2)^{-1}\Delta t\mathbf{D}(\mathbf{I}+\Delta t\mathbf{A}_1)^{-1}\Delta t\mathbf{P}$ is smaller than 1 in some norm. Here we choose $\|\cdot\|_2$.

The first natural idea when facing the problem of computing $\|\mathbf{M}\|_2$ is to use the following basic norm property:

$$\|\mathbf{M}\|_2 \leq \Delta t^2 \|(\mathbf{I} + \Delta t \mathbf{A}_2)^{-1}\|_2 \|\mathbf{D}\|_2 \|(\mathbf{I} + \Delta t \mathbf{A}_1)^{-1}\|_2 \|\mathbf{P}\|_2. \tag{23}$$

However, the inequality in (23) is too coarse and the right hand side of the previous expression is way bigger than 1. Any other combination of breaking $\|\mathbf{M}\|_2$ lead us to the same situation. In conclusion, in order to compute $\|\mathbf{M}\|_2$ we need to compute the iteration matrix $\mathbf{M}$ and apply the norm afterwards.

The iteration matrix $\mathbf{M}$ is not easy to compute for different reasons. First of all, the matrices $\mathbf{I} + \Delta t \mathbf{A}_1$ and $\mathbf{I} + \Delta t \mathbf{A}_2$ are sparse block tridiagonal matrices, and consequently, their inverses are not a straight forward computation. A block-by-block algorithm for inverting a block tridiagonal matrix is explained in [13]. However, the algorithm is based on the iterative application of the Schur complement [15], and it results in a sequence of block matrices and inverses of block matrices that it is impossible to compute exactly. Moreover, the diagonal block matrices of $\mathbf{I} + \Delta t \mathbf{A}_1$ and $\mathbf{I} + \Delta t \mathbf{A}_2$ are tridiagonal matrices but their inverses are full matrices [5] which does not help for computing $\mathbf{M}$.

Due to these difficulties, we propose here to approximate $\mathbf{M}$. One can observe that $\mathbf{I}+\Delta t\mathbf{A}_1$ and $\mathbf{I}+\Delta t\mathbf{A}_2$ are strictly diagonally dominant matrices, and therefore, we propose to approximate them by the their block diagonal. Explicitly, if we define $\tilde{\mathbf{A}}_1 := \Delta x^2 \mathbf{A}_1$, we approximate

$$\mathbf{I} + \Delta t \mathbf{A}_1 = \mathbf{I} + \frac{D_1 \Delta t}{\Delta x^2}\tilde{\mathbf{A}}_1 \approx \text{blockdiag}\left(\mathbf{I} + \frac{D_1 \Delta t}{\Delta x^2}\tilde{\mathbf{A}}_1\right) = \begin{pmatrix} \mathbf{A} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \mathbf{A} \end{pmatrix}, \tag{24}$$

where $\mathbf{A}$ is defined by (compare with (7))

$$\mathbf{A} := \begin{pmatrix} 1+4\frac{D_1\Delta t}{\Delta x^2} & -\frac{D_1\Delta t}{\Delta x^2} & & 0 \\ -\frac{D_1\Delta t}{\Delta x^2} & 1+4\frac{D_1\Delta t}{\Delta x^2} & \ddots & \\ & \ddots & \ddots & -\frac{D_1\Delta t}{\Delta x^2} \\ 0 & & -\frac{D_1\Delta t}{\Delta x^2} & 1+4\frac{D_1\Delta t}{\Delta x^2} \end{pmatrix}_{N_x\times N_x}.$$

Analogously, if we define $\tilde{\mathbf{A}}_2 := \Delta x^2 \mathbf{A}_2$, we approximate

$$\mathbf{I} + \Delta t \mathbf{A}_2 = \mathbf{I} + \frac{D_2\Delta t}{\Delta x^2}\tilde{\mathbf{A}}_2 \approx \text{blockdiag}\left(\mathbf{I} + \frac{D_2\Delta t}{\Delta x^2}\tilde{\mathbf{A}}_2\right) = \begin{pmatrix} \mathbf{B} & & & \mathbf{0} \\ & \mathbf{A} & & \\ & & \ddots & \\ \mathbf{0} & & & \mathbf{A} \end{pmatrix}, \quad (25)$$

where $\mathbf{B}$ is defined by (compare with (12))

$$\mathbf{B} := \begin{pmatrix} 1+3\frac{D_2\Delta t}{\Delta x^2} & -\frac{D_2\Delta t}{\Delta x^2} & & 0 \\ -\frac{D_2\Delta t}{\Delta x^2} & 1+3\frac{D_2\Delta t}{\Delta x^2} & \ddots & \\ & \ddots & \ddots & -\frac{D_2\Delta t}{\Delta x^2} \\ 0 & & -\frac{D_2\Delta t}{\Delta x^2} & 1+3\frac{D_2\Delta t}{\Delta x^2} \end{pmatrix}_{N_x\times N_x}.$$

Observe that $\mathbf{A}$ and $\mathbf{B}$ are nonsingular and thus [11]

$$\|\mathbf{A}^{-1}\|_2 = \frac{1}{\sigma_{\min}(\mathbf{A})}, \quad \|\mathbf{B}^{-1}\|_2 = \frac{1}{\sigma_{\min}(\mathbf{B})}, \quad (26)$$

where $\sigma_{\min}$ is the smallest singular value. Note also that the eigenvalues of $\tilde{\mathbf{A}}$ are $\lambda_i(\tilde{\mathbf{A}}) = -4+2\cos\left(\frac{i\pi}{N_x+1}\right)$ for $i = 1,..,N_x$ and the eigenvalues of $\hat{\mathbf{A}}$ are $\lambda_i(\hat{\mathbf{A}}) = -3+2\cos\left(\frac{i\pi}{N_x+1}\right)$ for $i = 1,..,N_x$. Furthermore, as $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{I}$ are Hermitian matrices, the following inequalities hold [15],

$$\sigma_{\min}(\mathbf{I}+\mathbf{A}) \geq \sigma_{\min}(\mathbf{I}) + \sigma_{\min}(\mathbf{A}), \quad \sigma_{\min}(\mathbf{I}+\mathbf{B}) \geq \sigma_{\min}(\mathbf{I}) + \sigma_{\min}(\mathbf{B}). \quad (27)$$

Now we can compute the product of the four matrices that conform the matrix iteration $\mathbf{M}$ using the approximations (24) and (25) for the matrices $\mathbf{I} + \Delta t\mathbf{A}_1$ and $\mathbf{I} + \Delta t\mathbf{A}_2$ respectively. Thus,

$$\mathbf{M} \approx \Delta t^2 \begin{pmatrix} \mathbf{B} & & \mathbf{0} \\ & \mathbf{A} & \\ & & \ddots \\ \mathbf{0} & & \mathbf{A} \end{pmatrix}^{-1} \mathbf{D} \begin{pmatrix} \mathbf{A} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \mathbf{A} \end{pmatrix}^{-1} \mathbf{P}$$

$$= \frac{D_1 D_2 K_1 \Delta t^2}{K_2 \Delta x^4} \begin{pmatrix} \mathbf{B}^{-1} & & \mathbf{0} \\ & \mathbf{A}^{-1} & \\ & & \ddots \\ \mathbf{0} & & \mathbf{A}^{-1} \end{pmatrix} \begin{pmatrix} \mathbf{0} & \mathbf{I} & -\mathbf{I} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{A}^{-1} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \mathbf{A}^{-1} \end{pmatrix} \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{I} & \mathbf{0} \end{pmatrix}$$

$$= \frac{D_1 D_2 K_1 \Delta t^2}{K_2 \Delta x^4} \begin{pmatrix} \mathbf{0} & \mathbf{B}^{-1} & -\mathbf{B}^{-1} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{A}^{-1} & \mathbf{0} \end{pmatrix}$$

$$= \frac{D_1 D_2 K_1 \Delta t^2}{K_2 \Delta x^4} \begin{pmatrix} -\mathbf{B}^{-1}\mathbf{A}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}.$$

So, we have that

$$\|\mathbf{M}\|_2 \approx \left\| \frac{D_1 D_2 K_1 \Delta t^2}{K_2 \Delta x^4} \begin{pmatrix} -\mathbf{B}^{-1}\mathbf{A}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \right\|_2 = \frac{D_1 D_2 K_1 \Delta t^2}{K_2 \Delta x^4} \|\mathbf{B}^{-1}\mathbf{A}^{-1}\|_2$$

$$\leq \frac{D_1 D_2 K_1 \Delta t^2}{K_2 \Delta x^4} \|\mathbf{B}^{-1}\|_2 \|\mathbf{A}^{-1}\|_2 = \frac{D_1 D_2 K_1 \Delta t^2}{K_2 \Delta x^4} \frac{1}{\sigma_{\min}(\mathbf{A})} \frac{1}{\sigma_{\min}(\mathbf{B})}$$

$$\leq \frac{D_1 D_2 K_1 \Delta t^2}{K_2 \Delta x^4} \frac{1}{1 + \frac{D1\Delta t}{\Delta x^2} |\lambda_i(\tilde{\mathbf{A}})|} \frac{1}{1 + \frac{D2\Delta t}{\Delta x^2} |\lambda_i(\hat{\mathbf{A}})|} \leq \frac{D_1 D_2 K_1 \Delta t^2}{K_2 (\Delta x^2 + D_2 \Delta t)(\Delta x^2 + 2 D_1 \Delta t)},$$

where the last equality holds due to (26) and the previous to last inequality holds by (27).

Therefore,

$$\epsilon := \frac{D_1 D_2 K_1 \Delta t^2}{K_2 (\Delta x^2 + D_2 \Delta t)(\Delta x^2 + 2 D_1 \Delta t)} \longrightarrow \frac{K_1}{2 K_2} \quad \text{when} \quad \Delta x \to 0 \tag{28}$$

is the estimate of $\|\mathbf{M}\|_2$ using the approximations explained above.

5 NUMERICAL RESULTS

First of all we present a numerical solution of the model explained above. Figure 1 shows the initial condition and the discrete solution for $\Delta x = 1/30$, $\Delta t = t_f/10$, $t_f = 0.01$ and $\tau = 1e-4$. One can observe how the absolut maximum of the initial condition decreases with time and the absolut minimum increases. This is due to the heat interaction between both subdomains.

Now we want to compare the estimate ϵ with the convergent rates for different examples varying the thermal diffusivities and conductivities D_1, D_2, K_1 and K_2. Input data for the different cases can be checked in table 1.

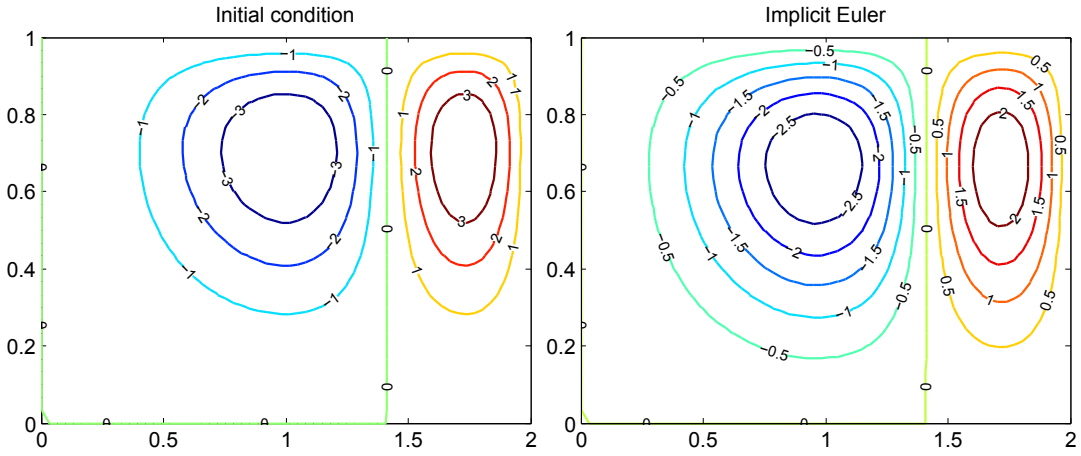


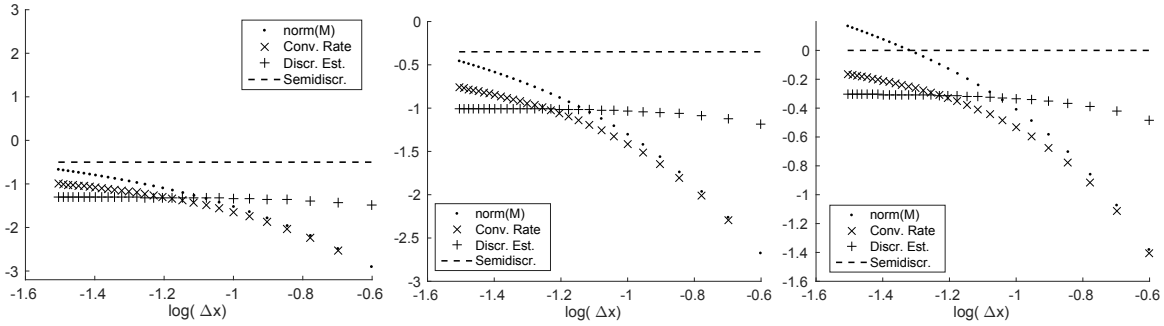
Figure 1: Contour lines of initial condition and discrete solution for $\Delta x = 1/30$, $\Delta t = t_f/10$, $t_f = 0.01$ and $\tau = 1e-4$.

Table 1: Data for different test cases. β is the semi discrete estimate in equation (19), and the last column represents the asymptotics of the discrete estimate ϵ in (28).

Case	D_1	D_2	K_1	K_2	β	$K_1/2K_2$
0	0.1	1	0.1	1	0.32	0.05
1	0.1	0.5	0.2	1	0.45	0.1
2	1	1	1	1	1	0.5
3	0.1	1000	0.1	1	10	0.05
4	1000	0.1	0.1	1	0.001	0.05
5	0.1	1000	1	0.1	1000	5

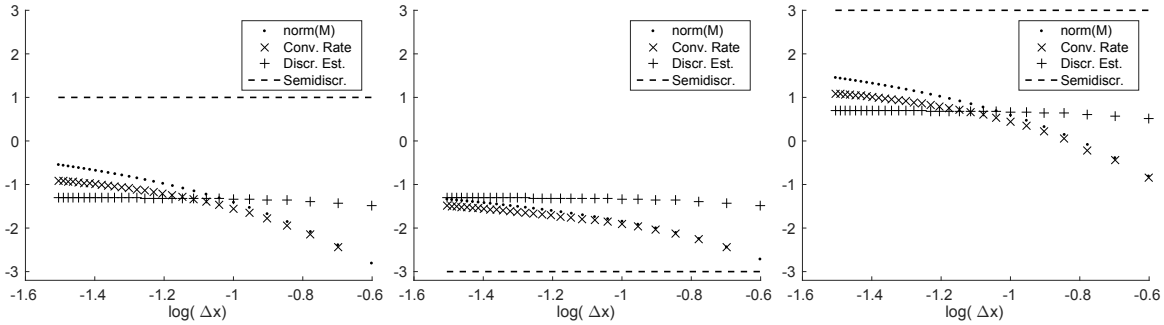
Figure 2 shows the cases 0, 1 and 2 specified in table 1. We observe that the semidiscrete estimate β is an upper bound and the discrete estimate ϵ intersects at certain point with the convergence rates curve. However, the asymptotic behavior of both estimates is fairly close to the asymptotic behavior of the convergence rates.

Figures 3a and 3b shows the comparison between cases 3 and 4 in table 1. Here, the thermal conductivities K_1 and K_2 are the same in both plots but the thermal diffusivities are switched (meaning that D_1 in case 3 corresponds to D_2 in case 4 and D_2 in case 3 corresponds to D_1 in case 4). We can observe that the asymptotics of the convergence rates do not vary that much in both plots. This pattern is been observed in many numerical experiments when we kept the same values for K_1 and K_2 and we varied D_1 and D_2. This result leads us to the conclusion that the behavior of the convergence rates do not have a strong dependence on the thermal diffusivities D_1 and D_2. This observation matches with the asymptotic behavior of the discrete estimate ϵ (see equation (28)). It is important to notice that the constants D_1 and D_2 have an effect in the convergence rates, but this



(a) Case 0. (b) Case 1. (c) Case 2.

Figure 2: Cases 0, 1 and 2 from table 1. The dotted line corresponds to $\|\mathbf{M}\|_2$, the dashed line corresponds to the semidiscrete estimate β, the crosses correspond to the discrete estimate ϵ and the remaining line corresponds to the convergence rates. The curves are restricted to the discrete values $\Delta x = 1/30, 1/29, .., 1/2$ and $\Delta t = 0.2$. All the curves are plotted in logarithm scale.



(a) Case 3. (b) Case 4. (c) Case 5.

Figure 3: Cases 3, 4 and 5 from table 1. The dotted line corresponds to $\|\mathbf{M}\|_2$, the dashed line corresponds to the semidiscrete estimate β, the crosses correspond to the discrete estimate ϵ and the remaining line corresponds to the convergence rates. The curves are restricted to the discrete values $\Delta x = 1/30, 1/29, .., 1/2$ and $\Delta t = 0.2$. All the curves are plotted in logarithm scale.

effect vanishes when Δx tends to zero.

The comparison between figures 3a and 3b is also interesting because we can observe a clear difference between the semidiscrete and the discrete cases. In particular, in figure 3a the semidiscrete estimate predicts divergence, but we obtain convergence in the discrete case for all Δx.

Now we see when one keeps the same values for D_1 and D_2 and varies K_1 and K_2. We can see an example of this in figures 3a and 3c where a comparison between cases 3 and 5 is shown. First of all, notice that the plot 3c does not converge, and therefore, the curve named "Conv. Rates" indicates the speed of the iteration in this case. However, although the method does not converge in plot 3c we can observe that ϵ predicts the speed of the iteration well.

Finally, we want to highlight that the experimental convergence rates and $\|\mathbf{M}\|_2$ have

a similiar behavior in figures 2 and 3. However, when Δx tends to zero, $\|\mathbf{M}\|_2$ separates from the convergence rates. This shows that due to the unsymmetry the 2-norm is not a good estimate of the spectral radius when Δx tends to zero.

In conclusion, the main difference observed between the discrete and the semidiscrete cases is that the estimated convergence condition does not depend on the thermal diffusivities D_1 and D_2 for the discrete case.

6 CONCLUSIONS

We have described an approach for solving the coupling of two heat equations on two identical squared domains. In particular, the coupled PDEs were discretized into a system of algebraic equations. Afterwards, a fixed point iteration was performed and the exact iteration matrix was found. And finally, due to the complexity of computing the spectral radius of the iteration matrix, it was approximated by its block diagonal. Numerical results show the accuracy of this approach as well as a comparison between the semi discrete and the fully discrete cases.

The main difference observed between the discrete and the semidiscrete cases is that the estimated convergence condition does not depend on the thermal diffusivities D_1 and D_2 for the discrete case. Moreover, the estimated norm of the iteration matrix is observed to be below the semi discrete estimate β. This proves a faster convergence speed for the fully discrete case than for the semi discrete.

There are a variety of future directions for this work. More complicated domains can be taken into consideration. One can also couple a fluid together with a solid, the so-called fluid structure interaction. Another future direction will be to study the convergence speed of an actual non-linear application.

REFERENCES

[1] Birken, P. Termination criteria for inexact fixed point methods. *Numer. Linear Algebra Appl.*, accepted.

[2] Birken, P., Gleim, T., Meister, A. and Kuhl, D. Fast Solvers for Unsteady Thermal Fluid Structure Interaction. *Int. J. Numer. Meth. Fluids*, submitted.

[3] Buchlin, J.M. Convective Heat Transfer and Infrared Thermography. *J. Appl. Fluid Mech.* (2010) **3(1)**:55–62.

[4] Farhat, C. CFD-based Nonlinear Computational Aeroelasticity. In Stein, E., de Borst, R. and Hughes, T.J.R., editors, *Encyclopedia of Computational Mechanics*, volume 3: Fluids, chapter 13, pages 459-480. John Wiley & Sons, (2004).

[5] Fonseca, C.M. and Petronilho, J. Explicit inverses of some tridiagonal matrices. *Linear Algebra and its Applications.* (2001) **325**:7–21.

[6] Heck, U., Fritsching, U. and Bauckhage, K. Fluid flow and heat transfer in gas jet quenching of a cylinder. *International Journal of Numerical Methods for Heat & Fluid Flow* (2011) **11**:36–49.

[7] Henshaw, W.D. and Chand, K.K. A composite grid solver for conjugate heat transfer in fluid–structure systems. *Journal for Computational Physics.* (2009) **228**:2708–3741.

[8] Hinderks, M. and Radespiel, R. Investigation of Hypersonic Gap Flow of a Reentry Nosecap with Consideration of Fluid Structure Interaction. *AIAA Paper* (2006) **6**.

[9] Kowollik, D.S.C., Horst, P. and Haupt, M.C. Fluid-structure interaction analysis applied to thermal barrier coated cooled rocket thrust chambers with subsequent local investigation of delamination phenomena. *Progress in Propulsion Physics* (2013) **4**:617-636.

[10] Mehta, R.C. Numerical Computation of Heat Transfer on Reentry Capsules at Mach 5. *AIAA-Paper* (2005) **178**.

[11] Meyer, C.D. *Matrix analysis and applied linear algebra*. Siam, (2000).

[12] Quarteroni, A. and Valli, A. *Domain decomposition methods for partial differential equations*. Oxford Science Publications, (1999).

[13] Reuter, M.G. and Hill, J.C. An efficient, block-by-block algorithm for inverting a block tridiagonal, nearly block Toeplitz matrix. *Computational Science and Discovery.* (2012) **5**.

[14] Stratton, P., Shedletsky, I. and Lee, M. Gas Quenching with Helium. *Solid State Phenomena* (2006) **118**:221-226.

[15] Zhang, F. *The Schur complement and its applications*. Springer, (2005).

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis(Eds)

COUPLING REQUIREMENTS FOR WELL POSED AND STABLE MULTI-PHYSICS PROBLEMS

Jan Nordström*, Fatemeh Ghasemi†

*Department of Mathematics, Computational Mathematics, Linköping University, SE-581 83 Linköping, Sweden.
e-mail: jan.nordstrom@liu.se, web page: http://www.mai.liu.se/ janno11/

†Department of Mathematics, Computational Mathematics, Linköping University, SE-581 83 Linköping, Sweden
e-mail: fatemeh.ghasemi@liu.se - Web page: http://www.mai.liu.se/ fatgh43/

Key words: multi-physics problem, well posed problems, stability, coupling procedure, high order finite differences, summation-by-parts operatorsm, weak interface conditions

Abstract. We discuss well-posedness and stability of multi-physics problems by studying a model problem. By applying the energy method, boundary and interface conditions are derived such that the continuous and semi-discrete problem are well-posed and stable. The numerical scheme is implemented using high order finite difference operators on summation-by-parts (SBP) form and weakly imposed boundary and interface conditions. Numerical experiments involving a spectral analysis corroborate the theoretical findings.

1 INTRODUCTION

Roughly speaking, a well posed initial boundary value problem require that a unique solution that can be estimated in terms of the data, exist. The most common procedure for showing well posedness is the so called energy-method where one multiply the governing partial differential equations (PDEs) with the solution, integrate by parts and impose boundary conditions [1]. The same general knowledge is not wide-spread when it comes to the mathematical coupling of multi-physics problems. The reason for that is the more complex and to some extent more unclear nature of coupling conditions compared to imposing boundary conditions.

Firstly, accuracy relations must exist such that combinations of variables for one set of PDEs at the interface is equal to combinations of variables for the other set. Secondly, the number of accuracy relations must fit both problems. Too many conditions ruin existence and too few ruin uniqueness. If the number of accuracy relations are too few, additional conditions requiring external data must be added. Thirdly, the accuracy relations must be such that no artificial growth or decay is generated.

We will investigate the problems mentioned above and generalize the investigation in [2, 3] where we derived the coupling conditions by only demanding a well posed problem. Coupling of hyperbolic PDEs of different size at the interface will be our primary focus. Once the coupling conditions are known for the continuous multi-physics problem we will discretize using high order finite differences on summation-by-parts form and include the coupling conditions weakly using the SAT technique [4, 5].

2 THE MODEL PROBLEM

We will consider the following system,

$$\begin{aligned} u_t + Au_x &= 0, \qquad -1 \leq x \leq 0, \qquad t > 0, \\ u(x,0) &= f(x), \end{aligned} \tag{1}$$

and the scalar equation

$$\begin{aligned} v_t + bv_x &= 0, \qquad 0 \leq x \leq 1, \qquad t > 0, \\ v(x,0) &= g(x). \end{aligned} \tag{2}$$

In (1), $u = (u_1, u_2)^T$ is a vector of unknowns, $f(x) = (f_1(x), f_2(x))^T$ is a vector of given data and for simplicity we choose

$$A = \begin{bmatrix} 0 & a \\ a & 0 \end{bmatrix}, a > 0. \tag{3}$$

Two boundary/interface conditions are needed for the system (1) while equation (2) needs one boundary/interface condition.

2.1 The interface conditions

We apply the energy method to both equations and add them together to get

$$\frac{d}{dt}(\|u\|_2^2 + \alpha\|v\|_2^2) = -u^T Au|_{x=0} + \alpha b v^2|_{x=0} = w^T E w, \tag{4}$$

where α is positive free weight, $w = [u_1, u_2, v]^T$ and E is

$$E = \begin{bmatrix} 0 & -a & 0 \\ -a & 0 & 0 \\ 0 & 0 & \alpha b \end{bmatrix}. \tag{5}$$

In (4), the boundary terms at the outer boundaries $x = \pm 1$ are ignored. The eigenvalues of E are $\{a, -a, \alpha b\}$. If $b < 0$, one of the eigenvalues is positive and we need one condition at $x = 0$, otherwise we need two conditions.

In order to couple the problems we need at least one accuracy condition. Let

$$v = C^T u, \quad C = [c_1, c_2]^T. \tag{6}$$

The relation (6) inserted in equation (4) leads to

$$\frac{d}{dt}(\|u\|_2^2 + \alpha\|v\|_2^2) = u^T(0,t)Du(0,t), \tag{7}$$

where $D = (\alpha b C C^T - A)$. The characteristic polynomial related to the eigenvalues λ of D is

$$\lambda^2 - \alpha b(c_1^2 + c_2^2)\lambda + 2\alpha a b c_1 c_2 - a^2. \tag{8}$$

To simplify the following discussion we let $2s_1 = -\alpha b(c_1^2 + c_2^2)$ and $s_2 = 2\alpha a b c_1 c_2 - a^2$, which yields the roots

$$\lambda_{1,2} = -s_1 \pm \sqrt{s_1^2 - s_2}. \tag{9}$$

First we consider $b < 0$. This leads to a positive s_1. If $c_1 c_2 \leq a/2b\alpha$, then $s_2 \geq 0$ and both roots of the characteristic polynomial are negative, which means that D is negative definite. This means that if c_1 and c_2 have opposite sign, the coupled problems satisfy an energy estimate for all choices of α. But if c_1 and c_2 have the same sign, the energy estimate is not satisfied for any value of α. Consequently the coupled problems with the interface condition $v = C^T u$ satisfy an energy estimate for $b < 0$ if and only if c_1 and c_2 have opposite signs.

Next, consider $b > 0$. This leads to negative s_1 and at least one of the eigenvalues must be positive, which means that we need an additional condition. As mentioned above, two conditions are needed at $x = 0$. One of them is an interface condition and the other one must be such that the right-hand side of (7) is negative semi-definite. We will refer to this additional condition as a boundary condition. If $c_1 c_2 \leq a/2b\alpha$, then $s_2 \leq 0$ and one of the eigenvalues of D is positive (λ^+) and the other one is negative (λ^-). Let $D = Y\Lambda Y^T$ and rewrite (7) as

$$\frac{d}{dt}(\|u\|_2^2 + \alpha\|v\|_2^2) = u^T(0,t)(Y\Lambda Y^T)u(0,t), \tag{10}$$

where $\Lambda = \mathrm{diag}\{\lambda^+, \lambda^-\}$ and Y is the matrix of eigenvectors to D. Let $\Lambda = \Lambda^+ + \Lambda^-$, where $\Lambda^+ = \mathrm{diag}\{\lambda^+, 0\}$ and $\Lambda^- = \mathrm{diag}\{0, \lambda^-\}$. Furthermore we have $D = D^+ + D^-$ where $D^+ = Y\Lambda^+ Y^T$ and $D^- = Y\Lambda^- Y^T$. Then (10) leads to

$$\frac{d}{dt}(\|u\|_2^2 + \alpha\|v\|_2^2) = (Y^T u(0,t))^T(\Lambda^+ + \Lambda^-)(Y^T u(0,t)), \tag{11}$$

The most general condition based on (11) is

$$(Y_+^T - R_r Y_-^T)u(0,t) = h(t), \qquad x = 0, \tag{12}$$

where Y_+ and Y_- are the eigenvectors related to the positive and negative eigenvalues, respectively. Letting $h(t) = 0$ and inserting (12) into (11) leads to

$$\frac{d}{dt}(\|u\|_2^2 + \alpha\|v\|_2^2) = (\lambda^- + R_r^2\lambda^+)(Y_-^T u(0,t))^2. \tag{13}$$

If $\lambda^- + R_r^2\lambda^+ \leq 0$, then the right-hand side of (13) is bounded and we have a well-posed coupling. Note that with $R_r = 0$ we have the so called characteristic boundary conditions.

Consequently an energy estimate is obtained if c_1 and c_2 are chosen such that $c_1 c_2 \leq a/2b\alpha$. This means that c_1 and c_2 must be less than an arbitrary positive number that we can choose. In short: all values of c_1 and c_2 lead to a well-posed problem if $b > 0$.

2.2 The semi-discrete problem

Let A be an $M \times N$ matrix and B a $P \times R$ matrix. The Kronecker product of these matrices is defined as

$$A \otimes B = \begin{bmatrix} a_{11}B & \cdots & a_{1N}B \\ \cdots & & \cdots \\ a_{M1}B & \cdots & a_{MN}B \end{bmatrix}. \tag{14}$$

First, we consider $b < 0$. The semi-discrete SBP-SAT formulations of (1) and (2) are,

$$\begin{aligned} \mathbf{u}_t + (D_u \otimes A)\mathbf{u} &= (P_u^{-1}E_N^u \otimes \Sigma)(C^T\tilde{u}_N - v_0)e_N^u, \\ \mathbf{v}_t + bD_v\mathbf{v} &= P_v^{-1}\sigma(v_0 - C^T\tilde{u}_N)e_0^v. \end{aligned} \tag{15}$$

In (15), the outer boundary conditions are ignored as in the continuous case, $D_{u,v} = P_{u,v}^{-1}Q_{u,v}$ are the difference operators, $P_{u,v}$ are positive definite matrices and $Q_{u,v}$ satisfy $Q_{u,v} + Q_{u,v}^T = diag[-1, \cdots, 1]$. The discrete grid functions, related to the grid vectors $x_u = (x_0 = -1, \cdots, x_N = 0)$ and $x_v = (y_0 = 0, \cdots, y_M = 1)$ are

$$\mathbf{u} = (u_{10}, u_{20}, \cdots u_{1N}, u_{2N}), \qquad \mathbf{v} = (v_0, \cdots, v_M). \tag{16}$$

The vectors $e_N^u = (0, \cdots, 0, 1, 1)^T$ and $e_0^v = (1, \cdots, 0)^T$ are $2N \times 1$ and $M \times 1$, respectively. $E_N^u = diag[0, \cdots, 1]$ and $E_0^v = diag[1, \cdots, 0]$ are $N \times N$ and $M \times M$, respectively. The penalty matrix Σ is given by

$$\Sigma = \begin{bmatrix} \sigma_1 & \sigma_2 \\ \sigma_3 & \sigma_4 \end{bmatrix}, \tag{17}$$

σ is penalty parameter and $\tilde{u}_N = [u_{1N}, u_{2N}]^T$.

Next, consider $b > 0$. The semi-discrete SBP-SAT formulations of (1) and (2) are,

$$\begin{aligned}\mathbf{u}_t + (D_u \otimes A)\mathbf{u} &= (P_u^{-1}E_N^u \otimes \Sigma)(C^T\tilde{u}_N - v_0)e_N^u + (P_u^{-1}E_N^u \otimes \Xi)((I_N \otimes \tilde{H})\mathbf{u} - e_N^u \otimes \tilde{h}),\\ \mathbf{v}_t + bD_v\mathbf{v} &= P_v^{-1}\sigma(v_0 - C^T\tilde{u}_N)e_0^v.\end{aligned} \tag{18}$$

where the penalty matrix Ξ and $\tilde{H}$ are given by

$$\Xi = \begin{bmatrix} \chi_1 & \chi_2 \\ \chi_3 & \chi_4 \end{bmatrix}, \quad \tilde{H} = \begin{bmatrix} 1 & -R_r \\ 0 & 0 \end{bmatrix} Y^T. \tag{19}$$

The boundary data $\tilde{h}$ is defined as $\tilde{h} = [0, h]^T$. Also, in the following analysis we will use the discrete norms

$$\|\mathbf{u}\|^2_{P_u \otimes I} = \mathbf{u}^T(P_u \otimes I)\mathbf{u}, \qquad \|\mathbf{v}\|^2_{P_v} = \mathbf{v}^T P_v \mathbf{v}. \tag{20}$$

2.2.1 Stability conditions at the interface

First we consider $b < 0$. The discrete energy method is applied to (15) by multiplying the two equations with $\mathbf{u}^T(P_u \otimes I)$ and $\mathbf{v}^T P_v$, respectively. The SBP properties of $D_{u,v}$ yields

$$\frac{d}{dt}(\|\mathbf{u}\|^2_{P_u \otimes I} + \alpha_d\|\mathbf{v}\|^2_{P_v}) = -\tilde{u}_N^T A\tilde{u}_N + \alpha_d b v_0^2 + 2\tilde{u}_N^T \Sigma H + 2\alpha_d \sigma v_0(v_0 - C^T\tilde{u}_N). \tag{21}$$

In(21), α_d is a positive weight (not necessarily the same as in the continuous case) and $H = [C^T\tilde{u}_N - v_0, C^T\tilde{u}_N - v_0]^T$. In order to mimic the continuous case, we choose $\sigma_2 = c_1\alpha b/2$ and $\sigma_4 = c_2\alpha b/2$ and $\sigma_1 = \sigma_3 = 0$. The final penalty matrix in block form becomes $\Sigma = \alpha b/2 \begin{bmatrix} 0 & C \end{bmatrix}$. By inserting that into (21) we get

$$\frac{d}{dt}(\|\mathbf{u}\|^2_{P_u \otimes I} + \alpha_d\|\mathbf{v}\|^2_{P_v}) = \tilde{u}_N^T D\tilde{u}_N + \alpha_d v_0^2(b + 2\sigma) - \sigma v_0 C^T\tilde{u}_N(\alpha b + 2\alpha_d\sigma). \tag{22}$$

If we choose $\sigma = -\alpha b/2\alpha_d$, for $\alpha_d \leq \alpha$ the right-hand side of (22) will be bounded due to the continuous result above.

Next, we consider $b > 0$ and let $h(t) = 0$. Multiplying (18) by $\mathbf{u}^T(P_u \otimes I)$ and $\mathbf{v}^T P_v$ leads to

$$\frac{d}{dt}(\|\mathbf{u}\|^2_{P_u \otimes I} + \alpha_d\|\mathbf{v}\|^2_{P_v}) \leq \tilde{u}_N^T(D + \Xi\tilde{H} + (\Xi\tilde{H})^T)\tilde{u}_N, \tag{23}$$

where we have chosen Σ and σ as for the case $b < 0$. By using $YY^T = I$, we can rewrite the right-hand side of (23) as

$$\tilde{u}_N^T(D + \Xi\tilde{H} + (\Xi\tilde{H})^T)\tilde{u}_N = (Y^T\tilde{u}_N)^T(\Lambda + (Y^T\Xi\tilde{H}Y) + (Y^T\Xi\tilde{H}Y)^T)(Y^T\tilde{u}_N). \tag{24}$$

Let $\tilde{\Xi} = Y^T \Xi$ and choose Ξ such that $\tilde{\Xi} = diag(\tilde{\chi}_1, \tilde{\chi}_2)$. We also use the following split,

$$(Y^T \tilde{u}_N) = \begin{bmatrix} (Y_+^T \tilde{u}_N) \\ (Y_-^T \tilde{u}_N) \end{bmatrix}. \tag{25}$$

Now, we can rewrite (24) as,

$$\tilde{u}_N^T (D + \Xi \tilde{H} + (\Xi \tilde{H})^T) \tilde{u}_N = \begin{bmatrix} (Y_+^T \tilde{u}_N) \\ (Y_-^T \tilde{u}_N) \end{bmatrix}^T \begin{bmatrix} \lambda^+ + 2\tilde{\chi}_1 & -R_r \tilde{\chi}_1 \\ -R_r \tilde{\chi}_1 & \lambda^- \end{bmatrix} \begin{bmatrix} (Y_+^T \tilde{u}_N) \\ (Y_-^T \tilde{u}_N) \end{bmatrix}. \tag{26}$$

By the choice $\tilde{\chi}_1 = -\lambda^+$, the right-hand side of (26) can be rewritten as

$$(\lambda^- + R_r^2 \lambda^+)(Y_-^T \tilde{u}_N)^2 - \lambda^+ ((Y_+^T \tilde{u}_N) - R_r (Y_-^T \tilde{u}_N))^2, \tag{27}$$

which is negative due to the continuous result. Consequently if we choose

$$\Sigma = \alpha b/2 \begin{bmatrix} 0 & C \end{bmatrix}, \ \sigma = -\alpha b/2\alpha_d, \ \Xi = Y \begin{bmatrix} -\lambda^+ & 0 \\ 0 & 0 \end{bmatrix},$$

then for $\alpha_d \geq \alpha$, (18) is stable.

2.2.2 Stability conditions at the left boundary

In order to have a well-posed problem, we need one condition at $x = -1$. We consider the homogeneous boundary condition

$$(X_+^T - R_l X_-^T) u(-1, t) = 0, \tag{28}$$

with $|R_l| < 1$. The SAT term at $x = -1$ is $(P_u^{-1} E_0^u \otimes \Pi)(I_N \otimes \hat{H})\mathbf{u}$, where

$$\Pi = \begin{bmatrix} \pi_1 & \pi_2 \\ \pi_3 & \pi_4 \end{bmatrix}, \quad \hat{H} = \begin{bmatrix} 1 & -R_l \\ 0 & 0 \end{bmatrix} X^T, \tag{29}$$

and X is the matrix of eigenvectors to A. It can be shown that an energy estimate is obtained if

$$\pi_1 = -a/(4(R_l + 1)), \qquad \pi_2 = 0, \qquad \pi_3 = a/(4(R_l - 1)), \qquad \pi_4 = 0. \tag{30}$$

2.2.3 Stability conditions at the right boundary

For the case $b < 0$, one condition at $x = 1$ is also needed. We choose the homogeneuos $v(1, t) = 0$. The SAT term at $x = 1$ is $P_v^{-1} \theta v_N^2 e_N^v$ where θ satisfies

$$\theta \leq b/2. \tag{31}$$

3 THE SPECTRUM

In this section, we consider the continuous and discrete spectrum for our problem.

3.1 The spectrum for the continuous problem

By applying the Laplace transform to (1) and (2) we get the following system of ordinary differential equations

$$\begin{aligned} s\hat{u} + A\hat{u}_x = 0, \quad -1 \le x \le 0, \\ s\hat{v} + b\hat{v}_x = 0, \quad 0 \le x \le 1. \end{aligned} \tag{32}$$

We have ignored the initial conditions, since they do not influence the spectra and make the ansatz $\hat{u} = e^{kx}\psi$ and $\hat{v} = e^{k_3 x}\psi_3$. This leads to

$$(sI + Bk)\Psi = 0, \qquad B = \begin{bmatrix} A & 0 \\ 0 & b \end{bmatrix}, \tag{33}$$

where $\Psi = [\psi_1, \psi_2, \psi_3]^T$. This system of equations have a non-trivial solution only when $\det(sI + Bk) = 0$, which leads to $k_1 = -\frac{s}{a}, k_2 = \frac{s}{a}$ and $k_3 = -\frac{s}{b}$.

The general solution including the eigenvectors is

$$\hat{w} = \alpha_1 e^{-\frac{s}{a}x} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + \alpha_2 e^{\frac{s}{a}x} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} + \alpha_3 e^{-\frac{s}{b}x} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \tag{34}$$

where $\hat{w} = [\hat{u}_1, \hat{u}_2, \hat{v}]^T$. The unknowns α_1, α_2 and α_3, will be determined by the boundary and interface conditions.

First we consider $b < 0$, with the conditions

$$\begin{aligned} (X_+^T - R_l X_-^T)u(-1,t) = 0, \\ C^T u(0,t) - v(0,t) = 0, \\ v(1,t) = 0, \end{aligned} \tag{35}$$

and $|R_l| \le 1$. The interface and boundary conditions are such that the coupled problem is well-posed. By applying these conditions to (34), we obtain

$$E\alpha = 0, \qquad E = \begin{bmatrix} c_1 + c_2 & c_1 - c_2 & -1 \\ 2e^{\frac{s}{a}} & -2R_l e^{-\frac{s}{a}} & 0 \\ 0 & 0 & e^{-\frac{s}{b}} \end{bmatrix}, \tag{36}$$

where $\alpha = [\alpha_1, \alpha_2, \alpha_3]^T$. A non-trivial solution, require

$$\det(E) = 2e^{-\frac{s}{b}}(e^{\frac{s}{a}}(c_2 - c_1) - R_l e^{-\frac{s}{a}}(c_1 + c_2)) = 0. \tag{37}$$

The zeros of $\det(E)$ which form the spectrum of (32) are

$$s = \begin{cases} \frac{a}{2}\ln(|\frac{R_l(c_1+c_2)}{c_2-c_1}|) + na\pi i, & n \in \mathbb{Z}, \quad \text{if} \frac{R_l(c_1+c_2)}{c_2-c_1} > 0, \\ \frac{a}{2}\ln(|\frac{R_l(c_1+c_2)}{c_2-c_1}|) + na\pi i + \frac{a\pi i}{2}, & n \in \mathbb{Z}, \quad \text{if} \frac{R_l(c_1+c_2)}{c_2-c_1} < 0. \end{cases} \tag{38}$$

The real part of s is negative if $|\frac{R_l(c_1+c_2)}{c_2-c_1}| < 1$. It is easy to verify that this holds for $|R_l| < 1$ and c_1, c_2 with opposite sign. This means that if we choose c_1, c_2 and R_l such that the coupled problem leads to an energy estimate, then the real part of s will be negative. Recall that this required that c_1 and c_2 must have opposite signs.

Next, consider $b > 0$. In this case we have the conditions

$$\begin{aligned} (X_+^T - R_l X_-^T)u(-1,t) &= 0, \\ C^T u(0,t) - v(0,t) &= 0, \\ (Y_+^T - R_r Y_-^T)u(0,t) &= 0, \end{aligned} \tag{39}$$

and $\lambda^- + R_r^2\lambda^+ \leq 0$ and $R_l^2 \leq 1$. The coupled problems with the conditions (39) satisfy an energy estimate. By applying (39) to (34) leads to the following system of equations

$$E\alpha = 0, \qquad E = \begin{bmatrix} c_1 + c_2 & c_1 - c_2 & -1 \\ 2e^{\frac{s}{a}} & -2R_l e^{-\frac{s}{a}} & 0 \\ y_{12} - R_r y_{11} + 1 - R_r & y_{12} - R_r y_{11} - 1 + R_r & 0 \end{bmatrix}, \tag{40}$$

where $\alpha = [\alpha_1, \alpha_2, \alpha_3]^T$. The zeros of $\det(E)$ in this case are

$$s = \begin{cases} \frac{a}{2}\ln(|\frac{R_l(y_{12}-R_r y_{11}+1-R_r)}{-y_{12}+R_r y_{11}+1-R_r}|) + na\pi i, & n \in \mathbb{Z}, \quad \text{if} \frac{R_l(y_{12}-R_r y_{11}+1-R_r)}{-y_{12}+R_r y_{11}+1-R_r} > 0, \\ \frac{a}{2}\ln(|\frac{R_l(y_{12}-R_r y_{11}+1-R_r)}{-y_{12}+R_r y_{11}+1-R_r}|) + na\pi i + \frac{a\pi i}{2}, & n \in \mathbb{Z}, \quad \text{if} \frac{R_l(y_{12}-R_r y_{11}+1-R_r)}{-y_{12}+R_r y_{11}+1-R_r} < 0. \end{cases} \tag{41}$$

The real part of s is negative if

$$|\frac{R_l(y_{12} - R_r y_{11} + 1 - R_r)}{-y_{12} + R_r y_{11} + 1 - R_r}| < 1. \tag{42}$$

Note that the determinant of E is independent of c_1 and c_2. This means that if (42) holds, then the real part of s is negative for all c_1 and c_2. Recall that for all values of c_1 and c_2, suitable choices of α such that the coupled problems satisfy an energy estimate could be found. This implies that there is no limitation on c_1 and c_2 in both the energy and spectral analysis. However, recall that this required an additional boundary condition.

3.2 The semi-discrete spectrum

Consider $b < 0$. The SBP-SAT approximation of (1) and (2), including (35) is

$$\begin{aligned} \mathbf{u}_t + (D_u \otimes A)\mathbf{u} &= (P_u^{-1}E_0^u \otimes \Pi)(I_N \otimes \hat{H})\mathbf{u} + (P_u^{-1}E_N^u \otimes \Sigma)(C^T\tilde{u}_N - v_0)e_N^u, \\ \mathbf{v}_t + bD_v\mathbf{v} &= P_v^{-1}\sigma(v_0 - C^T\tilde{u}_N)e_0^v + P_v^{-1}\theta v_N e_N. \end{aligned} \tag{43}$$

In order to determine the semi-discrete spectrum, we follow [6] and rewrite (43) in matrix form as

$$W_t = P^{-1}(H_i + H_c)W, \tag{44}$$

where $W = [\mathbf{u}_1, \mathbf{u}_2, \mathbf{v}]^T$ and

$$H_i = \begin{bmatrix} -(Q_u \otimes A) & 0 \\ 0 & -bQ_v \end{bmatrix}, \quad P^{-1} = \begin{bmatrix} P_u^{-1} \otimes I_2 & 0 \\ 0 & P_v^{-1} \end{bmatrix}. \tag{45}$$

The penalty matrix H_c which is zero except at the boundaries and interface has the structure

$$H_c = \begin{bmatrix} \Pi\hat{H} & & & & & \\ & \ddots & & & & \\ & & EC^T & -E & & \\ & & -\sigma C^T & \sigma & & \\ & & & & \ddots & \\ & & & & & \theta \end{bmatrix}, \tag{46}$$

where $E = [\sigma_2, \sigma_4]^T$.

Next, consider $b > 0$. The SBP-SAT approximation of (1) and (2), with conditions (39) is

$$\begin{aligned} \mathbf{u}_t + (D_u \otimes A)\mathbf{u} &= (P_u^{-1}E_0^u \otimes \Pi)(I_N \otimes \hat{H})\mathbf{u} + (P_u^{-1}E_N^u \otimes \Sigma)(C^T\tilde{u}_N - v_0)e_N^u \\ &\quad + (P_u^{-1}E_N^u \otimes \Xi)(I_N \otimes \tilde{H})\mathbf{u}, \\ \mathbf{v}_t + bD_v\mathbf{v} &= P_v^{-1}\sigma(v_0 - C^T\tilde{u}_N)e_0^v. \end{aligned} \tag{47}$$

The approximation (47) can be written on the the form (44) where in this case

$$H_c = \begin{bmatrix} \Pi\hat{H} & & & & & \\ & \ddots & & & & \\ & & EC^T + \Xi\tilde{H} & -E & & \\ & & -\sigma C^T & \sigma & & \\ & & & & \ddots & \\ & & & & & 0 \end{bmatrix}, \tag{48}$$

while H_i is the same as before and given above. The eigenvalues of the matrix $P^{-1}(H_i+H_c)$ form the discrete spectrum of (43) and (47).

4 NUMERICAL RESULTS

In this section, we use the method of manufactured solution in order to test the accuracy of the approximations. RK3 is used to discretize time. We also discuss the relation between the continuous and semi-discrete spectrum.

	$SBP21$		$SBP42$		$SBP63$		$SBP84$	
N	*error*	*rate*	*error*	*rate*	*error*	*rate*	*error*	*rate*
20	2e-2	-	2e-3	-	2e-3	-	1e-3	-
40	6e-3	1.877	3e-4	3.006	1e-4	4.035	3e-5	5.376
80	1e-3	2.046	3e-5	3.242	8e-6	4.224	8e-7	5.392
160	4e-4	1.985	3e-6	3.052	4e-7	4.470	2e-8	5.113
320	1e-4	2.004	4e-7	3.021	2e-8	4.375	6e-10	5.091
640	2e-5	1.998	6e-8	3.013	1e-9	4.077	2e-11	5.047

Table 1: error and rate q_u for $b < 0$.

	$SBP21$		$SBP42$		$SBP63$		$SBP84$	
N	*error*	*rate*	*error*	*rate*	*error*	*rate*	*error*	*rate*
20	2e-1	-	3e-2	-	3e-2	-	4e-3	-
40	4e-2	2.148	3e-3	3.154	1e-3	4.469	2e-4	4.288
80	1e-2	2.050	4e-4	3.046	6e-5	4.704	8e-6	4.741
160	2e-3	2.014	5e-5	3.011	2e-6	4.668	3e-7	4.916
320	6e-4	2.005	6e-6	3.003	1e-7	4.474	9e-9	4.798
640	2e-4	2.001	7e-7	3.002	4e-9	4.467	3e-10	4.832

Table 2: error and rate q_v for $b < 0$.

4.1 Accuracy

The analytical solution that we use in the method of manufactured solution is

$$u_1(x,t) = u_2(x,t) = \cos(2\pi(x-t)), v(x,t) = \sin(3\pi(x-bt)). \tag{49}$$

The rate of convergence is calculated as

$$q_u = \ln\left(\frac{\|(\mathbf{u}_1^{N_1}, \mathbf{u}_2^{N_1}) - (u_1,u_2)\|_{P_u\otimes I}}{\|(\mathbf{u}_1^{N_2}, \mathbf{u}_2^{N_2}) - (u_1,u_2)\|_{P_u\otimes I}}\right) / \ln\left(\frac{N_1}{N_2}\right), q_v = \ln\left(\frac{\|\mathbf{v}^{N_1} - v\|_{P_v}}{\|\mathbf{v}^{N_2} - v\|_{P_v}}\right) / \ln\left(\frac{N_1}{N_2}\right), \tag{50}$$

where u_1, u_2 and v are the analytical solutions and $u_1^{N_i}, u_2^{N_i}$ and v^{N_i} are the corresponding numerical solutions with N_i grid points.

First, we consider $b < 0$. The choosen coefficients are $\alpha = \alpha_d = 1, a = 1, b = -1$. To have a well-posed problem, we choose $|R_l| < 1$ and c_1, c_2 such that $c_1 c_2 \leq -1/2$. Let $R_l = 0.25$ and $c_1 = 1, c_2 = -2$. Tables 1 and 2 show the error and convergence rate q_u and q_v, respectively, for SBP operators with $2^{th}, 3^{th}, 4^{th}$ and 5^{th} order. Next, we consider $b > 0$. We again choose $\alpha = \alpha_d = 1, a = 1, b = 1$ and take $R_l = 0.25, R_r = 0.25, c_1 = 1$ and $c_2 = 1$ in order to have a well-posed problem. Tables 3 and 4 show the error and convergence rates for q_u and q_v, respectively. Clearly, the design order of accuracy is obtained.

	$SBP21$		$SBP42$		$SBP63$		$SBP84$	
N	*error*	*rate*	*error*	*rate*	*error*	*rate*	*error*	*rate*
20	2e-2	-	5e-3	-	3e-3	-	2e-3	-
40	6e-3	1.885	6e-4	2.951	2e-4	4.186	1e-4	4.239
80	1e-3	2.062	8e-5	2.979	8e-6	4.367	3e-6	5.050
160	4e-4	1.983	1e-5	2.994	4e-7	4.492	8e-8	5.222
320	9e-5	2.005	1e-6	2.999	2e-8	4.392	2e-9	5.200
640	2e-5	1.998	1e-7	2.999	1e-9	4.321	5e-11	5.193

Table 3: error and rate q_u for $b > 0$.

	$SBP21$		$SBP42$		$SBP63$		$SBP84$	
N	*error*	*rate*	*error*	*rate*	*error*	*rate*	*error*	*rate*
20	3e-2	-	8e-3	-	7e-3	-	3e-3	-
40	7e-3	2.002	1e-3	3.083	4e-4	4.287	2e-4	4.427
80	1e-3	2.003	1e-4	2.954	1e-5	4.523	7e-6	4.542
160	4e-4	2.000	2e-5	2.983	7e-7	4.442	3e-7	4.773
320	1e-4	2.000	2e-6	2.989	3e-8	4.437	9e-9	4.753
640	3e-5	2.000	3e-7	2.994	2e-9	4.436	3e-10	4.749

Table 4: error and rate q_v for $b > 0$.

4.2 The spectrum of the continuous and semi-discrete operators

Figures 1-3 show the discrete and continuous spectrum for different grids using the SBP42 operator. One can clearly see the convergence of the discrete spectrum to the continuous one as the grids are refined. This convergence hold both for positive and negative b and show that the solutions of the semi-discrete scheme converge to the continuous one.

5 SUMMARY AND CONCLUSIONS

We have discussed well-posedness and stability of multi-physics problems by analyzing a model problem. It was shown that for ceartain wave speeds, only interface conditions were required, while in other cases additional information in the form of boundary conditions must be supplied.

By applying the energy method, we derived boundary and interface conditions such that the continuous and semi-discrete problem are well-posed and stable. The numerical scheme was implemented using high order finite difference operators on SBP form and weakly imposed boundary and interface conditions using the SAT technique.

It was shown that we obtained design order of accuracy, and that the spectrum of the

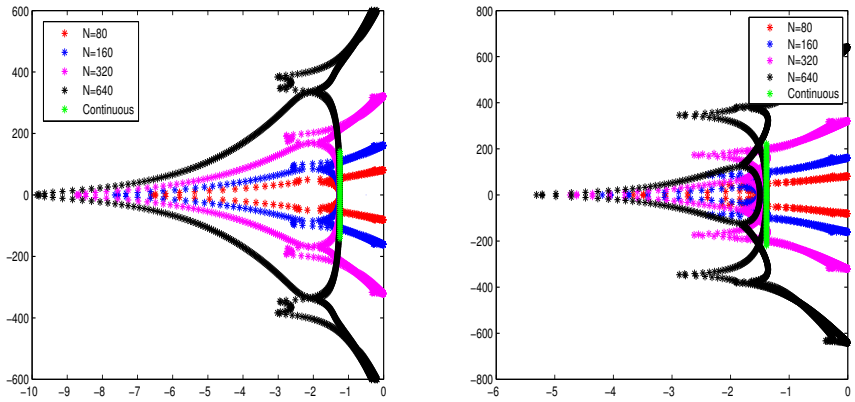


Figure 1: Global view: the discrete and continuous spectrum, $b < 0$ (left) and $b > 0$ (right).

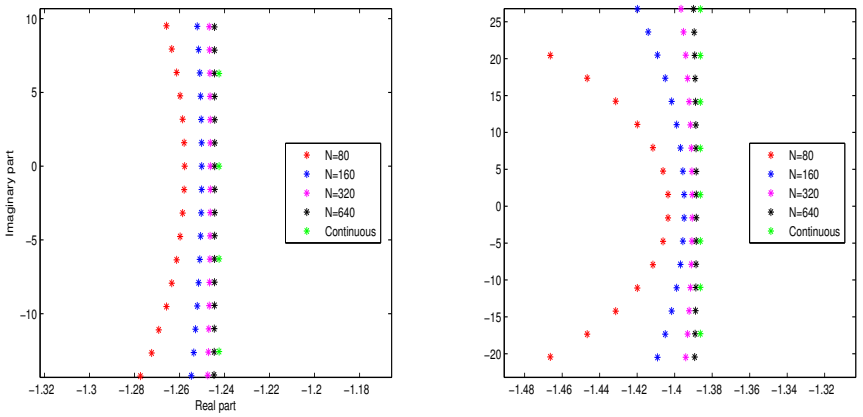


Figure 2: Medium view: the discrete and continuous spectrum, $b < 0$ (left) and $b > 0$ (right).

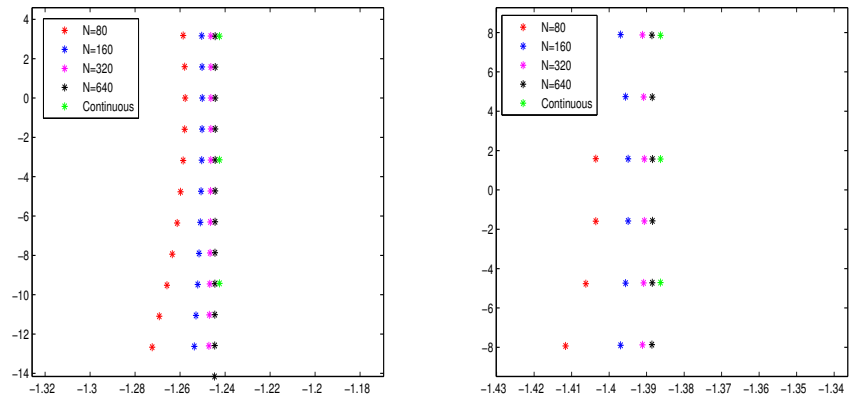


Figure 3: Zoomed view: the discrete and continuous spectrum, $b < 0$ (left) and $b > 0$ (right).

discrete operator converged to the spectrum of the continuous operator. The numerical experiments in combination with the theoretical derivations showed what type of analysis that is required to obtain accurate numerical simulations of multi-physics problems.

Future work will include a generalization of this investigation for hyperbolic problems, and an extension to coupling of incompletely parabolic problems such as the compressible Navier-Stoke's equations.

REFERENCES

[1] Nordström, J. and Svärd, M. Well-posed boundary conditions for the Navier-Stokes equations. *SIAM J. Numer. Anal.* (2005) **43**:1231-1255.

[2] Nordström, J. and Eriksson, S. Fluid structure interaction problems: the necessity of a well posed, stable and accurate formulation. *Commun. comput. phys(CiCP).* (2010) **8**:1111-1138.

[3] Lindström, J. and Nordström, J. A stable and high order accurate conjugate heat transfer problem. *J. Comput. Phys.* (2010) **229**:5440-5456.

[4] Nordström, J. and Berg, J. Conjugate heat transfer for the unsteady compressible Navier-Stokes equations using a multi-block coupling. *Comput Fluids.* (2013) **72**:20-29.

[5] Nordström, J., Gong, J., Van der Weide, E. and Svärd, M. A stable and conservative high order multi-block method for the compressible Navier-Stokes equations. *J. Comput. Phys.* (2009) **228**:9020-9035.

[6] Nordström, J. and Carpenter, M. H. High-order finite difference methods, multi-dimensional linear problems and curvilinear coordinates. *J. Comput. Phys.* (2001) **173**:149-174.

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis(Eds)

A COMPARISON OF VARIOUS QUASI-NEWTON SCHEMES FOR PARTITIONED FLUID-STRUCTURE INTERACTION

FLORIAN LINDNER[†], MIRIAM MEHL[†], KLAUDIUS SCHEUFELE[†], AND BENJAMIN UEKERMANN[*]

[†]Institute for Parallel and Distributed Systems, Universität Stuttgart
Universitätsstraße 38, 70569 Stuttgart, Germany
e-mail: miriam.mehl@ipvs.uni-stuttgart.de, web page: http://www.ipvs.uni-stuttgart.de/

[*]Institute for Advanced Study (IAS), Technische Universität München
Lichtenbergstraße 2a, 85748 Garching b. München, Germany
e-mail: uekerman@in.tum.de, web page: http://www5.in.tum.de

Key words: Fluid-Structure Interaction, Quasi-Newton, Partitioned Coupling

Abstract. During the last 5 years, quasi-Newton schemes have proven to be a robust and efficient way to couple partitioned fluid-structure interaction. We showed in previous work that they also allow to perform a parallel coupling. Bogaers et al. introduced a new variant based on a multi-vector update [14]. This variant renders a tuning of the reuse of old information unnecessary as all old iterations are implicitly covered in a Jacobian update. In this work, we compare this multi-vector variant in an inverse formulation to the classical IQN-ILS algorithm for serial as well as parallel coupling.

1 INTRODUCTION

The simulation of fluid-structure interactions is an important contribution to many fields in science and engineering – from aero-elasticity in aerospace engineering to hemodynamics in medical applications. At the same time, it is a very challenging type of multi-physics application as it tends to be ill-conditioned and unstable in particular for incompressible fluids.

Since fluid-structure interactions have been among the first multi-physics models considered for numerical simulation, the respective methods are in the meantime very sophisticated and, in particular, also a lot of very powerful monolithic solvers are available [1, 2, 3]. We focus, however, on partitioned approaches as our aim is to provide a general coupling tool, not only for fluid-structure interactions but also for multi-field coupling [4] and situation where flexible ad-hoc solutions combining different existing solvers are required. We have to design coupling numerics that are suited for black-box solvers where

nothing but the pure input and output data of the respective solver are accessible to the user. This enables us to use also commercial closed-source software as solvers can be coupled to our coupling software with minimal code changes or code wrappers. Thus, we consider the structural solver S and the flow solver F as mappings taking certain boundary values as an input and delivering other values at the coupling surface between fluid and structure as an output:

$$\begin{aligned} S : x_f &\mapsto x_d, \\ F : x_d &\mapsto x_f. \end{aligned}$$

Here, x_d denotes interface displacements or velocities at the wet surface between fluid and structure, x_f the forces or stresses modelling the impact of the fluid on the structure at the wet surface. This notation corresponds to the well-known Dirichlet-Neumann coupling between fluid and structure solvers. We are going to focus on this type of coupling throughout this paper. Alternative ways to define boundary conditions for fluid and structure solvers often require at least some knowledge on discretization details or Jacobians of the involved solvers. This is the case, e.g., for Robin-type boundary conditions [5].

Stability issues due to the physically strong coupling between fluid and structure, also referred to as the added mass effect (see [6, 7]), enforce implicit coupling within each time step. There are several ways how to realize the respective iterations which are different

1. in terms of the execution order of fluid and structure solvers,

2. in terms of the method used to stabilize and accelerate the coupling iteration.

Regarding the first, a staggered execution of fluid and structure solver, resulting in a Gauss-Seidel type iteration has been the common approach for a long time (cf. e.g. [8]). For compressible fluids, however, also the simultaneous execution of fluid and structure solvers, i.e., a Jacobian-type coupling, in many cases even as an explicit version has been in use ([8, 9]). We have presented a coupling based on the simultaneous execution of both solvers also for incompressible fluids and examples with strong instabilities in [10]. This iteration type is not feasible for a pure fixed-point iteration. Thus, sophisticated stabilization methods from 2. have to be used. We showed in [10] that neither constant nor adaptive Aitken underrelaxation was sufficient to accelerate the simultaneous coupling enough to make it competitive with the staggered iteration. Only quasi-Newton methods, that showed the best results also for the staggered coupling over the last years ([11, 12, 13]) showed a very good performance as presented in [10], even for the parallel coupling of more than two physical fields [4].

The combination of simultaneous solver execution with efficient quasi-Newton solvers results in an almost optimal partitioned fluid-structure interaction simulation method. However, the performance of the quasi-Newton methods depends on parameters such as the number of old time steps used to estimate Jacobians, for which only experience or try-and-error can be used to determine the optimal value. Therefore, we compare the

quasi-Newton approach used in [11] and [10] with an alternative approach presented in [14] for a different, block-iterative Newton solver. For this approach, information from passed time steps is implicitly used in the Jacobian estimation by suitable norm minimization conditions. The drawback, however, is that these methods explicitly determine and store Jacobians whereas the quasi-Newton approach from [11] only estimates the results of matrix-vector products involving parts of the Jacobian.

In the following[1], we first present the respective different coupling methods in Sect. 2, followed by numerical results in Sect. 3.

2 QUASI-NEWTON SCHEMES FOR PARTITIONED FLUID-STRUCTURE INTERACTION

We consider two different implicit coupling systems, namely a serial or staggered coupling and a parallel or vectorial coupling systems as mentioned above in Sect. 1. The parallel or vectorial system is preferable in case of massively parallel simulations in order to achieve a reasonable load balancing as described in [10]. We repeat the basics of both systems in Sect. 2.1.

Both equation systems are solved in two different ways: Using the known interface-quasi-Newton least squares approach as in [10, 11] and a multi-vector Jacobian approximation as described in [14] for a different Newton- iteration type, a block-iterative Newton method. The two solver alternatives are introduced in Sect. 2.2.

2.1 Execution Orders and Fixed Point Equations

In this section, we shortly recapitulate the introduction of fixed-point equations at the wet surface between fluid and structure depending on the execution order of the solvers.

Serial Implicit Coupling Scheme (S-System). If we execute the solvers in a staggered way, i.e., the flow solver first computes stresses or forces that are communicated to the structure solver that afterwards computes new wet surface displacements or velocities, the corresponding interface equation reads

$$x_d \stackrel{!}{=} S \circ F(x_d) \; . \tag{1}$$

Parallel Implicit Coupling Scheme (V-System) The above serial implicit coupling scheme offers some remarkable drawbacks regarding efficient parallelization. There is a substantial mismatch of work load between the structure and the fluid field solver, which does not allow for an efficient parallelization using the S-System. The only way to overcome these limitations in parallel efficiency is to evaluate the fluid and structure solver in parallel. Hereby, the V-System uses the original input/output relation for both

[1] A more detailed presentation of the findings of this paper including tables and figures has already been published in the master thesis of Klaudius Scheufele [15].

solvers but the boundary values are exchanged after each solve of the solvers executed in parallel. This leads to the vectorial fixed-point equation

$$\begin{pmatrix} x_f \\ x_d \end{pmatrix} \overset{!}{=} \begin{pmatrix} 0 & F \\ S & 0 \end{pmatrix} \begin{pmatrix} x_f \\ x_d \end{pmatrix} . \tag{2}$$

The vectorial system results in two independent instances of the S-System if solved by a pure fixed-point iteration, but quasi-Newton solvers turn out to be powerful enough such that one iteration of the V-System is comparable to one iteration of the S-System (cf. [4, 10]).

2.2 Quasi-Newton Solvers for the Interface Equations

In this section, we describe two quasi-Newton schemes, which can be applied to both aforementioned fixed-point equations (1) and (2). For sake of clarity, we introdue a unified notation:

$$x := \begin{cases} x_d \\ \begin{pmatrix} x_f \\ x_d \end{pmatrix} \end{cases} \quad \text{and} \quad H := \begin{cases} S \circ F & \text{for the staggered equation (1) ,} \\ \begin{pmatrix} 0 & F \\ S & 0 \end{pmatrix} & \text{for the parallel equation (2) .} \end{cases}$$

Thus, we have to solve the fixed-point equation

$$H(x) = x \;\Leftrightarrow\; R(x) := H(x) - x \overset{!}{=} 0 . \tag{3}$$

As a pure fixed-point iteration tends to be unstable and, in particular, two times slower for the parallel fixed-point equation compared to the staggered one (compare [10]), we use a quasi-Newton scheme as a stabilization and acceleration after each iteration $\tilde{x}^k = H(x^k)$. We, therefore, rewrite (3) to the equivalent inverse form

$$\tilde{R}(\tilde{x}) := \tilde{x} - H^{-1}(\tilde{x}) \overset{!}{=} 0 .$$

Now, the Newton iteration reads

$$\textbf{solve} \quad \left[I - J_{H^{-1}}(\tilde{x}^k)\right] \Delta\tilde{x}^k = J_{\tilde{R}}(\tilde{x}^k)\Delta\tilde{x}^k = -\tilde{R}(\tilde{x}^k) , \tag{4}$$

$$\textbf{set} \quad x^{k+1} = \tilde{x}^k + \Delta\tilde{x}^k . \tag{5}$$

As the exact Jacobian $J_{\tilde{R}}(x^k) := I - J_{H^{-1}}(\tilde{x}^k)$, however, is not accessible for black-box solvers, we work with an approximation $\widehat{J_{\tilde{R}}}(\tilde{x}^k)$. To minimize the computational cost, in particular of solving the system (4), we do not approximate the Jacobian itself, but its inverse $\widehat{J_{\tilde{R}}}^{-1}(\tilde{x}^k)$. To do so, we collect input-output data throughout our iterations within a time step and generate the following matrices:

$$W_k = (w_i^k)_{i=0}^{k-1} = \left[\Delta\tilde{x}_0^k, \Delta\tilde{x}_1^k, \cdots, \Delta\tilde{x}_{k-1}^k\right] , \qquad \text{with} \;\; \Delta\tilde{x}_i^k = \tilde{x}^k - \tilde{x}^i ,$$

$$V_k = (v_i^k)_{i=0}^{k-1} = \left[\Delta R_0^k, \Delta R_1^k, \cdots, \Delta R_{k-1}^k\right] , \qquad \text{with} \;\; \Delta R_i^k = R(x^k) - R(x^i) .$$

The inverse Jacobian approximately fulfills the secant equation, i.e.,

$$J_{\tilde{R}}^{-1}(\tilde{x}^k)\, V_k \approx W_k \ .$$

We use this as a system of equations for the entries of our approximate $\widehat{J_{\tilde{R}}}^{-1}(\tilde{x}^k)$. As k is in general much smaller than the number of degrees of freedom at the coupling interface, W_k and V_k are tall and thin matrices. Thus, we get an underdetermined system of equations for the entries of $\widehat{J_{\tilde{R}}}^{-1}(\tilde{x}^k)$:

$$\widehat{J_{\tilde{R}}}^{-1}(\tilde{x}^k) V_k = W_k \ . \tag{6}$$

In the following, we present two approaches to compute $\widehat{J_{\tilde{R}}}^{-1}(\tilde{x}^k)$ or the result of applying part of the approximate Jacobian to $-\tilde{R}(\tilde{x}^k) = -R(x^k)$, respectively, based on this secant equation.

2.2.1 Interface quasi-Newton Least Squares (IQN-LS)

The interface quasi-Newton least squares method presented in [11] uses the norm minimization

$$\left\| \widehat{J_{\tilde{R}}}^{-1}(\tilde{x}^k) \right\|_F \rightarrow \min \tag{7}$$

to enhance the secant equation (6) to a system with a unique solution for $\widehat{J_{\tilde{R}}}^{-1}(\tilde{x}^k)$. Here, $\| \cdot \|_F$ is the Frobenius norm. This gives the approximate inverse Jacobian

$$\widehat{J_{\tilde{R}}}^{-1}(x^k) = W_k \left(V_k^T V_k \right)^{-1} V_k^T$$

and the update formula

$$x^{k+1} = \tilde{x}^k + W_k \underbrace{(V_k^T V_k)^{-1} V_k \left(-R(x^k) \right)}_{=:\, \alpha} .$$

We do not have to explicitely compute the inverse Jacobian, but can restrict ourselves to compute only the vector α. This can be realized very efficiently by solving the least squares problem

$$\min_{\alpha \in \mathbb{R}^k} \| V_k \alpha + R(x^k) \|_2 \ ,$$

where $\| \cdot \|_2$ denotes the Euclidian norm.[2]

The convergence properties of the IQN-LS method can be greatly improved, if the input/output informations from previous time steps are incorporated into the secant equation, i.e., into W_k and V_k. To achieve this, the difference matrices $V^{n+1-R}, \cdots, V^n, V^{n+1}$

[2]Note that the update formula for x^{k+1} also shows that skipping the fixed point iteration step (computing $\tilde{x}^k = H(x^k)$) before using a quasi-Newton step would have lead to linearly dependent columns in W_k: We then would always correct x^k to x^{k+1} by adding multiples of differences $x^k - x^i$ from previous iterations as we would have to use $W_k = (\Delta x_0^k, \Delta x_1^k, \ldots, \Delta x_{k-1}^k)$ with $\Delta x_i^k = x^k - x^i$ in this case. Using induction over the iterations, we see that all columns of W_k would be in the space spanned by x^0 and $x^1 - x^0$.

and $W^{n+1-R}, \cdots, W^n, W^{n+1}$ from the previous $R \in \mathbb{N}$ time steps are stored and included in the secant equation of the current time step, i. e., we replace W_k and V_k by the enhanced versions

$$W_k^{(R)} = \left[W^{\{n+1-R\}}, W^{\{n-R\}}, \cdots, W_k^{\{n+1\}}\right],$$
$$V_k^{(R)} = \left[V^{\{n+1-R\}}, V^{\{n-R\}}, \cdots, V_k^{\{n+1\}}\right].$$

This additional information significantly improves the convergence as shown in [11] and Sect. 3. However, the optimal parameter R of reused time steps is highly problem dependent and there is no analytical method available to determine the optimal R. Thus, in practice, R has to be determined based on experiences and in a costly try-and-error process. Also, linear dependencies and contradicting information within the accumulated difference matrices need to be handled properly. The alternative quasi-Newton approach presented in the next section provides an automatic implicit incorporation of information from passed time steps and, thus, avoids these drawbacks of the IQN-LS method. However, this requires to explicitly compute $\widehat{J_{\tilde{R}}}^{-1}$ instead of only the short vector $\alpha \in \mathbb{R}^k$.

2.2.2 Interface quasi-Newton Multiple Vector Jacobian (IQN-MVJ)

The IQN-MVJ method presented here is a newly developed quasi-Newton fluid-structure coupling approach. It combines the idea of approximating the Newton iteration defined by (4) and (5) based on the secant equation (6) with the ideas presented in [14] for Jacobian approximations in the context of a block-iterative Newton method. To implicitly use information from previous time steps, the IQN-MVJ method uses a different norm minimization than the IQN-LS method in order to achieve uniqueness of the inverse Jacobian approximation:

$$\left\| \widehat{J_{\tilde{R}}}^{-1}(\tilde{x}^k) - \widehat{J_{\tilde{R}prev}}^{-1} \right\|_F \rightarrow \min , \tag{8}$$

where $\widehat{J_{\tilde{R}prev}}^{-1}$ denotes the last inverse Jacobian approximation of the previous time step. Thus, our approximations always stay as close as possible to the approximation from the last time step. This automatically guaranties that we profit from past information without having to explicitely use old W and V matrices again. We get the approximate inverse Jacobian

$$\widehat{J_{\tilde{R}}}^{-1}(\tilde{x}^k) = \widehat{J_{\tilde{R}prev}}^{-1} + \left(W_k - \widehat{J_{\tilde{R}prev}}^{-1} V_k\right) \left(V_k^T V_k\right)^{-1} V_k^T$$

and the update formula

$$\begin{aligned} x^{k+1} &= \tilde{x}^k + \widehat{J_{\tilde{R}}}^{-1}(\tilde{x}^k) \left(-R(x^k)\right) \\ &= x^k + \left(\widehat{J_{\tilde{R}prev}}^{-1} + \left(W_k - \widehat{J_{\tilde{R}prev}}^{-1} V_k\right) \left(V_k^T V_k\right)^{-1} V_k^T\right) \left(-R(x^k)\right). \end{aligned}$$

There is a close relation of the IQN-MVJ update scheme to the Broyden method: The Broyden method also minimizes distances between successively computed Jacobian approximations within a Newton iteration. However, Broyden minimizes the distance of approximations between two successive iterations, whereas we minimize the distance between Jacobian approximations in two successive time steps. This also implies that our updates are not only rank-one modifications as more than one new input/output pair is added for the update of the Jacobian approximation from the previous time step to the Jacobian approximation of the current time step.

In Algorithm 1, we oppose the two quasi-Newton approaches, the established IQN-LS method from [11] and our new variant. Summarizing Sect. 2, we state that we have described four different coupling methods, which are going to be considered for numerical test cases in Sect. 3. For clarity reasons, we give an overview of our combinations in Tab. 1.

Algorithm 1 Quasi-Newton schemes, **left:** the IQN-LS algorithm, a matrix free approach that approximates for the inverse Jacobian of the residual operator, **right:** the IQN-MVJ algorithm which explicitly stores the matrix of the Jacobian estimation. The inverse of the Jacobian of the residual operator is approximated using information from previous time steps implicitly.

IQN-LS
initial value x^0
$\tilde{x}^0 = H(x^0)$ and $R^0 = \tilde{x}^0 - x^0$
$x^1 = x^0 + 0.1 \cdot R^0$
for $k = 1 \dots$ **do**
  $\tilde{x}^k = H(x^k)$ and $R^k = \tilde{x}^k - x^k$
  $V_k = [\Delta R_0^k, \dots, \Delta R_{k-1}^k]$ with $\Delta R_i^k = R^i - R^k$
  $W_k = [\Delta \tilde{x}_0^k, \dots, \Delta \tilde{x}_{k-1}^k]$ with $\Delta \tilde{x}_i^k = \tilde{x}^i - \tilde{x}^k$
  decompose $V^k = Q^k U^k$
  solve the first k lines of $U^k \alpha = -{Q^k}^T R^k$
  $\Delta \tilde{x}^k = W\alpha$
  $x^{k+1} = \tilde{x}^k + \Delta \tilde{x}^k$
end for

IQN-MVJ(x)
initial value x^0 and $\widehat{J}_{\tilde{R}prev}^{-1} = 0$
$\tilde{x}^0 = H(x^0)$ and $R^0 = \tilde{x}^0 - x^0$
$x^1 = x^0 + 0.1 \cdot R^0$
for $k = 1 \dots$ **do**
  $\tilde{x}^k = H(x^k)$ and $R^k = \tilde{x}^k - x^k$
  $V_k = [\Delta R_0^k, \dots, \Delta R_{k-1}^k]$ with $\Delta R_i^k = R^i - R^k$
  $W_k = [\Delta \tilde{x}_0^k, \dots, \Delta \tilde{x}_{k-1}^k]$ with $\Delta \tilde{x}_i^k = \tilde{x}^i - \tilde{x}^k$
  $\widehat{J}_{\tilde{R}}^{-1}(\tilde{x}^k) = \widehat{J}_{\tilde{R}prev}^{-1} + \left(W_k - \widehat{J}_{\tilde{R}prev}^{-1} V_k\right) \left(V_k^T V_k\right)^{-1} V_k^T$
  $\Delta \tilde{x}^k = -\widehat{J}_{\tilde{R}}^{-1}(\tilde{x}^k) R^k$
  $x^{k+1} = \tilde{x}^k + \Delta \tilde{x}^k$
end for

	Serial System	**Vectorial System**
IQN-LS	S-IQN-LS	V-IQN-LS
IQN-MVJ	S-IQN-MVJ	V-IQN-MVJ

Table 1: Numerical coupling methods examined in this paper: each method is defined by choosing either the serial fixed-point equation (1) or the vectorial fixed-point equation (2) and either the least-squares quasi-Newton solver described in Sect. 2.2.1 or the multi-vector Jacobian approximation quasi-Newton approach from Sect. 2.2.2.

3 NUMERICAL RESULTS

3.1 Software

All numerical experiments were conducted using the coupling library preCICE and the simulation toolbox OpenFOAM. We give a brief summary of both below.

preCICE is a library for flexible numerical coupling of single-physics solvers. preCICE[3] is developed at the Technische Universität München and the Universität Stuttgart. It uses a partitioned black-box coupling approach, thus requiring only minimal modifications to existing solvers. Its software architecture as a library in conjunction with a high-level API fosters quick and minimal-invasive integration into existing codes. Integration can be realized in less than 30 lines of code. preCICE offers a wide variety of runtime configurable aspects of numerical coupling like serial and parallel as well as explicit and implicit coupling schemes. The latter category includes the schemes presented in this paper. Subcycling enables to combine problems that show convergence on different time scales. Coupling between non-matching grids can be achieved by data mapping methods ranging from simple projection methods to approaches based on radial-basis functions. preCICE is written in C++ and features a clean and modern software design with extensive unit and integration testing while maintaining minimal external dependencies and easy extensibility. For more information, refer to [12].

OpenFOAM The fluid and structure simulations were done using a software based on OpenFOAM[4] resp. the foam-extend-3.1 project[5]. The preCICE adapter and the actual solver were developed by David Blom et.al. from TU Delft. The fluid solver uses a 2^{nd} order finite volume discretization of the incompressible Navier-Stockes equation. Instead of the standard PISO (pressure implicit with splitting of operator) algorithm it uses a coupled solution algorithm as described in [16]. Time integration is performed by the second order backward differencing scheme. The fluid equations are formulated in the arbitrary-lagrangian-eulerian perspective whereas the mesh movement uses radial basis function interpolation [17]. The structural domain is modeled using a fully langrangrian formulation and a Saint Venant-Kirchhoff model. For more information, refer to e.g. [13].

3.2 Three-dimensional flow over an elastic structure

Scenario description. This case depicts a three-dimensional laminar and incompressible flow over an elastic structure. The scenario is assumed to be symmetric in the x/y-plane, hence, the simulation is only performed in one half of the domain. The geometry of the scenario and physical results are shown in Fig. 1 (a). The computational

[3] `http://www5.in.tum.de/wiki/index.php/PreCICE_Webpage`

[4] `http://www.openfoam.org/`

[5] `http://www.extend-project.de/`

Table 2: Three-dimensional flow over an elastic structure. Average numbers of coupling iterations for IQN-LS and IQN-MVJ, and for serial and parallel coupling. The iteration numbers are averaged over the first 40 time steps. Different numbers of reused time steps are evaluated.† For a well-conditioned V-system we scaled the forces by a factor of $4.0 \cdot 10^4$ for IQN-LS and 10^4 for IQN-MVJ.

Reuse (R)	0	1	2	3	4	5	6	7	8
S-IQN-LS(R)	5.15	4.75	4.50	4.45	4.65	5.05	5.65	5.70	6.15
S-IQN-MVJ(R)	5.15	5.22	5.57	5.87	6.32	6.70	7.00	7.45	7.62
V-IQN-LS(R)†	9.30	7.85	7.07	6.90	6.70	6.80	7.05	7.15	7.45
V-IQN-MVJ(R)†	6.65	6.75	6.82	7.27	7.47	7.72	8.02	8.10	8.52

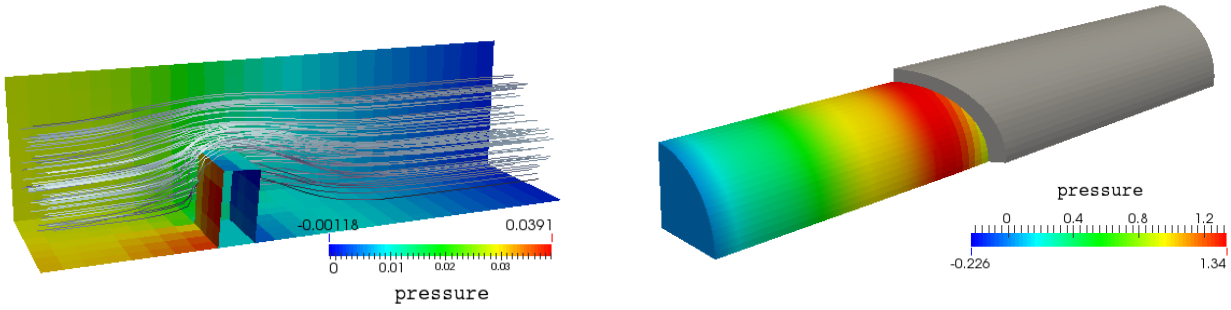
domain has a length of 1.5 m along with a width and height of 0.4 m, whereas the elastic rectangular structure with dimensions 0.2 m × 0.2 m × 0.2 m is mounted on the wall. The fluid flow is driven by a parabolic velocity profile v with peak velocity $v_{max} = 0.2\,\frac{\mathrm{m}}{\mathrm{s}}$ which is imposed as a Dirichlet boundary condition at the inflow boundary. The profile is faded in smoothly for $t < 2.0$. At the opposing side standard outflow conditions are applied, on the symmetry surface free-slip conditions as well as no-slip conditions on all remaining boundaries. The fluid density is $1 \times 10^3\,\frac{\mathrm{kg}}{\mathrm{m}^3}$, the dynamic viscosity $1 \times 10^{-3}\,\mathrm{Pa\,s}$. The density of the structure is $1 \times 10^3\,\frac{\mathrm{kg}}{\mathrm{m}^3}$, the Poisson ratio 0.4 and the Young's modulus is set to $1 \times 10^4\,\frac{\mathrm{N}}{\mathrm{m}^2}$.

Results. The simulation domain is decomposed into 1632 cells for the fluid and 32 cells for the structure mesh. Forty time steps with a time step size of 0.1 s are performed. A relative convergence measure of 10^{-5} for the forces as well as for the displacements at the fluid-structure interface is used. The results of the simulation are shown in Tab. 2. The average number of iterations over the first 40 time steps are shown for the IQN-LS as well as for the IQN-MVJ quasi-Newton scheme with different numbers of reused time steps. To precondition the V-system, we scale the forces in advance, such that forces and displacements have the same order of magnitude. It can be seen from Tab. 2, that the IQN-LS and the IQN-MVJ approach yield quite similar results for the S-system (serial coupling) as well as for the V-system (parallel coupling). As the IQN-MVJ implicitly incorporates information from previous time steps via the norm minimization in eq. (8) there is no benefit in collecting columns from previous time steps for the V and W matrices. Hence, the IQN-MVJ scheme is more robust and there is no need to tune for an optimal number of reused time steps (R).

3.3 Wave propagation in a three-dimensional elastic tube

Scenario description. This test case simulates the wave propagation in a straight, three-dimensional elastic tube (cf. e.g. [11]). The length of the tube is 0.05 m. The fluid

domain has a diameter of 0.003 m, whereas the tube thickness is 0.001 m. Both ends of the tube are fixed. For the initial duration of 0.003 s, the boundary condition for the pressure inlet is set to a fixed value of 1333.2 Pa. Thereafter, the inlet pressure is set to zero. At the outlet, the pressure is set to zero. The fluid has a density of $1 \times 10^3 \frac{\text{kg}}{\text{m}^3}$, and a dynamic viscosity of 3×10^{-3} Pa s. The density of the elastic structure is $1.2 \times 10^3 \frac{\text{kg}}{\text{m}^3}$, the Young's modulus $3 \times 10^5 \frac{\text{N}}{\text{m}^2}$, and the Poisson's ratio 0.3. A pressure pulse propagates through the tube as shown in Fig. 1 (b).



(a) (b)

Figure 1: Geometry and physics of the simulated scenarios. (a) Three-dimensional flow over an elastic structure. Streamlines and pressure contours in the fluid domain at $t = 2.8\,s$. (b) Wave propagation in a three-dimensional elastic tube. Geometry and pressure contours on the fluid-structure interface at $t = 8.9 \cdot 10^{-3}\,s$

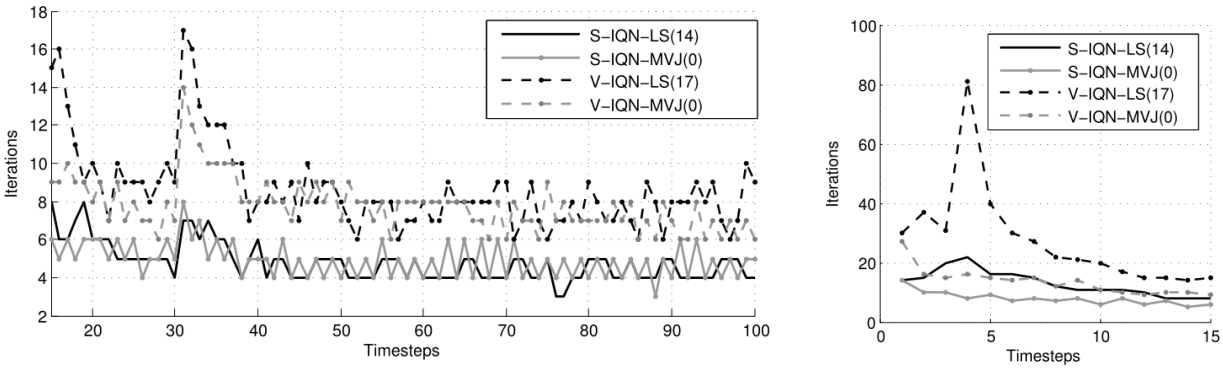
Results. The fluid mesh consists of 17600 cells, and the structure mesh contains 800 cells. Hundred time steps of 1×10^{-4} s are performed. A relative convergence measure of 10^{-5} is used for both, the displacements and the traction on the fluid-structure interface. The results of the simulation are shown in Tab. 3. It can be observed that the behavior of S-IQN-LS(9) and S-IQN-MVJ(0) is quite similar, cf. Fig. 2 (a), which confirms the results of Sect. 3.2. Also, it seems to be advantageous to reuse more time steps for the parallel coupling than for the serial coupling. Furthermore, for parallel coupling, it turns out that the IQN-MVJ scheme is more robust and performs slightly better than the IQN-LS. This dominance of the IQN-MVJ(0) over the IQN-LS(17) is mainly due to the fact that the IQN-MVJ is far more robust during the attack time of the simulation, as can be seen from Fig. 2 (b).

4 CONCLUSIONS

We applied the multi-vector update scheme from Bogaers et al. [14] on the inverse Jacobian formulation – the similar setting that was used by Degroote et al. to derive the classical IQN-ILS scheme [11]. We tested this new coupling scheme, the IQN-MVJ, as well as the IQN-ILS for serial and parallel coupling on two 3D scenarios. The multi-vector

Table 3: Wave propagation in a three-dimensional elastic tube. Average numbers of coupling iterations over first 100 time steps for IQN-ILS and IQN-MVJ, and for serial and parallel coupling. Different numbers of reused time steps are evaluated. † For a well-conditioned V-system, we scale the forces by a factor of 10^{10} for IQN-LS and $4.0 \cdot 10^8$ for IQN-MVJ.

Reuse(R)	0	1	3	5	7	9	11	13	15	17
S-IQN-LS(R)	9.97	9.45	6.80	6.35	6.17	6.05	6.05	6.06	6.15	6.21
S-IQN-MVJ(R)	5.36	8.97	9.47	9.32	9.25	9.21				
V-IQN-LS(R)†	21.41	18.63	13.73	12.90	12.28	11.95	11.71	13.36	13.36	11.47
V-IQN-MVJ(R)†	8.67	15.18	17.66	17.60	16.99	17.13				



(a) (b)

Figure 2: Wave propagation in a three-dimensional elastic tube. Iteration numbers for IQN-LS and IQN-MVJ, and for serial and parallel coupling. (a) timesteps 16 to 100, (b) timesteps 1 to 15.

update renders, as expected, a fine-tuning of the reuse of old information unnecessary. Similar to our previous findings for IQN-ILS [10], IQN-MVJ almost retains the convergence order, when moving from the serial coupling to the parallel one. In upcoming work, we want to automate the pre-conditioning of the parallel coupling.

Acknowledgement The financial support of the Institute for Advanced Study (IAS) of the Technische Universität München and of SPPEXA, the German Science Foundation Priority Programme 1648 – Software for Exascale Computing is thankfully acknowledged.

REFERENCES

[1] Bazilevs, Y., Takizawa, K. and Tezduyar, T. E. *Computational Fluid-Structure Interaction - Methods and Applications.* John Wiley and Sons, Vol. I., (2013).

[2] Gee, M., Küttler, U. and Wall, W. A. Truly monolithic algebraic multigrid for fluid-

structure interaction *Int. J. for Numer. Meth. in Eng.* (2011) **85**:987–1016.

[3] Crosetto, P., Deparis, S., Fourestey, G. and Quarteroni, A. Parallel algorithms for fluid-structure interaction problems in haemodynamics *Siam SISC* (2011) **33**:1598–1622.

[4] Bungartz, H.-J., Lindner F., Mehl M. and Uekermann, B. A plug-and-play coupling approach for parallel multi-field simulations *Comp. Mech.* (2014):1–11.

[5] Nobile, F. and Vergara, C. An effective fluid-structure interaction formulation for vascular dynamics by generalized Robin conditions *SIAM SISC* (2008) **30**:73–763.

[6] Causin, P., Gerbeau, J. F. and Nobile, F. Added-mass effect in the design of partitioned algorithms for fluid-structure problems. *Comput. Methods Appl. Mech. Eng.* (2005) **194**:4506–4527.

[7] Van Brummelen, E. H. Added mass effects of compressible and incompressible flows in fluid-structure interaction. *J. Appl. Mech.* (2009) **76**:1–7.

[8] Farhat, C. and Lesoinne, M. Two efficient staggered algorithms for the serial and parallel solution of three-dimensional nonlinear transient aeroelastic problems. *Comput. Method. Appl. M.* (2000) **182**:499–515.

[9] Ross, M.R., Felippa, C.A., Park, K.C. and Sprague, M.A. Treatment of acoustic fluid-structure interaction by localized Lagrange multipliers: Formulation. *Comput. Methods Appl. Mech. Eng.* (2008) **197**:305–3079.

[10] Uekermann, B., Bungartz, H. J., Gatzhammer, B. and Mehl, M. A parallel, black-box coupling algorithm for FSI *Proc. ECCOMAS Coupl. Prob.* (2013).

[11] Degroote, J., Bathe, K.-J. and Vierendeels, J. Performance of a new partitioned procedure versus a monolithic procedure in fluid-structure interaction *Comput. Struct.* (2009) **87**:793–801.

[12] Gatzhammer, B. Efficient and flexible partitioned simulation of fluid-structure interactions *Ph.D. thesis, Technische Universität München, Inst. für Informatik* (2015).

[13] Blom, D. S., van Zuijlen, A. H. and Bijl, H. Acceleration of strongly coupled fluid-structure interaction with manifold mapping *Proc. 11th. WCCM* (2014) 4484–4495.

[14] Bogaers, A. E. J., Kok, S., Reddy, B. D. and Franz, T. Quasi-Newton methods for implicit black-box FSI coupling *Comp. Meth. in Appl. Mech. and Eng.* (2014) **279**:113–132.

[15] Scheufele, K. Robust quasi-Newton methods for partitioned fluid-structure interaction *Master thesis, Universität Stuttgart, Fakultät für Informatik* (2015).

[16] Darwish, M., Sraj, I. and Moukalled, F. A coupled finite volume solver for the solution of incompressible flows on unstructured grids *J. of Comp. Phys.* (2009) **228**:180–201.

[17] de Boer, A., van Zuijlen, A. H. and Bijl, H. Radial basis functions for interface interpolation and mesh deformation *Lecture Notes in CSE* (2010) **71**:143–178.

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis(Eds)

A DYNAMICALLY ADAPTIVE LATTICE BOLTZMANN METHOD FOR PREDICTING WAKE PHENOMENA IN FULLY COUPLED WIND ENGINEERING PROBLEMS

RALF DEITERDING* AND STEPHEN L. WOOD†

*German Aerospace Center (DLR)
Institute for Aerodynamics and Flow Technology
Bunsenstr. 10, 37073 Göttingen, Germany
e-mail: ralf.deiterding@dlr.de

†University of Tennessee - Knoxville
The Bredesen Center
Knoxville, TN 37996, USA
e-mail: swood@utk.edu

Key words: Lattice Boltzmann Method, Fluid-Structure Coupling, Wind Turbines, Hinged Wing, Adaptive Mesh Refinement

Abstract. The essential components of a dynamically adaptive, parallel lattice Boltzmann method particularly tailored for coupled wind engineering are described. By utilizing a level set approach for geometry embedding the method can handle rotating and moving structures effectively and is thereby genuinely suited for fluid-structure coupling problems involving low-Mach number flows. The approach is validated for the canonical six degrees of freedom test case of a driven two-segment hinged wing. Subsequently, the wake field in an array of three Vestas V27 wind turbines at prescribed rotation rate and under constant inflow condition is simulated for two different scenarios. These results demonstrate that the time-explicit nature and the low dissipation properties of the lattice Boltzmann scheme in combination with dynamic mesh adaptation are able to predict well-resolved vortex structures created by realistic rotor speeds far downstream of the turbines at moderate computational costs.

1 Introduction

Many aerodynamic wind engineering problems are characterized by a strongly coupled interaction between moving structures and fluid flow. An example of particular technical relevance are horizontal axis wind turbines. Here, the incoming flow drives the motion and elastic deformation of the rotor, which in itself generates large-scale wake structures that can affect downstream turbines considerably. When multiple wind turbines are placed in

an array, the question of optimal placement arises. If a turbine is exposed to a major vortex field, its energy output will generally be reduced and additionally induced structural vibrations will cause disproportionately accelerated material fatigue. In order to simulate the flow field in a turbine array layout, e.g., for a dominant wind direction, vortex or disc actuator models are presently most frequently adopted. Computational fluid dynamics (CFD) with accurate consideration of the moving structures is still rarely applied, which is due to the complexities involved in solving the weakly or incompressible Navier-Stokes equations on moving three-dimensional meshes effectively, cf. [10].

As an alternative to the implicit, typically pressure-correction based CFD solution algorithms generally applied in wind engineering [10], we adopt in here the lattice Boltzmann method (LBM). The LBM is based on solving the Boltzmann equation in a specially chosen, discrete phase space and is fully explicit in time [7]. The scheme is normally constructed on uniform Cartesian grids and geometrically complex boundaries are considered with an immersed boundary approach, making the method well suited for considering moving structures. Here, we utilize a level set distance function to represent embedded objects. Dynamic mesh adaptation is applied in addition in order to increase the local resolution based on the level set function and features detected in the flow field [4]. Distributed memory parallelization is adopted to allow for large-scale simulations.

The paper is organized as follows: In Section 2, we recall the construction principles of the LBM and our embedded boundary treatment method. In Section 3, the block-based mesh adaptation procedure and in particular the incorporation of the LBM are presented. Section 4 explains our approach in dealing with embedded geometries. Section 5 discusses a coupled validation configuration of a two-segment hinged wing with torsion damper and Section 6 presents first simulations of entire wind turbines with prescribed motion of the rotor that demonstrate the benefit of the proposed overall approach. The conclusions are given in Section 7.

2 Lattice Boltzmann method

The lattice Boltzmann method is based on computing approximations of the Boltzmann equation with a simplified collision operator

$$\partial_t f + \mathbf{u} \cdot \nabla f = \omega(f^{eq} - f) \tag{1}$$

on a rectangular grid of characteristic domain length L with isotropic mesh spacing Δx under the assumption of a small Knudsen number $\mathrm{Kn} = l_f/L \ll 1$, where the mean free path length l_f is replaced with Δx. A crucial idea of the LBM is to approximate Eq. (1) in a specially chosen discrete phase space, in which a partial density distribution function $f_\alpha(\mathbf{x}, t)$ is associated to every discrete lattice velocity $\mathbf{e}_\alpha$. The total density distribution is given as $\rho(\mathbf{x}, t) = \sum_\alpha f_\alpha(\mathbf{x}, t)$ and the macroscopic moments as $\rho(\mathbf{x}, t)u_i(\mathbf{x}, t) = \sum_\alpha \mathbf{e}_{\alpha i} f_\alpha(\mathbf{x}, t)$. A splitting approach is then adopted that first solves the homogeneous transport equation with the time-explicit update step

$$\mathcal{T}: \qquad \tilde{f}_\alpha(\mathbf{x} + \mathbf{e}_\alpha \Delta t, t + \Delta t) = f_\alpha(\mathbf{x}, t). \tag{2}$$

Here, we apply the D3Q19 model for which the lattice velocities are defined as

$$\mathbf{e}_\alpha = \begin{cases} 0, & \alpha = 0, \\ (\pm 1, 0, 0)c, (0, \pm 1, 0)c, (0, 0, \pm 1)c, & \alpha = 1, \ldots, 6, \\ (\pm 1, \pm 1, 0)c, (\pm 1, 0, \pm 1)c, (0, \pm 1, \pm 1)c, & \alpha = 7, \ldots, 18, \end{cases}$$

with $c = \Delta x/\Delta t$. The physical speed of sound c_s is related to c by $c_s = c/\sqrt{3}$. The right-hand of Eq. (1) is integrated subsequently by the collision operator

$$\mathcal{C} : f_\alpha(\cdot, t + \Delta t) = \tilde{f}_\alpha(\cdot, t + \Delta t) + \omega \Delta t \left(\tilde{f}_\alpha^{eq}(\cdot, t + \Delta t) - \tilde{f}_\alpha(\cdot, t + \Delta t) \right) \tag{3}$$

with equilibrium function

$$f_\alpha^{eq}(\rho, \mathbf{u}) = \rho t_\alpha \left[1 + \frac{3\mathbf{e}_\alpha \mathbf{u}}{c^2} + \frac{9(\mathbf{e}_\alpha \mathbf{u})^2}{2c^4} - \frac{3\mathbf{u}^2}{2c^2} \right] \tag{4}$$

with $t_0 = 1/3$, $t_\alpha = 1/18$ for $\alpha = 1, \ldots, 6$ and $t = 1/36$ for $\alpha = 7, \ldots, 18$. The variation in hydrodynamic pressure for the equilibrium function (4) reads $\delta p = \sum_\alpha f_\alpha^{eq} c_s^2 = (\rho - \rho_0) c_s^2$.

Applying a Chapman-Enskog expansion procedure, it can be shown [8] that the sketched LBM converges to a solution of the weakly compressible Navier-Stokes equations

$$\partial_t \rho + \nabla \cdot (\rho \mathbf{u}) = 0, \tag{5a}$$

$$\partial_t \mathbf{u} + \mathbf{u} \cdot \nabla \mathbf{u} = -\nabla p + \nu \nabla^2 \mathbf{u}. \tag{5b}$$

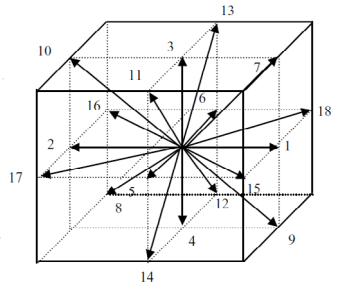


Figure 1: The velocities $\mathbf{e}_\alpha$ of the D3Q19 lattice.

It can be shown further, cf. [7], that the kinematic viscosity ν and collision frequency ω are connected by the relation

$$\omega = \tau^{-1} = \frac{c_s^2 \Delta t}{\nu + \Delta t c_s^2/2}. \tag{6}$$

While the sketched model can be used directly to simulate laminar flows, it is mandatory to apply a turbulence model in addition in high Reynolds number situations. In the context of LBM, it is common to adopt a large eddy simulation approach and assume that the partial density distribution functions used in the scheme represent the resolved scales. The subgrid scale turbulence is then considered by adding a turbulent viscosity ν_t to the physical one and by utilizing the effective viscosity $\nu^\star = \nu + \nu_t = \frac{1}{3}\left(\tau^\star - \frac{1}{2}\right) c\Delta x$ with $\tau^\star = \tau + \tau_t =: 1/\omega^\star$ in (3) throughout the scheme. Like Hou *et al.* [8], we apply the Smagorinsky model to evaluate ν_t, for which $\nu_t = (C_{sm}\Delta x)^2 \bar{S}$ is used, with

$$\bar{S} = \sqrt{2 \sum_{i,j} \bar{\mathbf{S}}_{ij} \bar{\mathbf{S}}_{ij}}, \qquad \bar{\mathbf{S}}_{ij} = -\frac{1}{2\rho_0 c_s^2 \tau} \sum_a \mathbf{e}_{\alpha i} \mathbf{e}_{\alpha j} (\bar{f}_\alpha - \bar{f}_\alpha^{eq}) \tag{7}$$

yielding

$$\tau_t = \frac{1}{2} \left(\sqrt{\tau^2 + 18\sqrt{2}(\rho_0 c^2)^{-1} C_{sm}^2 \Delta x \bar{S}} - \tau \right). \tag{8}$$

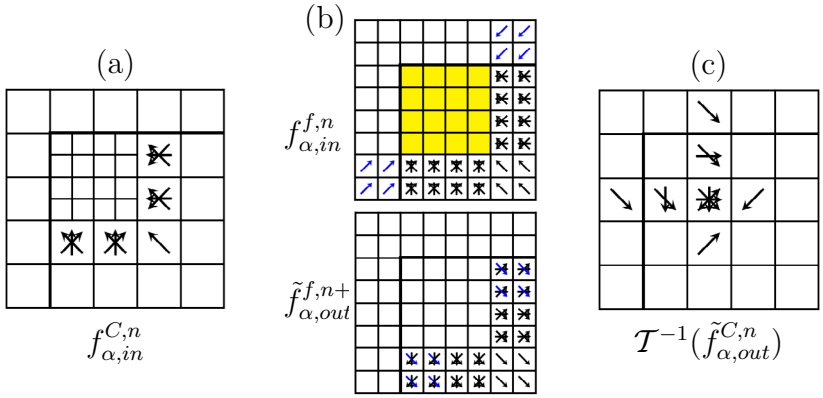


Figure 2: Visualization of partial density distribution functions involved in necessary data exchange at a coarse-fine boundary. The thick black lines indicate a physical boundary. (a) Coarse distributions going into fine grid; (b) ingoing interpolated fine distributions in halos (top), outgoing distributions in halos after two fine-level transport steps (bottom); (c) averaged distributions replacing coarse values before update is repeated in cells next to boundary.

3 Dynamic mesh adaptation

For local dynamic mesh adaptation we have adopted the block-structured adaptive mesh refinement (SAMR) method after Berger & Collela [1]. In order to fit smoothly into our existing, fully parallelized finite volume SAMR software system AMROC [4], we have implemented the LBM cell-based, which makes the scheme also conservative in ρ and ρu_i. In the SAMR approach, finite volume cells are clustered with a special algorithm into non-overlapping rectangular grids. The grids have a suitable layer of halo cells for synchronization and applying inter-level and physical boundary conditions. Refinement levels are integrated recursively. The spatial mesh width Δx_l and the time step Δt_l are refined by the same factor r_l, where we assume $r_l \geq 2$ for $l > 0$ and $r_0 = 1$. Note that in an adaptive LBM the collision frequency ω_l is not a constant but needs to be adjusted according to Eq. (3) for the update on each level. In addition to this, the interface region requires a specialized treatment. Distinguishing between the transport and collision operators, $\mathcal{T}$ and $\mathcal{C}$, cf. Eqs. (2) and (3), the steps of our method for a refinement factor of 2 are:

1. Complete update on coarse grid: $f^{C,n+1}_{\alpha} := \mathcal{C}\mathcal{T}(f^{C,n}_{\alpha})$
2. Use coarse grid distributions $f^{C,n}_{\alpha,in}$ that propagate into the fine grid, cf. Fig. 2(a), to construct initial fine grid halo values $f^{f,n}_{\alpha,in}$, cf. Fig. 2(b).
3. Complete transport $\tilde{f}^{f,n}_{\alpha} := \mathcal{T}(f^{f,n}_{\alpha})$ on whole fine mesh. Collision $f^{f,n+1/2}_{\alpha} := \mathcal{C}(\tilde{f}^{f,n}_{\alpha})$ is applied only in the interior cells (yellow in Fig. 2(b)).
4. Repeat 3. to obtain $\tilde{f}^{f,n+1/2}_{\alpha} := \mathcal{T}(f^{f,n+1/2}_{\alpha})$ and $f^{f,n+1}_{\alpha} := \mathcal{C}(\tilde{f}^{f,n+1/2}_{\alpha})$.

5. Average outgoing distributions from fine grid halos (Fig. 2(c)), that is $\tilde{f}^{f,n+1/2}_{\alpha,out}$ in the inner halo layer and $\tilde{f}^{f,n}_{\alpha,out}$ (outer halo layer) to obtain $\tilde{f}^{C,n}_{\alpha,out}$.

6. Revert transport for averaged outgoing distributions, $\bar{f}^{C,n}_{\alpha,out} := \mathcal{T}^{-1}(\tilde{f}^{C,n}_{\alpha,out})$, and overwrite those in the previous coarse grid time step, cf. Fig. 2(d).

7. Parallel synchronization of $f^{C,n}_{\alpha}, \bar{f}^{C,n}_{\alpha,out}$ on entire level.

8. Repeat complete update on coarse grid cells next to coarse-fine boundary only: $f^{C,n+1}_{\alpha} := \mathcal{CT}(f^{C,n}_{\alpha}, \bar{f}^{C,n}_{\alpha,out})$

This algorithm is computationally equivalent to the method by Chen *et al.* [2] but tailored to the SAMR recursion that updates coarse grids in their entirety before fine grids are computed. Because of the nonlinearity of the collision operator $\mathcal{C}$ it becomes necessary under this paradigm to repeat the LBM update for those coarse grid cells that share a face or corner with a fine grid.

4 Embedded structure handling

We represent non-Cartesian boundaries implicitly on the adaptive Cartesian grid by utilizing a scalar level set function φ that stores the distance to the boundary surface. The boundary surface it located exactly at $\varphi = 0$ and the boundary outer normal in every mesh point can be evaluated as $\mathbf{n} = -\nabla\varphi/|\nabla\varphi|$ [3]. We treat a fluid cell as an embedded ghost cell if its *midpoint* satisfies $\varphi < 0$.

In order to implement non-Cartesian boundary conditions with the LBM, we have chosen to pursue a 1st order accurate ghost fluid approach that was already available in AMROC [3]. In our technique, the density distributions in embedded ghost cells are adjusted to model the boundary conditions of a non-Cartesian reflective wall moving with velocity $\mathbf{w}$ before applying the unaltered LBM. The last step involves interpolation and mirroring of ρ, $\mathbf{u}$ across the boundary to ρ' and $\bar{\mathbf{u}}$ and modification of the macro velocity in the immersed boundary cells to $\mathbf{u}' = 2\mathbf{w} - \bar{\mathbf{u}}$, cf. [4]. From the newly constructed macroscopic values the density distributions in the embedded ghost cells are simply set to $f^{eq}_{\alpha}(\rho', \mathbf{u}')$.

Real-world geometries are considered in AMROC as triangular surface meshes. The computation of the level set distance information in every Cartesian mesh point could principally be accomplished by simply iterating over the entire surface mesh; yet, this would lead to detrimental performance for increasing mesh size. The problem is equivalent to determining for every Cartesian cell the closest facet on the surface mesh. For this purpose, we employ a specially developed algorithm based on characteristic reconstruction and scan conversion developed by Mauch [9] that is used to compute the distance exactly only in a small band around the embedded structure.

The dynamics of multi-body systems undergoing interaction with the fluid are modeled as sets of triangulated surface meshes configured in kinetic chains. The dynamics of

these mechanisms are solved by a recursive Newton-Euler method at each time step [12]. Considering an arbitrary link with a coordinate frame located at point $\mathbf{P}$ that is not coincident with its associated body's center of mass, the force and torque applied by the preceding link are

$$\begin{pmatrix} \mathbf{F} \\ \boldsymbol{\tau}_{\mathrm{P}} \end{pmatrix} = \begin{pmatrix} m\mathbf{1} & -m[\mathbf{c}]^{\times} \\ m[\mathbf{c}]^{\times}\mathbf{I}_{\mathrm{cm}} & -m[\mathbf{c}]^{\times}[\mathbf{c}]^{\times} \end{pmatrix} \begin{pmatrix} \mathbf{a}_{\mathrm{P}} \\ \boldsymbol{\alpha} \end{pmatrix} + \begin{pmatrix} m[\boldsymbol{\omega}]^{\times}[\boldsymbol{\omega}]^{\times}\mathbf{c} \\ [\boldsymbol{\omega}]^{\times}(\mathbf{I}_{\mathrm{cm}} - m[\mathbf{c}]^{\times}[\mathbf{c}]^{\times})\,\boldsymbol{\omega} \end{pmatrix}. \tag{9}$$

Here, we additionally define the total force and torque acting on a body, $\mathbf{F} = (\mathbf{F}_{FSI} + \mathbf{F}_{prescribed}) \cdot \boldsymbol{\mathcal{C}}_{xyz}$ and $\tau = (\tau_{FSI} + \tau_{prescribed}) \cdot \boldsymbol{\mathcal{C}}_{\alpha\beta\gamma}$ respectively. Where $\boldsymbol{\mathcal{C}}_{xyz}$ and $\boldsymbol{\mathcal{C}}_{\alpha\beta\gamma}$ are the translational and rotational constraints, respectively. $\mathbf{F}_{FSI}$ and τ_{FSI} are determined for each body by integrating the fluid pressure on the triangular facets of the respective body's surface mesh. Each surface mesh is associated with a kinetic link in a chain that begins with a base link in the global coordinate frame. Links are connected by joints that may be independently constrained in six degrees of freedom relative to the preceding link. The evolution of the triangular surface mesh as well as the velocity $\mathbf{w}$ in each node are communicated to the LBM fluid solver in dedicated coupling time steps. The data exchange corresponds to the time step of an SAMR level but this does not have to be the finest refinement level available, cf. [5].

Triangulated surface meshes can be generated by lofting closed 3D curves or read from the standard STL format. Each surface mesh is associated with a kinetic link in a chain that begins with a base link in the global coordinate frame. Links are connected by joints that may be independently constrained in six degrees of freedom relative to the preceding link. Constraints, $\boldsymbol{\mathcal{C}}$, which may be prescribed motions or reaction forces are enforced during the backward calculation step of the Newton-Euler method as joint forces are calculated by proceeding from the distill link of a kinetic chain to the base link. The angular and linear position, velocity and acceleration of each link in terms of its preceding link are calculated during the forward calculation step. This formulation readily facilitates the analysis of motions, forces, and moments on each link and triangulated surface in the global coordinate frame or in any of the link coordinate frames.

5 Fluid-structure interaction validation

A canonical problem of fluid-structure interaction and wake prediction proposed by Toomey & Eldredge [11] is selected for validating the fluid-structure interaction simulation method. This model, depicted in Fig. 3, utilizes a system of two articulated rigid bodies connected by a torsion spring and damper. The kinematics of the centroid of the driven wing are prescribed, while the trailing body responds passively to the aerodynamic and inertial/elastic forces. The principle unknown in this rigid body dynamics problem is the hinge angle θ. The parametric kinematic equations

$$X_t(t) = \frac{A_0}{2}\frac{G_t(ft)}{\max Gt}C(ft), \quad \alpha_1(t) = -\beta\frac{G_r(ft)}{\max Gr}, \quad G_t(t) = \int_t \tanh[\sigma_t cos(2\pi t')]dt', \tag{10}$$

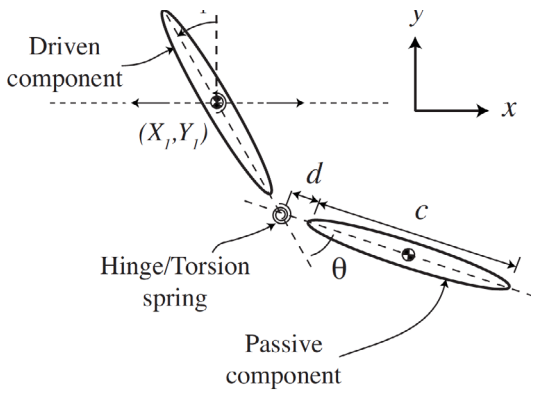


Figure 3: Model system consisting of two rigid elliptical sections connected by a hinge with torsion spring and damper.

Table 1: Kinematic parameters

$A_0\,(cm)$	7.1
$c\,(cm)$	5.1
$d\,(cm)$	0.25
β	$\pi/4$
σ_t	0.628
σ_r	0.628
Φ	0
Re_t	73, 370
Re_r	100, 500
$\rho_b\,(kg/m^3)$	5080
$f\,(Hz)$	0.15

Table 2: Nondimensional mean and peak force and moments

	$Re_r = 100$		$Re_r = 500$	
	Mean	Peak	Mean	Peak
F_x	2.78	430.2	3.78	429.8
$F_x\,(\%)$	-3.3	-9.8	-3.4	-9.9
F_y	39.0	206.2	52.3	210.2
$F_y\,(\%)$	-3.4	-15	-5	-7
M	-3.3	248	0.9	228.5
$M\,(\%)$	26	30	17	19

$$G_r(t) = \tanh[\sigma_r cos(2\pi t + \Phi)], \quad C(t) = \frac{\tanh(8t-2)+\tanh(2)}{1+\tanh(2)}. \tag{11}$$

describe the motion of the driven body. The parameters utilized in this work and in Case 1 of [11] to specify the kinematics through the translational, $G_t(t)$, and rotational, $G_r(t)$, shape functions are given in Table 1. The start-up conditioner, $C_t(t)$, is applied to the translational kinematics to avoid an impulsive start. The translational and rotational Reynolds numbers are based on the peak translational, V, and rotational, $2\pi\beta\sigma_r fc/\tanh(\sigma_r)$, velocities as shown in

$$Re_t = Vc/\nu, \quad Re_r = 2\pi\beta\sigma_r fc^2/(\tanh(\sigma_r)\nu). \tag{12}$$

The mean and peak values of the dimensionless fluid dynamic force, $F_{x,y} = 2F^*_{x,y}/(\rho_f^2 c^3)$, and moment, $M = 2M^*/(\rho_f f^2 c^4)$, generated by the wing motion are presented in Table 2.

The relative error between the results of this work and those in [11] is also presented in Table 2. The wing deflection and vorticity production at Re = 500 are depicted in Fig. 4. Figure 5 displays the hinge deflection angle for experiments and simulations in

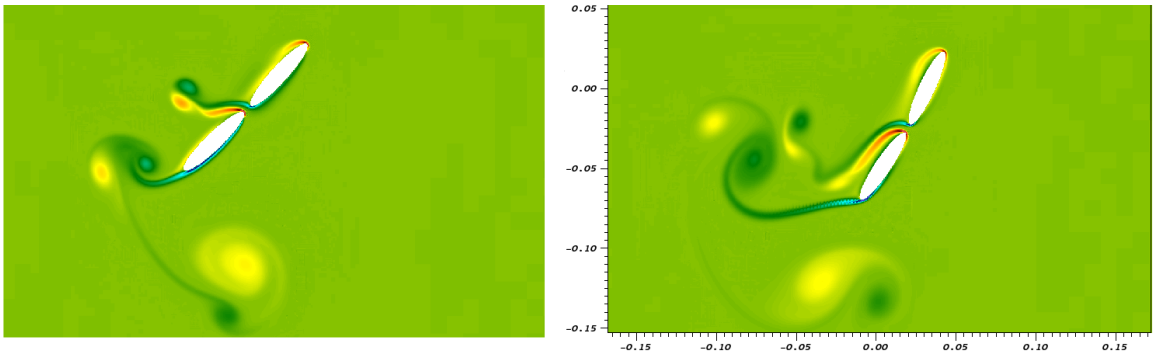


Figure 4: Computed vorticity field at t/T= 0.6 (left), 0.8 (right).

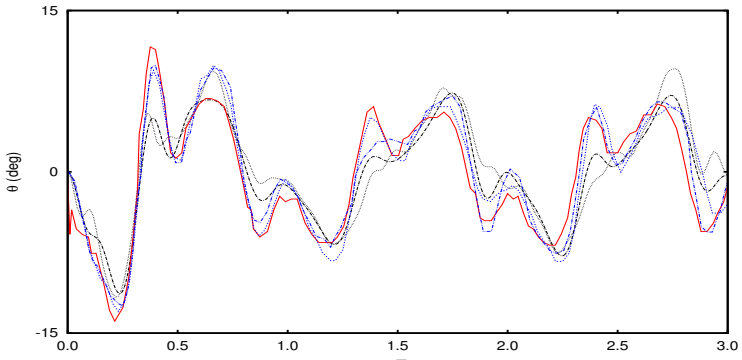


Figure 5: Hinge deflection angle over time. Experimental results (–); Toomey-Eldredge Re=100 (- -), Re=500 (–··); Current Re=100 (- · -), Re=500 (· · ·).

[11] and our simulations through three periods of motion. The dimensional torsion spring and damper coefficients utilized are $K^* = 5.5 \times 10^{-2}\,\mathrm{kg\,m^2/s^2}$ and $R^* = 6.0 \times 10^{-4}\,\mathrm{kg\,m/s^2}$ respectively. These values differ from those used in the viscous vortex particle method (VVPM) simulations by Toomey & Eldredge, however, the wing behavior and predicted loads are comparable and in good agreement with the experiments in [6, 11]. Note that although a no-slip boundary condition is applied at the wing surface, no interfacial shear forces are considered yet in this work. It follows that the fluid loads, particularly the moment, differ between this work and the VVPM used in [11], as shown in Table 2, while the hinge deflection presented in Fig. 5 is in good agreement. Without considering structural loading from shear forces yet, this work predicts the peak forces within 15% of those simulated by the VVPM and reveals the rotation of the driven body influencing hinge deflection with the same periodic trend as the experiments conducted by Toomey & Eldredge [6, 11].

6 Simulation of wakes behind wind turbines

Utilizing the developed LBM solver for moving geometries, we have carried out a simulation campaign to test the suitability of the overall approach to simulate the flow

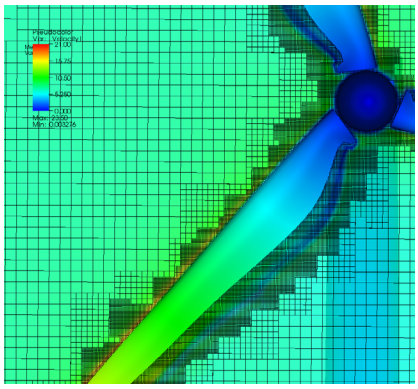


Figure 6: Snapshot of SAMR Mesh in rotor midplane.

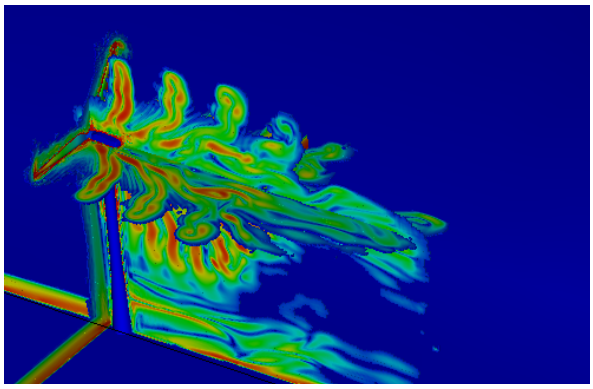


Figure 7: Emerging vortex field at $t \approx 9.3\,\mathrm{s}$. Rotation was started at $t = 4\,\mathrm{s}$.

fields created by wind turbines. We have built a tailored flexible surface mesh model of a Vestas V27 turbine. The V27 has a rotor diameter of $D = 27\,\mathrm{m}$, a tower height of $\sim$ 35 m and achieves its maximal energy output of 225 kW at wind velocities from 14 to 25 m/s. A prototypical ground topology is also included into the surface mesh model that represents the time-dependent geometry with $\sim 23,300$ facets per turbine. It is assumed that the inflow wind direction is always in direction of the turbine middle axis and the pitch blade angle is at 0 degree.

In the first computation, the wind velocity is $u_1 = 7\,\mathrm{m/s}$ with a boundary layer profile of 5 m height assumed near the ground. A simulation domain of extensions $200\,\mathrm{m} \times 100\,\mathrm{m} \times 100\,\mathrm{m}$ is used and discretized with a base grid of $400 \times 200 \times 200$ cells. Dynamic refinement with three additional levels refined by the factors $r_{1,2} = 2$ and $r_3 = 4$ is applied. The two highest levels are reserved to refine the moving surface mesh of rotor and tower at a geometry resolution of $\Delta x = 3.125\,\mathrm{cm}$; level 1 is used to dynamically adapt to the wake region using an empiric error estimation criterion on $|\mathbf{u}|$. Figure 6 shows a snapshot of the Cartesian cells in the rotor midplane and the moving structure colored by the length of the prescribed velocity vector. The computation is run for 141,344 highest level iterations to a final time of $t_e = 30\,\mathrm{s}$, where a constant rotation with 15 rpm (corresponding to a power generation of $\sim 52.5\,\mathrm{kW}$) is started at $t = 4\,\mathrm{s}$.

Figure 7 displays by color the length of the vorticity vector in two 2D planes shortly after starting the rotation. The creation of an emerging helical main vortex structure in the downwind direction emanating from the blade tips can be inferred. This computation was run on 96 cores of a cluster with Intel Xeon-Westmere CPUs and required $\sim 10,400\,\mathrm{h}$ CPU in total, i.e., 108 h wall clock time.

The second test setup considers three V27 turbines and corresponds to the U.S. Department of Energy's Scaled Wind Farm Technology (SWIFT) facility. Two turbines are positioned $3D$ apart in the wind direction; the third turbine is placed $5D$ downstream exactly in the rotor middle axis of the first one. This allows direct comparison of the wake field between two interacting turbines with an undisturbed one. The computational

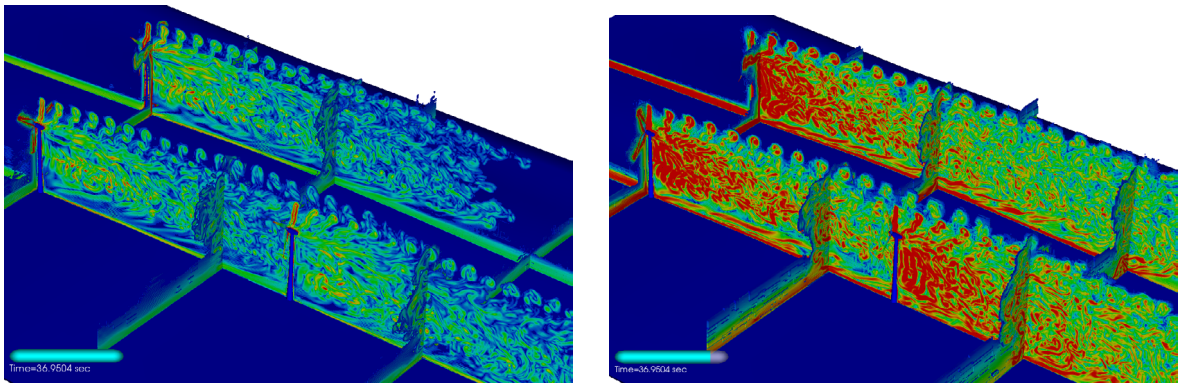


Figure 8: 2D Color planes depict the length of the vorticity vector for 7 m/s (left) and 25 m/s (right) inflow velocity at $t \approx 37$ s.

setup is similar to the previous simulation, where a domain of 488 m × 240 m × 100 m is discretized with a base resolution of 448 × 240 × 100 cells and again refined isotropically by the factors 2,2,4 (resolution near the rotors is $\Delta x = 6.25$ cm). Dynamic refinement of the wake field is permitted now up to level 2, yielding again a resolution in the wake of $\Delta x = 25$ cm.

Two configurations are compared: the case with $u_1 = 7$ m/s inflow velocity and all turbines operating at 15 rpm and a simulation with $u_1 = 25$ m/s and 43 rpm, which corresponds to the maximally allowed rotation rate under normal operations. 94,224 highest level iterations to a final time of $t_e = 40$ s are computed. Figure 8 depicts the wake fields for both cases after $t \approx 37$ s simulated time. The color coding in both graphics uses the same scale and it is obvious that vorticity production is considerably increased in the second configuration. In both simulations, the radius of the main vortices are only sightly increasing and vorticity seems overall exceptionally well preserved. A strong influence of the tower on the wake field is apparent and also the difference between the wake fields behind the isolated turbine versus the two turbines aligned in the wind direction is striking. A three-dimensional visualization of the wake field for the second configuration in its quasi-steady state in Fig. 9 reveals this difference especially well. Because of the higher wind velocity the number of vortex rotations reduces in the second configuration from ~ 15 to ~ 12. In both simulations, incident pressure and velocity on the downstream turbine are reduced (not specially shown). Its rotation apparently induces a less pronounced helical vortex than the upstream turbines, as can be seen in Fig. 9.

Level	Grids	Cells
0	3,234	10,752,000
1	11,921	21,020,256
2	66,974	102,918,568
3	896	5,116,992

Table 3: Grids and cells at t_e.

The used SAMR grids and total number of cells on each level for the second configuration at t_e are given in Table 3. Note that a corresponding uniform mesh would require $44 \cdot 10^9$ cells and take four times more time steps than level 2, which contains the majority of cells. A benchmark run on 288 cores on a cluster of

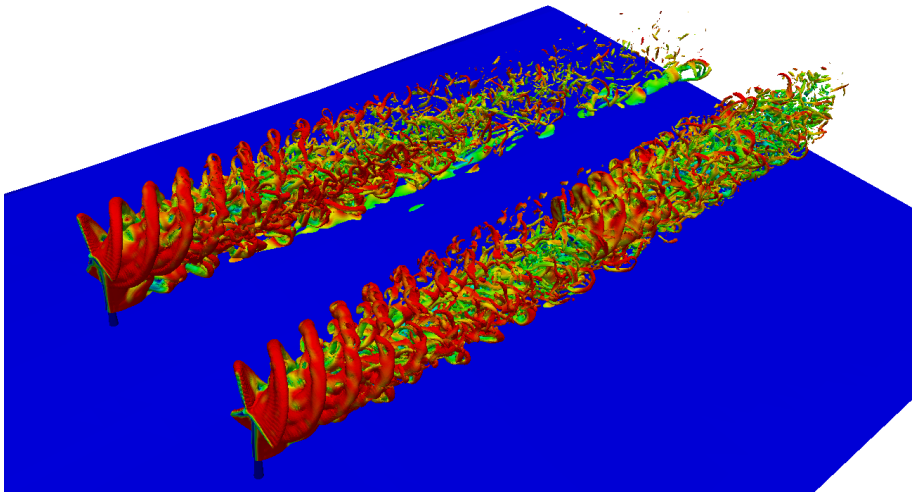


Figure 9: Quasi-steady three-dimensional wake field for the configuration with $u_1 = 25\,\mathrm{m/s}$ and 43 rpm visualized by coloring an iso-surface based on the length of the vorticity vector by the local total velocity.

Intel-Ivybridge CPUs required just 38.5 h wall clock time (11,090 h CPU) to advance the coupled, adaptive computation further from $t_e = 40$ to 50 s, which gives evidence for the practical feasibility of carrying out these computations on compute clusters of moderate size.

7 Conclusions

The first prototype of a dynamically adaptive, three-dimensional lattice-Boltzmann method for simulating coupled wind engineering problems has been developed. The results show that the overall approach can easily handle large structural motion and high coupling frequencies, while the low dissipation properties of the lattice Boltzmann scheme in combination with dynamic mesh adaptation are able to resolve even complex vortex structures at moderate computational costs. First coupled validation has been achieved for a canonical FSI problem from [11]. As a first verification for wind turbine simulation, we have demonstrated that our approach is able to simulate the propagation of wake fields created by the prescribed rotation of accurate Vestas V27 wind turbine rotor geometry, including the interaction with the tower, with apparent good quality and reasonable compute time. Immediate future work will concentrate on fully validating the approach, considering also the exchange of viscous shear forces beside hydrodynamic pressures, for available laboratory benchmarks, e.g., the Mexico wind tunnel experiments [10] and then to incorporate the dynamic elastic response of the blades into the turbine model.

Acknowledgments

Stephen L. Wood was supported by the TN-SCORE Energy Scholar program funded by NSF EPS-1004083 during this work.

REFERENCES

[1] M. Berger and P. Colella. Local adaptive mesh refinement for shock hydrodynamics. *J. Comput. Phys.*, 82:64–84, 1988.

[2] H. Chen, O. Filippova, J. Hoch, K. Molvig, R. Shock, C. Teixeira, and R. Zhang. Grid refinement in lattice Boltzmann methods based on volumetric formulation. *Physica A*, 362:158–167, 2006.

[3] R. Deiterding. A parallel adaptive method for simulating shock-induced combustion with detailed chemical kinetics in complex domains. *Computers & Structures*, 87:769–783, 2009.

[4] R. Deiterding. Block-structured adaptive mesh refinement - theory, implementation and application. *European Series in Applied and Industrial Mathematics: Proceedings*, 34:97–150, 2011.

[5] R. Deiterding and S. L. Wood. Parallel adaptive fluid-structure interaction simulations of explosions impacting building structures. *Computers & Fluids*, 88:719–729, 2013.

[6] J. D. Eldredge, J. Toomey, and A. Medina. On the roles of chord-wise flexibility in a flapping wing with hovering kinematics. *Journal of Fluid Mechanics*, 659:94–115, 9 2010.

[7] D. Hähnel. *Molekulare Gasdynamik*. Springer, 2004.

[8] S. Hou, J. Sterling, S. Chen, and G. D. Doolen. A lattice Boltzmann subgrid model for high Reynolds number flows. In A. T. Lawniczak and R. Kapral, editors, *Pattern formation and lattice gas automata*, volume 6, pages 151–166. Fields Inst Comm, 1996.

[9] S. P. Mauch. *Efficient Algorithms for Solving Static Hamilton-Jacobi Equations*. PhD thesis, California Institute of Technology, 2003.

[10] J. G. Schepers and K. Boorsma. Final report of iea task 29: Mexnext (phase 1) – Analysis of Mexico wind tunnel measurements. Technical Report ECN-E-12-004, European research Centre of the Netherlands, 2012.

[11] J. Toomey and J. D. Eldredge. Numerical and experimental study of the fluid dynamics of a flapping wing with low order flexibility. *Physics of Fluids (1994-present)*, 20(7):–, 2008.

[12] L. Tsai. *Robot Analysis: The Mechanics of Serial and Parallel Manipulators*. Wiley, 1999.

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis (Eds)

A HYBRID LES/CAA METHOD APPLIED TO A 3D SHEAR FLOW SIMULATION

O. LABBÉ, C. PEYRET

ONERA - The French Aerospace Lab
F-92322 Châtillon, France
email: odile.labbe@onera.fr, www.onera.fr
email: christophe.peyret@onera.fr, www.onera.fr

Key words: Shear flow, Finite Volume method, Discontinuous Galerkin Method, CWIPI, LES/CAA

Abstract. An aero-acoustic coupling method based on the splitting into noise sources generation and acoustic propagation in separate physical domains is presented in this paper. The key idea is to limit, as much as possible, the CFD domain to the noise generation region and to accurately propagate the acoustic waves with a CAA solver. The approach presented here is a domain decomposition method based on the coupling of different equations, methods and meshes, which allows a simulation of both flow and acoustics in one single coupled calculation suitable for far field predictions with reflecting bodies through a coupling boundary.

1 INTRODUCTION

The numerical simulation of both generation and propagation of acoustic waves in one computation is actually a challenge due to the spatio-temporal scale disparities of the problem. The full time-dependent, compressible Navier-Stokes equations describe both aerodynamic and acoustic phenomena, but require a detailed numerical compressible flow simulation, using a grid fine enough to minimize the introduction of sound propagation errors. A large number of grid elements would be necessary to resolve all the scales accurately in an unsteady simulation. Moreover, the handling of long-distance sound propagation remains difficult with usual CFD solvers due to the numerical damping and dissipation. On the other hand, a single approach contains naturally the interaction of the acoustic perturbations with the flow field and with embedded geometries. The dissipation problems generally lead to carry out hybrid methods, in which the computational domain is split into different regions, such that the governing flow field (source region) or acoustic field (acoustic region) can be solved with different equations, numerical techniques and computational grids. Various hybrid methodologies exist, differing from each other in the type of applied propagation equations or in the way the coupling between source and propagation regions is made. Bailly and Bogey [1] proposed a review of the progress in the computational aeroacoustics field and discussed connections between CFD and CAA using hybrid approaches. The coupling methods commonly used for hybrid CFD/CAA applications are divided into two categories: one based on equivalent source formulations and the other based on an acoustic continuation

of source region simulation. For the first one, once the sound source is predicted, the approach to describe its propagation is the extension of near-field CFD results to the acoustic far-field with surface or volume integral methods [2]-[6]. The second hybrid CFD/CAA approach solves the Acoustic Perturbation Equations [7]-[9] or the Linearized Euler Equations [10], [11] to extend the CFD solutions to the far-field. These propagation solvers are generally high order accurate but necessitate a mean flow definition. Generally these hybrid methods do not take into account any acoustic feedback, except for the domain decomposition performed by [12], [13], where the coupling approach connects different classes of methods on structured and unstructured grids for the solution of Navier-Stokes, Euler and linearized Euler equations. The approach presented in this paper is also a domain decomposition method based on the coupling of different equations, grids and methods, which allows a simulation of both flow and acoustics in one single calculation suitable for far field predictions with reflecting bodies. The CFD domain solving the Navier-Stokes equations is reduced to the region of viscous effects and initial turbulence development which generally accounts for a small part of the flow. The acoustic propagation is solved with the full non-linear Euler equations with a coupling boundary located in the turbulent flow. The acoustic solver is based on high order Discontinuous Galerkin schemes [14], which offer a high accuracy, low dispersion and low dissipation. This class of solvers is able to accurately propagate waves over large distances and allows using unstructured grids, which present significant advantages such as a highly flexible refinement even in complex geometries.

The coupling method presented in this paper consists in separating the whole computational domain into two complementary parts. The first domain is assumed to contain the region where turbulence develops and its size is reduced as much as possible. The flow in this domain is solved using the Navier-Stokes equations. The second domain is devoted to the propagation of the perturbations generated in the first one, in a flow which is not necessary uniform and may contain reflecting bodies. It may also include noise production, in which viscous effects can be negligible. The problem is solved using the full Euler equations. Such LES/CAA couplings pose, a priori, the problem of the continuity of the solution on both sides of the exchange surface. Indeed, the LES must be performed using sufficiently fine meshes and time steps small enough to correctly simulate the development of turbulence. On the other hand, to be efficient, the CAA must be carried out using meshes and time steps adapted to the acoustic scales which are generally much larger than the turbulence scales. So, in this approach the spatial scales will be very different on both sides of the coupling boundary, but the time step will be identical.

In a previous study [15], this splitting method was applied to a hot jet simulation, that only dealt with a one way coupling from LES to CAA and thus mainly focused on the acoustic purpose. Here, this splitting method is applied to a 3D shear flow. The computational domain is a 3D cylinder, which includes a smaller one, and can be considered as a model of jet/wing interactions. A shear flow (100m/s-50m/s) is imposed at the inflow of the small cylinder and a rigid wall conditions is imposed at the external boundary of the big cylinder, which induces wave reflections impacting on the shear flow. In this approach, both the LES and CAA solvers are part of the same simulation framework. While the LES solver is based on a finite-volume method for the prediction of the flow field, the CAA approach is based on a high-

order discontinuous Galerkin solver for the acoustic field. The exchanges between both solvers are made through the MPI communications.

2 NUMERICAL METHODS

2.1 Finite-volume method for the flow simulation

The aerodynamic solver FUNk [16] developed at ONERA, is based on the unsteady compressible Navier-Stokes equations expressed in conservative form. The LES equations are obtained using Favre filtering and the filtered equations are closed by means of a subgrid scale viscosity and the Prandtl analogy. The model used to compute the subgrid viscosity is the selective mixed scale model introduced for compressible flows by Lenormand et al. [17]. The spatial discretization method is based on the cell-centered Finite Volume methodology (FVM) on structured grid. An upwind biased scheme, with a third-order MUSCL interpolation scheme of AUSM+(P) family without any shock capturing feature is used for the convective terms. A second-order-accurate centered scheme is used for viscous fluxes. The time integration is carried out by means of a third-order compact Runge-Kutta scheme. The solver has been extensively validated and used for various flow problems.

2.2 Discontinuous Galerkin approximation

The CAA computation is carried out with the SPACE solver developed at ONERA [18], [19] which solves the Euler equations using a nodal Discontinuous Galerkin Method. A nodal DGM with an optimized Lagrangian polynomials basis of order N+1 is used to solve the conservative form of Euler equations:

$$\partial_t w + \nabla . F(w) = 0 \tag{1}$$

Considering a domain Ω with an envelope Γ, the derivation of the DG method starts with the weak form of the equation. Therefore, Eq. 1 is multiplied by a test function Ψ and integrated over the domain Ω. A variational formulation of Euler equations reads:

$$\left\{ \forall \Psi \in V(\Omega); \quad \int_{\Omega} (\partial_t w + \nabla . F(w))\ \Psi d\Omega = 0 \right\} \tag{2}$$

Subdividing the domain $\Omega = \bigcup_i \Omega_i$ and $\Gamma = \partial\Omega = \bigcup_j \Gamma_j$ and using integration by parts on the flux term, we obtain the weak formulations of the differential equations:

$$\left\{ \forall \Psi \in V(\Omega); \sum_i \int_{\Omega_i} \partial_t w \Psi d\Omega - \int_{\Omega_i} F(w).\nabla \Psi d\Omega + \oint_{\partial\Omega_i | \Gamma} \widetilde{F}(w, w^+) n \Psi d\Gamma + \sum_j \oint_{\Gamma_j} F(w) n \Psi d\Gamma = 0 \right\} \tag{3}$$

where n is the surface normal vector in the reference system and $\partial\Omega_i$ is the envelope of geometric cell Ω_i. On the element boundaries $\partial\Omega_{/\Gamma}$, a numerical flux combines values w and $w+$ from both sides to a single flux. The Local Lax-Friedrichs Numerical Fluxes have been chose:

$$\widetilde{F}(w,w^+)n = \tfrac{1}{2}(F(w)+F(w^+))n - \alpha\lambda_{\max}(w^+ - w) \qquad \text{where } \alpha \in [0,1]$$

and $\lambda_{\max} = \max(|v.n|, |v.n+c|, |v.n-c|, |v^+.n^+|, |v^+.n^+ + c^+|, |v^+.n^+ - c^+|)$

with $c = \sqrt{\gamma p/\rho}$

The solution *w* is approximated using a polynomial basis N

$$w(\xi,t) \approx \sum_{i,j,k}^{N} \overline{w_{ijk}}(t)\varphi_{ijk}(\xi_{ijk}) \qquad \varphi_{ijk}(\xi) = l_i(\xi_i)l_j(\xi_j)l_k(\xi_k)$$

where the basis functions φ_{ijk} are the product of one-dimensional Lagrange polynomials l_i of degree N in each spatial direction and $\overline{w_{ijk}}(t)$) are the sampling of solutions at (x_i, y_j, z_k). The nodal basis is defined with a set of interpolation points $\{\xi_{i=0}^{N}\}$ on the interval $\xi \in [-1, 1]$, which in this work are the Gauss-Lobatto nodes. The fluxes are approximated using the same approach.

The integrals in Eq. 3 are approximated by Gauss-Lobatto quadrature. Generally, Gauss quadrature of an arbitrary function f (x) on the interval [−1, 1], with N+1 nodes can be written as:

$$\int_{-1}^{1} f(x)dx = \sum_{i=0}^{N} \omega_i f(x_i) = \sum_{i=0}^{N} \omega_i \sum_{i=0}^{N} l_j(x_i) f_j = \sum_{i=0}^{N} \omega_i \, f_i ,$$

because $l_j(x_i) = \delta_{ij}$, where the weights ω_i and the integration nodes x_i are specific to the chosen quadrature. These weights are pre-calculated and stored to make the algorithm efficient. With the interpolation points $\{\xi_i\}$ collocated at the Gauss nodes, all sums collapse into single values.

In the next step, the variational formulations are integrated in time to obtain the solution at the next time step, for which a low-storage three-order Runge-Kutta scheme is used, the same as the one used in the CFD solver.

The coupling from LES to CAA is achieved by a numerical integration along the coupling surface Γ_c :

$$\int_{\Gamma_c} F(w)n\,\Psi d\Gamma = 0$$

For each CAA cell along the coupling boundary, the integral is computed using the values of LES solution interpolated on Gauss points.

2.3 Coupling procedure

The CWIPI [20], [21] library aims at providing a fully parallel communication layer for mesh based coupling between several parallel codes with MPI communications. CWIPI is a static tool in the sense that all the components of the simulations are started at the beginning, exchange data during the run phase and finish together at the end. The coupling is made through 1D, 2D or 3D exchange zones that can be discretized in different ways in the coupled

codes. The library takes into account all types of geometrical elements (polygon, polyhedral) with an unstructured description. CWIPI functionalities involve the construction of the communication graph between distributed geometric interfaces through geometrical localization, interpolation on non coincident meshes, exchanges of coupling fields for massively parallel applications as well as visualization files.

At the coupling boundary, the DGM solver is supplied at each stage of the Runge-Kutta method with the conservative variables from the CFD solver at the Gauss points belonging to the coupling boundary. These variables are computed with an interpolation of same order as the DGM scheme starting from the nodal values of the LES cells containing the Gauss points under consideration. And vice-versa, the CAA solver fills the ghost cells of Finite Volume method in order to apply the scheme at the coupling boundary. This coupling procedure works with structured/unstructured and non-conform meshes on both sides of the exchange surface.

3 LES/CAA COUPLING APPLIED TO 3D SHEAR FLOW SIMULATION

The simulation configuration consists of two cylinders fitted together as shown in Figure 1. The length and diameter of the external cylinder are 1m and 0.0794m, whereas the dimensions of the internal cylinder are respectively 0.06m and 0.04m. The meshes are built around a square of 14 cells × 14 cells with radial layers. There are 281 points in the axial direction and 57 points in the azimuthal direction. The internal cylinder has 225 120 elements, while the external one 1 061 760. The cylinder mesh is uniform in the axial direction with Δx =0.0005 m up to x=0 .06 and then a geometric progression is used in such a way that the cell size at the exit boundary Δx=0.02261 m.

In x=0, a velocity of 100m/s is imposed in the disc centered in 0 with a radius of 0.01m, on the remaining part of the entrance boundary the velocity is 50m/s as indicated in Figure 1, which shows the plane z=0, giving the different boundary conditions. In x=1m, a non-reflective condition is imposed at this boundary. The external envelope of the cylinder is a wall, on which is applied a slip condition. The coupling boundary is located at the interface of both cylinders. An explicit three-stage Runge-Kutta scheme is used in both solvers with a time step $\Delta t = 10^{-7}$ s. The exchanges between the solvers FUNk and SPACE occur at each stage of the RK3 scheme.

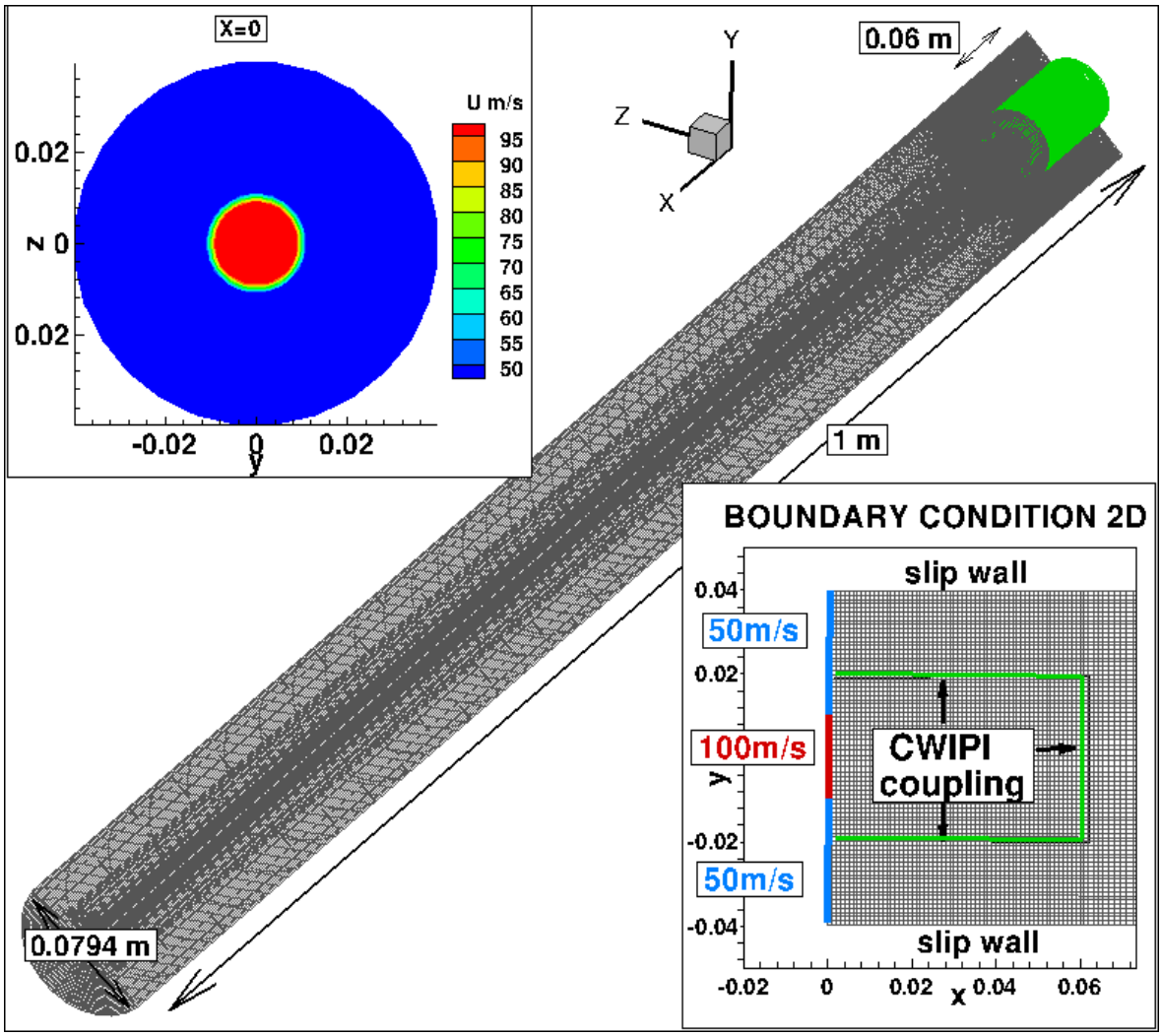


Figure 1 : Two cylinders fitted together, 2D view (z=0) of the boundary conditions and velocity at the entrance (x=0).

A first validation has been carried out with the aerodynamic solver FUNk, with and without CWIPI coupling in order to control if exchanges are correctly made. The results are identical. All the following results are plotted at t=0.001s, when the shear flow is crossing the coupling boundary at x=0.06m. In Figure 2 the pressure iso-contours are plotted in the plane z=0, the black bold line marks the coupling CWIPI boundary. An enlargement around this boundary is displayed with the iso-contours of velocity u. Moreover, the plane x=0.061m downstream to the coupling boundary shows the iso-velocity v. This case acts as reference for other simulations.

The aim of the hybrid method is first to use less points and high order in the CAA domain than in the same part of the previous CFD/CFD simulation. For that, each cell of the CAA part is divided by two in each direction. Then, the DGM method can access to high order schemes.

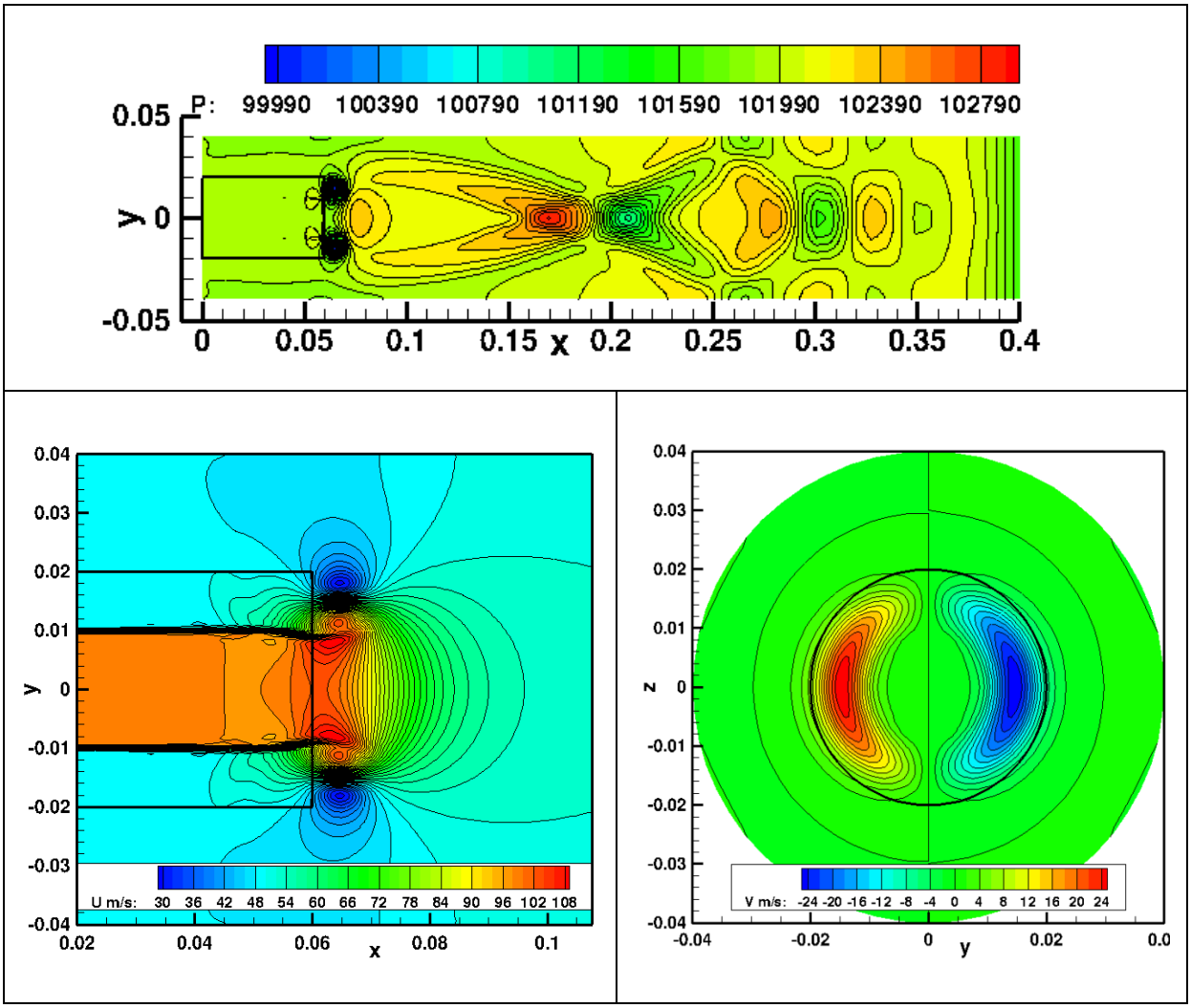


Figure 2 : FUNk/FUNk iso-contours: pressure and u-velocity (z=0), v-velocity (x=0.061).

The first FUNk/SPACE simulation uses the same mesh in the LES domain, but contains only 132 720 elements in the external cylinder, instead of 1 061 760 hexahedra in FUNk. A Q_1 Lagrangian basis is used for each SPACE hexahedra, which corresponds to approximately the same order of the LES scheme. The same iso-contours as previously are displayed in Figure 3. The results are similar to those obtained in the FUNk/FUNk simulation with much fewer points.

The second FUNk/SPACE simulation was to use a Q_2 hexahedra in the SPACE solver with the same coarse mesh, without changing anything for the FUNk solver. The results are plotted in Figure 4. The iso-pressure contours show that the pressure field is richer, especially between x=0.25 and 0.4m. That means the acoustic propagation is less numerically dissipative and dispersive, which is one of the main aims of this coupling method. The iso velocity-contours display spurious oscillations close to the coupling boundary, maybe due to the coarse mesh in this region. To solve this problem, either the use of refined hexahedra or the use of

tetrahedra in the CAA domain only in the vicinity of the coupling boundary and in the remaining part of the domain a stretched mesh is sufficient.

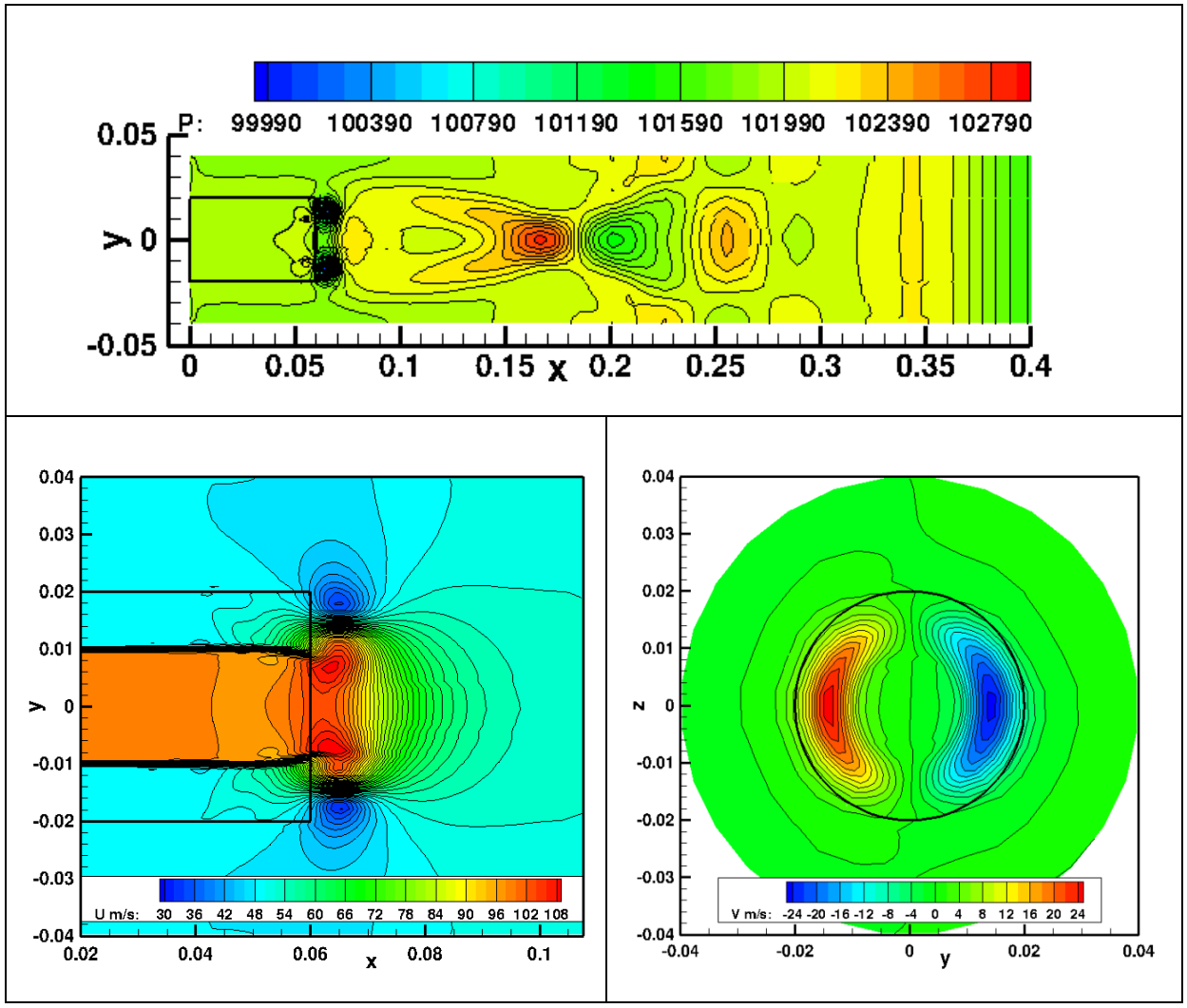


Figure 3 : FUNk/SPACE(Q_1) iso-contours: pressure and u-velocity (z=0), v-velocity (x=0.061).

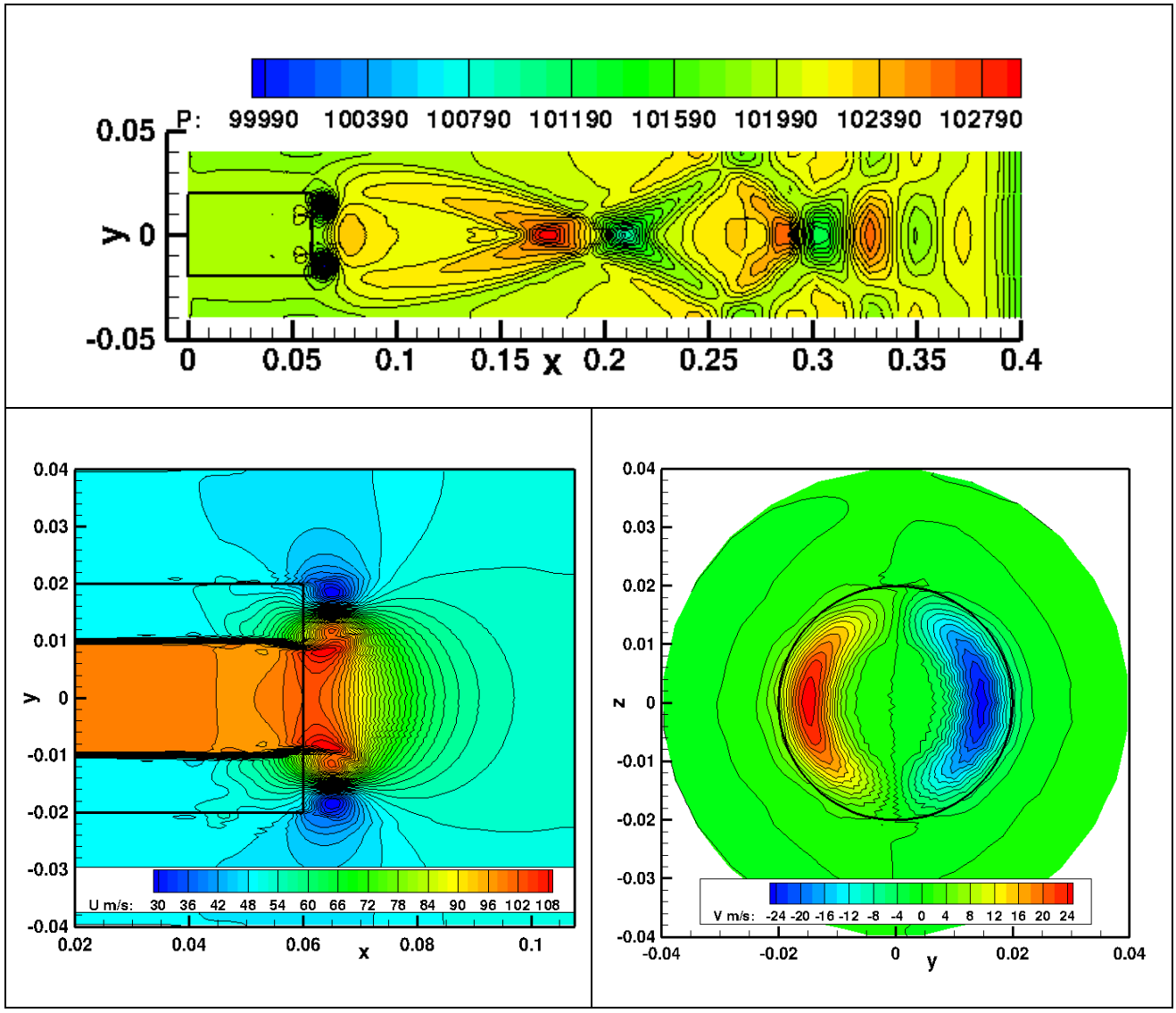


Figure 4: FUNk/SPACE (Q_2): iso-contours: pressure and u-velocity (z=0), v-velocity (x=0.061).

4 CONCLUSION

A domain decomposition approach for the coupling of aerodynamic and acoustic calculations is presented and applied to a 3D shear flow simulation. The aerodynamic domain is computed by a LES based on a Finite Volume method with a structured grid, while the acoustic one is performed using a Discontinuous Galerkin method solving the full non-linear Euler equations on an unstructured mesh. These results constitute the first validation of the hybrid method. First of all, the iso-contours obtained in the FUNk/SPACE (Q_1) are similar to

those obtained in the reference case FUNk/FUNk. So, with less (8 times) elements in the external cylinder, the same results are recovered. The first aim of the hybrid method is obtained. Secondly, with the order increase (Q_2) of the DGM, the results show that the acoustic contain is improved and the ability of the CAA method to better propagate acoustic waves compared to the LES.

The next step in this study is to exploit the DGM ability of highly flexible refinement in order to improve the exchanges at the vicinity of the coupling boundary and to use a coarse mesh elsewhere. Moreover, this hybrid method will be very useful applied to a jet simulation, in which reflecting bodies may occur in the flow field.

REFERENCES

[1] Bailly, C., and Bogey, C., Contributions of Computational Aeroacoustics to Jet Noise Research and Prediction, *International Journal of Computational Fluid Dynamics*, **18**, No. 6, (2004), 481-491.

[2] Gamet, L., and Estivalezes, J.L., Application of Large-eddy Simulations and Kirchhoff Method to Jet Noise Prediction, *AIAA Journal*, **36**, No. 12, (1998), 2170-2178.

[3] Lyrintzis, A., Surface integral methods in computational aeroacoustics -From the (CFD) near-field to the (Acoustic) far-field, *International Journal of Aeroacoustics*, **2**, No 2, (2003), 95-128.

[4] Rahier, G., Prieur, J., Vuillot, F., Lupoglazoff, N. and Biancherin, A., Investigation of integral surface formulations for acoustic post-processing of unsteady aerodynamic jet simulations, Aerospace Science and Technology, 2004, Vol. 8, pp. 453-467.

[5] Uzun, A., Lyrintzis, A., and Blaisdell, G., Coupling of Integral Acoustics Methods with LES for Jet Noise Prediction, 42nd AIAA Aerospace Sciences Meeting and Exhibit, Reno, Nevada, 2004-517.

[6] Perez, G., Prieur, J., Rahier, G., and Vuillot, F., Theoretical and Numerical Discussion on Volume Integral Methods for Jet Noise Prediction, 13th AIAA/CEAS Aeroacoustics Conference, 2007-3593, Rome, Italy.

[7] Ewert, R., and Schröder, W., Acoustic perturbation equation based on flow decomposition via source filtering, *Journal of Computational Physics*, **188**, (2003), 365-398.

[8] De Roeck, W., Baelmans, and M., Desmet, W., Aerodynamic/Acoustic Splitting Technique for Computational Aeroacoustics Applications at Low Mach Numbers, *AIAA Journal*, **46**, No. 2, (2008), 463-475.

[9] Cunha, G. and Redonnet, S., Towards a robust and accurate CFD/CAA Coupling Procedure for Hybrid Methods in Aeroacoustics, Part 1: On the Optimization of CFD/CAA Coupled Calculations, 18th AIAA/CEAS Aeroacoustics Conference, Colorado Springs, Co, AIAA 2012-2063.

[10] Viswanathan, K., and Sankar, L. N., Toward the direct calculation of noise: fluid/acoustic coupled simulation, *AIAA Journal*, **33**, No 12, (1995), 2271–2279.

[11] Bogey, C., Bailly, C., and Juvé, D., Noise investigation of a High Subsonic, Moderate Reynolds Number Jet Using a Compressible Large Eddy Simulation, *Theoretical and Computational Fluid Dynamics*, **16**, (2003), 273-297.

[12] Utzmann, J., Munz, C.-D., Dumbser, M., Sonnendrücker, E., Salmon, S., Jund, S., and

Frinod, E., Fluid-Acoustic Coupling and Wave Propagation, Num. Sim. of Turbulent Flows & Noise Generation, NNFM 104, Springer-Verlag (Ed.) (2009), 47-74.

[13] Birkefeld, A., Beck, A., Dumbser, M., Munz, C.-D., König, D., and Schröder,W., Advances in the Computational Aeroacoustics with the Discontinuous Galerkin Solver NoisSol, 16th AIAA/CEAS Aeroacoustics Conference Stockholm, Sweden, (2010), AIAA 3823.

[14] Lockard, D. P., and Atkins, H. L., Efficient Implementations of the Quadrature-Free Discontinuous Galerkin Method, 14th AIAA Conference, Norfolk, VA, (1999), AIAA 3309.

[15] Labbé, O. Peyret, C. Rahier, G. and Huet, M., A CFD/CAA coupling method applied to jet noise prediction. *Computers & Fluids*, **86** (2013) 1–13.

[16] Larchevêque, L., Sagaut, P., Mary, I., Labbé, O., and Comte, P., Large-eddy simulation of a compressible flow past a deep cavity, *Physics of Fluids*, **15**, No 1, (2003), 193-210.

[17] Lenormand, E., Sagaut, P., Phuoc, T. and Comte, P., Subgrid-scale models for large-eddy simulation of compressible wall bounded flows, *AIAA Journal,* **38**, 8, 1340-1350, (2000).

[18] Delorme, P., Mazet, P.-A., Peyret, C., and Ventribout, Y., Computational Aeroacoustics Applications based on a Discontinuous Galerkin Method, *Comptes rendus de Mécanique*, **333**, (2005), 676-682.

[19] Léger, R., Peyret, C., and Piperno, S., Coupled Discontinuous Galerkin/Finite Difference Solver on Hybrid Meshes for Computational Aeroacoustics, *AIAA Journal*, **50**, No. 2, 338- 349, (2012).

[20] http://sites.onera.fr/cwipi/

[21] http://www.aerospacelab-journal.org/sites/www.aerospacelab-journal.org/files/AL2-11_0.pdf page 7

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis(Eds)

COMPARISON OF ADVANCED TURBULENCE MODELING APPROACHES FOR FLUID-STRUCTURE INTERACTION

A. Ali*†, T. Reimann†, D.C. Sternel† and M. Schäfer†

†Institute of Numerical Methods in Mechanical Engineering
Technische Universität Darmstadt
Dolivostraße 15, 64293 Darmstadt, Germany
e-mail: ali@fnb.tu-darmstadt.de, web page: http://fnb.tu-darmstadt.de

Key words: Fluid-Structure Interaction, Turbulence, Delayed Detached-Eddy Simulation, Large Eddy Simulation

Abstract. In this study we present the results of the benchmark of a turbulent Fluid-Structure Interaction test case. An implicit partitioned approach is employed to couple the fluid and structure subproblems. We employ three different techniques to model the turbulence in fluid motion. A 2-d unsteady Reynolds Averaged Navier-Stokes approach, with an elliptic relaxation based turbulence model ($\zeta - f$), successfully captures the oscillation mode. Further investigations are performed with a Delayed Detached Eddy Simulation and a Large Eddy Simulation model. The $\zeta - f$ model is used as a baseline unsteady Reynolds Averaged Navier-Stokes model for the Delayed Detached Eddy Simulation. A comparison of the structural deflections from the simulations show a reasonable agreement with the experiment. In light of the presented results, the suitability of the modeling approaches is discussed.

1 INTRODUCTION

Fluid-Structure Interaction (FSI) phenomena are important to study the design of many engineering applications. Experiments for most of the real-world FSI applications are not feasible – measurement techniques can not look inside all the important parameters, and are too expensive. With increasing computational power, simulation of such multi physics scenarios are becoming feasible and can give new knowledge. The capabilities of these numerical methods to study FSI needs to be validated against benchmark test cases.

A great number of FSI applications have turbulent fluid motion, thus making it important to study the FSI phenomena in turbulent flows. The numerical and experimental studies on FSI, conducted in the last decade mostly focused on laminar flows. In an effort

to provide experimental data for validation of numerical tools, Gomes and Lienhart [6] proposed the first validation test case for FSI with incompressible turbulent fluid motion. The structural model of this test case, contains a thin flexible sheet attached to a revolvable cylinder with a rectangular mass attached to the other end of the sheet. The structure is placed inside a vertical tunnel with the flow *Reynolds Number* (R_e) of 15,000, based on the cylinder diameter. The structure exhibits a periodic oscillation close to its natural frequency. This instability mechanism is characterized as *Instability Induced Excitation* (IIE) [14] with the structure swiveling in the first mode.

Recently De Nayer and Kalmbach [4], as well as Kalmbach and Breuer [11] have proposed two different test cases with turbulent fluid motion. The proposed benchmark test cases offer a simpler structure geometry and provide a different excitation mechanism and motion mode, which is not present in [6]. The R_e based on the cylinder diameter is in sub-critical regime [15] ($10^3 < R_e < 2 \times 10^5$). This flow configuration is considered challenging for the turbulence models, since the boundary layer is laminar and transition to turbulence occurs in the separated shear layers and the wake.

The present investigation aims to access the capabilities of different turbulence modeling techniques for coupled FSI problems. The turbulent test case presented in [6] is simulated employing three different turbulence modeling techniques. The $\zeta - f$ model proposed in [8] is utilized to perform the study in 2-d *Unsteady Reynolds Averaged Navier-Stokes* (URANS) flow simulation. The Smagorinsky model [18] with dynamic procedure suggested by Germano [5] is used to carry out a *Large Eddy Simulation* (LES). A hybrid URANS/LES based on the *Detached Eddy Simulation model* (DDES) [19] with $\zeta - f$ as baseline URANS model, is also tested. The DDES formulation of $\zeta - f$ model has been discussed and verified in [24]. The structural deflections, swiveling frequency and the end mass phase delay from simulations are compared with the experimental data.

2 GOVERNING EQUATIONS

For the fluid subdomain Ω_f, the fluid is assumed to be Newtonian with incompressible fluid motion. The basic conservation equations governing transport of mass and momentum are given as

$$\frac{\partial v_i}{\partial x_i} = 0, \tag{1}$$

$$\rho_f \frac{Dv_i}{Dt} = -\frac{\partial p}{\partial x_i} + \mu_f \frac{\partial^2 v_i}{\partial x_j^2} + \rho_f f_i, \tag{2}$$

where v_i is the velocity vector, p is the static pressure, μ_f is the dynamic fluid viscosity, ρ_f is the fluid density and f_i represent the external force vector.

For the structure subdomain Ω_s, we define a material point X in the reference configuration. The function χ represents the transformation from X to x as

$$x_i = \chi(X_j, t), \tag{3}$$

where X_j are the components of position vector X and x_i represents the components of current spatial position vector x. The displacements are then defined as

$$u_i = x_i - X_i. \tag{4}$$

The basic equation of momentum balance for solid domain Ω_s is written as

$$\rho_s \frac{\partial^2 \chi(X_j, t)}{\partial t^2} = \frac{\partial S_{ji} F_{ij}}{\partial X_j} + \rho_s f_i, \tag{5}$$

where S_{ji} is the second Piola-Kirchhoff stress tensor, ρ_s is the density of solid material and f_s represents the external forces on the solid. $F_{ij} = \partial x_i / \partial X_j$ is the deformation gradient. In the present study, the material is modeled utilizing a simple hyper-elastic material model, the Saint Venant-Kirchhoff law (for details see [16, 23]). For the second Piola-Kirchhoff stress tensor the model states

$$S_{ij} = \lambda_s E_{kk} \delta_{ij} + 2\mu_s E_{ij}, \tag{6}$$

where the Green-Lagrange strain tensor is represented as

$$E_{ij} = \frac{1}{2} \left(F_{ki} f_{kj} - \delta_{ij} \right), \tag{7}$$

with λ_s and μ_s as Lamé constants.

The two subproblems are coupled at the boundary with suitable interface and boundary conditions. The standard boundary conditions apply on the fluid boundaries Γ_f and the structure boundaries Γ_s. The following conditions on velocities and stresses are applied at the fluid-structure interface

$$\left(v_i^f \right)_{\Gamma_f \cap \Gamma_s} = \dot{u}_i^b \quad \text{and} \quad \left(\sigma_{ij}^s \right)_{\Gamma_f \cap \Gamma_s} = \left(\sigma_{ij}^f \right)_{\Gamma_f \cap \Gamma_s}, \tag{8}$$

where $\dot{u}_i^b$ is the velocity of the interface and σ_{ij}^s and σ_{ij}^f represent the Cauchy stress tensor of the solid and the fluid domain, respectively.

3 MODELING APPROACH

This section gives a brief description of turbulence modeling approaches in this study.

3.1 $\zeta - f$ Model

The $\zeta - f$ model proposed in [8] is a linear eddy-viscosity model. The model is capable of predicting the anisotropic behavior of turbulence near walls by evaluating the eddy-viscosity ν_t based on the wall normal velocity scale ratio ζ as $\nu_t = C_\mu \zeta k \tau$, where $\zeta = \overline{v_2'^2}/k$ and k is the kinetic energy of turbulence. The constitutive model equations are given as

$$\frac{Dk}{Dt} = \mathcal{P} - \epsilon + \frac{\partial}{\partial x_i}\left[\left(\nu + \frac{\nu_t}{\sigma_k}\right)\frac{\partial k}{\partial x_i}\right], \tag{9}$$

$$\frac{D\epsilon}{Dt} = \frac{C_{\epsilon 1}\mathcal{P} - C_{\epsilon 2}\epsilon}{\tau} + \frac{\partial}{\partial x_i}\left[\left(\nu + \frac{\nu_t}{\sigma_\epsilon}\right)\frac{\partial \epsilon}{\partial x_i}\right], \tag{10}$$

$$\frac{D\zeta}{Dt} = f - \frac{\zeta}{k}\mathcal{P} + \frac{\partial}{\partial x_i}\left[\left(\nu + \frac{\nu_t}{\sigma_\zeta}\right)\frac{\partial \zeta}{\partial x_i}\right], \tag{11}$$

$$L^2\nabla f - f = \frac{1}{\tau}\left(c_1 + C_2{}'\frac{\mathcal{P}}{\epsilon}\right)\left(\zeta - \frac{2}{3}\right), \tag{12}$$

where ϵ is the dissipation of turbulent kinetic energy, $\mathcal{P}$ is the production of turbulent kinetic energy and f is the elliptic relaxation term that models the pressure-velocity correlations. L and τ are the length and time scales of turbulence, whereas other unknown terms in the given set of equations are model constants. For a detailed model description see [8].

3.2 Dynamic Smagorinsky Model

The LES simulation in this study is performed by applying the Smagorinsky model [18] to estimate the *Sub-Grid Scale* (SGS) turbulent viscosity ν_{SGS} as

$$\nu_{SGS} = C_s\Delta^2|\overline{S}|, \tag{13}$$

where $|\overline{S}| = (2\overline{S_{ij}S_{ij}})^{1/2}$ is the magnitude of strain-rate tensor $\overline{S_{ij}} = \partial\overline{v_i}/\partial x_j + \partial\overline{v_j}/\partial x_i$, C_s is the model constant and $\Delta = (\Delta_1\Delta_2\Delta_3)^{1/3}$ is the filter width, with Δ_i representing filter width in each spatial direction. An overbar on $\overline{S_{ij}}$ represents a filtered quantity.

The model constant $C_s = C_s(x,t)$ is calculated dynamically as proposed by Germano et al. [5]. The resulting equation system to estimate C_s is solved using least squares method as suggested by Lilly [12]. The C_s values are clipped as $C_s(x,t) = \max\{C_s(x,t), 0\}$ to avoid negative values of C_s.

3.3 $\zeta - f$ DDES

The *Detached Eddy Simulation* (DES) concept first proposed by Spalart [20], is to combine URANS and LES to have a model with better prediction of turbulence than a URANS and computationally less expensive than an LES. In the DES approach, the URANS model is modified to achieve a SGS model in regions where grid is fine enough for an LES. The switching between two modes is based on the length scale l_{turb} as

$$l_{turb} = \min(l_{RANS}, C_{DES}\Delta), \tag{14}$$

where $\Delta = \max(\Delta_1, \Delta_2, \Delta_3)$ and C_{DES} is a model constant. l_{turb} is introduced in the URANS model by modifying the dissipation term in the transport equation (9) for the

turbulent kinetic energy as $\epsilon = k^{3/2}/l_{turb}$. l_{turb} either becomes the original URANS length scale ($l_{RANS} < C_{DES}\Delta$) or the SGS length scale ($l_{RANS} > C_{DES}\Delta$).

The Delayed Detached Eddy Simulation (DDES) model was proposed in [19] as an improvement for some deficiencies of the original DES model. The modification and verification of $\zeta - f$ model to perform a DDES are presented in [24]. For DDES, the DES length scale is modified to incorporate a shielding function to preserve the URANS mode in boundary layers, because of the grid clustering near the boundaries $C_{DES}\Delta < l_{RANS}$. A quantity r_d is defined as

$$r_d = \frac{\nu_t + \nu}{\sqrt{v_{i,j}v_{i,j}}\kappa^2 d^2}, \tag{15}$$

where ν is the molecular viscosity, $v_{i,j}$ are the velocity gradients, κ is the Kármán constant and d is the wall distance. The shielding function f_d is defined as a function of r_d as

$$f_d = 1 - \tanh([8r_d]^3). \tag{16}$$

The function f_d is designed to be 0, to prevent activation of LES mode in boundary layer regions. The new length scale for DDES is then defined as

$$l_{DDES} = l_{RANS} - f_d \max(0, d - C_{DES}\Delta\psi), \tag{17}$$

where ψ is the term added to eliminate the influence of low-R_e turbulence models in the SGS mode, which is given as

$$\psi = \sqrt{\left(\frac{C_{\epsilon 1}}{C_{\epsilon 2}C_\mu \zeta}\right)^{3/2}}. \tag{18}$$

4 EXPERIMENTAL SETUP

The structural model for this test case consists of a flexible stainless steel sheet of 0.4 mm thickness, with density $\rho_{flexible\ sheet} = 7855$ kg/m^3 and Young's modulus $E_{flexible\ sheet} = 2\times10^{11}$ N/m^2. The flexible sheet is attached to a revolvable circular cylinder of aluminum. A rectangular stainless steel mass is attached to the other end of the sheet. The physical

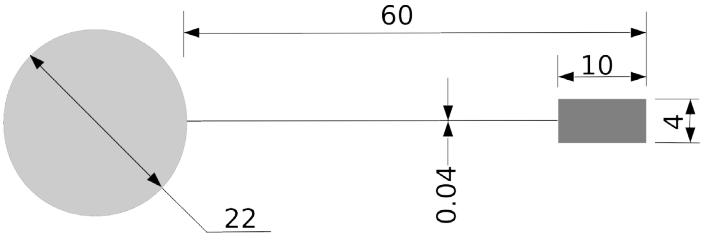


Figure 1: Structural model and dimensions in mm.

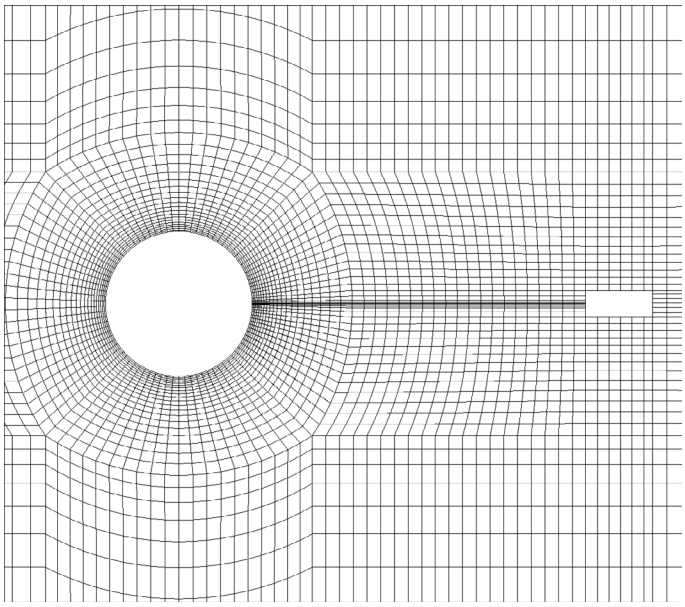


Figure 2: CFD grid around structure.

dimensions and shape of the structure are illustrated in Figure 1. Both, the front cylinder and the rectangular mass, can be considered rigid. The structure is placed inside a vertical tunnel with a test section cross-sectional area of 180 mm $\times$ 240 mm and a length of 338 mm. The fluid is water at 25°C, with kinematic fluid viscosity $\nu_{fluid} = 0.97 \times 10^{-6}$ m^2/s and density $\rho_{fluid} = 998$ kg/m^3. The bulk fluid velocity at the inlet is 0.68 m/s. The structure exposed to the incoming flow velocity oscillates around a mean position, and the flexible sheet deflects in the first mode. The detailed experimental setup, the measurement techniques, and the physical properties of the structure are described in [6].

5 COMPUTATIONAL APPROACH

The structural subproblem is solved employing the finite-element solver FEAP [21]. For the fluid domain, the finite-volume solver FASTEST [13] with block-structured body-fitted grids is used. The parallelization in FASTEST is achieved with domain decomposition and communication via MPI. The data transfer between the two codes and the interpolation on non-matching grid interfaces is performed via the interface coupling code MpCCI [9]. For details concerning the coupling algorithm see [17]. Both, the structural and fluid solver employ fully implicit second-order temporal discretizations. For the spatial discretization, the structural solver involves hexahedra elements with enhanced-strain formulation, whereas the fluid solver utilizes a second-order MUSCL [25] for the 2-d URANS simulation, second-order *Central Differencing Scheme* (CDS) for the LES and a blending between CDS and the GAMMA scheme [10] for the DDES. The blending between two schemes is done via a function introduced in [22], which ensures the calculation of fluxes with the CDS in LES regions and with the GAMMA scheme in RANS regions of

the DDES.

Table 1: Number of CV and time step size

Abbreviation	Turbulence model	No. of CVs	Time step Δt[s]	Averaged no. of periods
Sim-1	$\zeta - f$	0.37×10^6	2.5×10^{-4}	9
Sim-2	$\zeta - f$ DDES	12.0×10^6	1.5×10^{-4}	13
Sim-3	Dyn. Smag.	40.0×10^6	1.5×10^{-4}	7

Table 1 summarize the number of *Control Volumes* (CV), the time step sizes Δt, and the number of motion cycles performed for averaging of the structural deflections. Figure 2 presents a view of the grid used for 2-d URANS (Sim-1), in x-y plane around the structure, where every 4th grid-line is shown. The grids are designed to have $y^+ < 1$, for the first cell adjacent to solid walls. The time step sizes are lower bound by the artificial added mass effect [2]. The convergence of the coupled problem was observed to deteriorate, when reducing the time step size. The CFL number based on the time step sizes varied between 1.4 and 2.0. One reason for large variations in CFL number is the grid movement with structural deflections, where maximum CFL numbers are observed when the structural deflection or the velocity approaches a maximum.

6 RESULTS AND DISCUSSION

The periodic motion of the structure is adequately predicted by the simulations, whereas quantitative comparison among the simulations and the experiment is based on the averaged structural deflections. The averaging is performed in time-phase as suggested in [6], after cycle-to-cycle variations of the end mass displacements reach a minimum. The number of motion periods averaged for each simulation are listed in Table 1.

Figure 3 compares the absolute velocity contours of the experiment and Sim-1 at different phase angles. The arrangement of flow instabilities from the simulation is comparable with the experiment, despite the 2-d approach in Sim-1. Table 2 draws a quantitative comparison for the oscillation frequency of the structure f_{FSI}, the end mass phase delay ϕ_{shift} and y extrema of the end mass normalized by the cylinder diameter. The quantities in Table 2 are time-phase averaged.

The time-phase averaged cylinder rotation angle, and the end mass excursions are plotted in Figure 4a and Figure 4b, respectively. A slight asymmetry in the simulation data can be observed from Table 2 and Figure 4. This asymmetry is more noticeable in Sim-1 and Sim-3, where the motion cycles performed for averaging are less than that of Sim-2. The end mass displacement and the cylinder rotation from Sim-1 are underestimated, with y extrema of the end mass 20% lower than that of experimental values. The overdamping

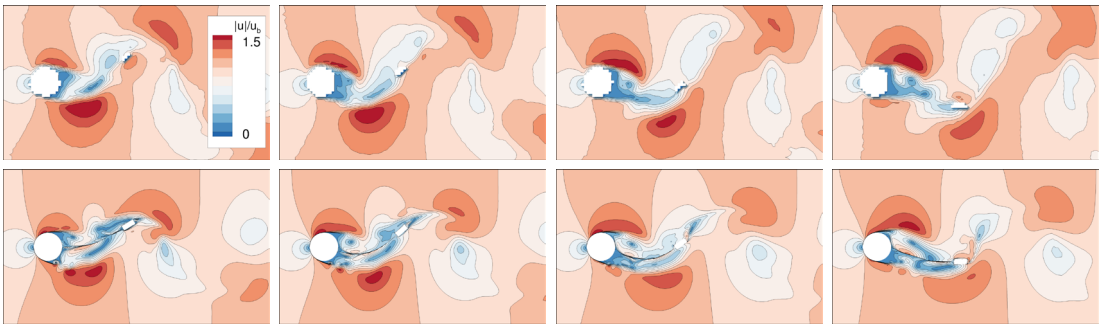


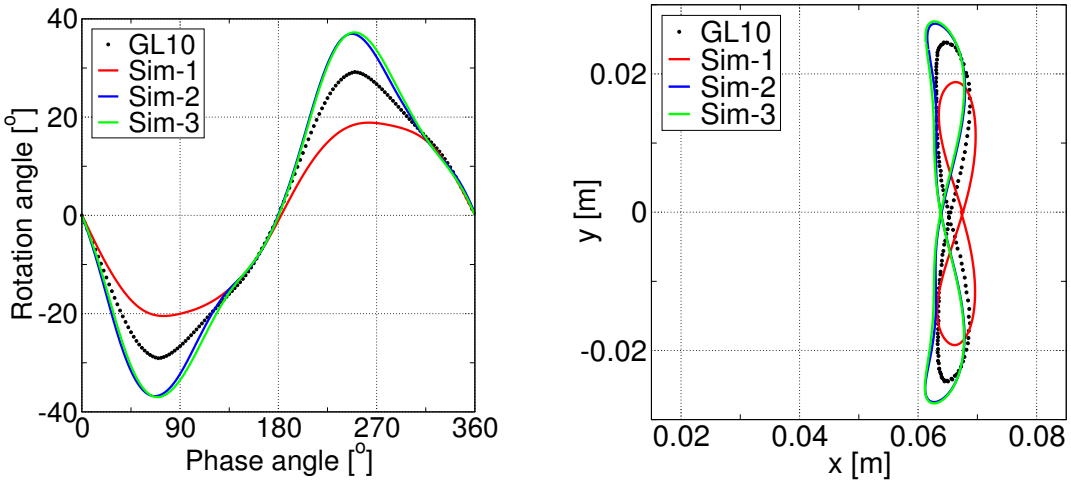
Figure 3: Phase-resolved contours of absolute velocity at different phase angles for Sim-1, comparison of experiment (top) and simulation (bottom).

Table 2: Oscillation frequency, the end mass phase shift and y extrema of the end mass displacement

	f_{FSI}[hz]	ϕ_{shift}[deg]	$(u_y)^*_{max}$	$(u_y)^*_{min}$
GL10 [6]	4.45	95	1.12	-1.11
Sim-1	4.58	109	0.85	-0.90
Sim-2	4.37	84	1.24	-1.25
Sim-3	4.37	83	1.25	-1.26

of the 3-d flow configuration in a 2-d flow simulation can explain the under-prediction of the the structural deflections. A study conducted by Breuer [1] for 2-d simulation of a circular cylinder at $R_e = 3900$, reports the overdamping of turbulent fluid motion.

Sim-2 and Sim-3 exhibit a close agreement in predicted values of the structural deflections. y extrema of the end mass displacement are about 13% higher than the experimental extrema of the end mass. A reason for this pronounced increase in the structural deflections could be the negligence of structural damping. A study of the laminar version of this test case [7] produced good agreement when simulating the *Movement Induced Instability* (MII), whereas large differences are observed for IIE where the structure oscillates in the first mode, as it is the case with this turbulent benchmark test case. The new test cases proposed in [4, 3] (also introduced in Section 1) study the effects of material damping in two different modes of the structural oscillation. The study depicts a higher importance of the material damping in the first mode of the structural oscillation, where the damping model significantly effects the structural deflections in numerical simulation. Nevertheless the material damping for the rubber (used in [4, 3]) would be higher than steel, and the premise that material damping is the cause of the over-prediction in Sim-2 and Sim-3, might not apply. Other possible reason for the differences between simulations



(a) (b)

Figure 4: (a)Cylinder rotation angle plotted against time phase angle. (b)Trailing edge coordinates.

and the experiment could be the ignored side walls, taken as symmetry boundaries in the simulations to reduce the computational cost. The effect of side wall boundary layers are ignored on the assumption that the test section width in experiment is too high (about 8 times the diameter of cylinder) for the side wall boundary layers to have a significant effect.

7 CONCLUSIONS

We have presented the results for a turbulent FSI benchmark test case. Three different turbulence modeling techniques have been studied. The 2-d URANS depicts a reasonable agreement with the experiment, regardless of the 3-d flow configuration. The excessive fluid damping in 2-d URANS is considered to be the cause of underrated structural deflections. The LES and the DDES simulation reproduce a close agreement between each other and an acceptable agreement with the experiment. The probable causes of overestimation of deflection in two simulations are also discussed. Further investigations with a variation of the structural material model are planned, as well as the simulation of the test cases [4, 3] with $\zeta - f$ model.

8 ACKNOWLEDGMENTS

The authors gratefully acknowledge the Research Training Group 1344 of the German Research Foundation for the funding of this project.

REFERENCES

[1] M. Breuer. Large eddy simulation of the subcritical flow past a circular cylinder: numerical and modeling aspects. *International Journal for Numerical Methods in Fluids*, 28(9):1281–1302, 1998.

[2] P. Causin, J.F. Gerbeau, and F. Nobile. Added-mass effect in the design of partitioned algorithms for fluid–structure problems. *Computer methods in applied mechanics and engineering*, 194(42):4506–4527, 2005.

[3] G. De Nayer and M. Breuer. Numerical fsi investigation based on les: Flow past a cylinder with a flexible splitter plate involving large deformations (fsi-pfs-2a). *International Journal of Heat and Fluid Flow*, 50:300–315, 2014.

[4] G. De Nayer, A. Kalmbach, M. Breuer, S. Sicklinger, and R. Wüchner. Flow past a cylinder with a flexible splitter plate: A complementary experimental–numerical investigation and a new fsi test case (fsi-pfs-1a). *Computers & Fluids*, 99:18–43, 2014.

[5] M. Germano, U. Piomelli, P. Moin, and W.H. Cabot. A dynamic subgrid-scale eddy viscosity model. *Physics of Fluids A: Fluid Dynamics (1989-1993)*, 3(7):1760–1765, 1991.

[6] J.P. Gomes and H. Lienhart. Experimental benchmark: Self-excited fluid-structure interaction test cases. In H.J. Bungartz, M. Mehl, and M. Schäfer, editors, *Fluid Structure Interaction II*, volume 73 of *Lecture Notes in Computational Science and Engineering*, pages 383–411. Springer Berlin Heidelberg, 2010.

[7] J.P. Gomes, S. Yigit, H. Lienhart, and M. Schäfer. Experimental and numerical study on a laminar fluid-structure interaction reference test case. *Journal of Fluids and Structures*, 27(1):43 – 61, 2011.

[8] K. Hanjalić, M. Popovac, and M. Hadžiabdić. A robust near-wall elliptic-relaxation eddy-viscosity turbulence model for CFD. *International Journal of Heat and Fluid Flow*, 25(6):1047–1051, 2004.

[9] MpCCI - Mesh-Based Parallel Code Coupling Interface. User guide v3.0. *Fraunhofer, SCAI*, 2004.

[10] H. Jasak, H.G. Weller, and A.D. Gosman. High resolution NVD differencing scheme for arbitrarily unstructured meshes. *International Journal for Numerical Methods in Fluids*, 31(2):431–449, 1999.

[11] A. Kalmbach and M. Breuer. Experimental PIV/V3V measurements of vortex-induced fluid–structure interaction in turbulent flow - A new benchmark FSI-PfS-2a. *Journal of Fluids and Structures*, 42:369–387, 2013.

[12] D.K. Lilly. A proposed modification of the Germano subgrid-scale closure method. *Physics of Fluids A: Fluid Dynamics (1989-1993)*, 4(3):633–635, 1992.

[13] FASTEST - User Manual. Institute of numerical methods in mechanical engineering. *Technische Universität Darmstadt*, 2005.

[14] E. Naudascher and D. Rockwell. *Flow-induced vibrations: An engineering guide*. Courier Dover Publications, 2012.

[15] B.R. Noack. On the flow around a circular cylinder. part ii: turbulent regime. *ZAMM-Journal of Applied Mathematics and Mechanics/Zeitschrift für Angewandte Mathematik und Mechanik*, 79(S1):227–230, 1999.

[16] R.W. Ogden. *Non-linear elastic deformations*. Courier Dover Publications, 1997.

[17] M. Schäfer, D.C. Sternel, G. Becker, and P. Pironkov. Efficient numerical simulation and optimization of fluid-structure interaction. In H.J. Bungartz, M. Mehl, and M. Schäfer, editors, *Fluid Structure Interaction II*, volume 73 of *Lecture Notes in Computational Science and Engineering*, pages 131–158. Springer Berlin Heidelberg, 2010.

[18] J. Smagorinsky. General circulation experiments with the primitive equations: I. The basic experiment. *Monthly weather review*, 91(3):99–164, 1963.

[19] P.R. Spalart, S. Deck, M.L. Shur, K.D. Squires, M.Kh. Strelets, and A. Travin. A new version of detached-eddy simulation, resistant to ambiguous grid densities. *Theoretical and Computational Fluid Dynamics*, 20(3):181–195, 2006.

[20] P.R. Spalart, W.H. Jou, M. Strelets, and S.R. Allmaras. Comments on the feasibility of LES for wings, and on a hybrid RANS/LES approach. In C. Liu and Z. Liu, editors, *Advances in DNS/LES: Proceedings of the First AFOSR International Conference on DNS/LES, Louisiana Tech University, Ruston, Louisiana, USA*. Greyden Press Columbus, OH, August 1997.

[21] R.L. Taylor. Feap - A Finite Element Analysis Program, Version 7.5 User Manual. *University of California at Berkeley, Berkeley, CA*, 2005.

[22] A. Travin, M. Shur, M. Strelets, and P.R. Spalart. Physical and numerical upgrades in the detached-eddy simulation of complex turbulent flows. In R. Friedrich and W. Rodi, editors, *Advances in LES of Complex Flows*, volume 65 of *Fluid Mechanics and its Applications*, pages 239–254. Springer Netherlands, 2002.

[23] C. Truesdell and W. Noll. *The non-linear field theories of mechanics*. Springer, 2004.

[24] S. Türk, T. Reimann, D.C. Sternel, and M. Schäfer. On the performance of detached-eddy simulation in fluid-structure interaction problems. In S. Idelsohn, M. Papadrakakis, and B. Schrefler, editors, *Computational Methods for Coupled Problems in Science and Engineering, COUPLED PROBLEMS 2013*, June 2013.

[25] L. Xue. *Entwicklung eines effizienten parallelen Lösungsalgorithmus zur dreidimensionalen Simulation komplexer turbulenter Strömungen.* PhD thesis, Technische Universität Berlin, 1998.

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis(Eds)

COUPLED CFD-CAA APPROACH FOR ROTATING SYSTEMS

M. Kaltenbacher*, A. Hüppe*, A. Reppenhagen† and W. Kühnel‡

*Institute of Mechanics and Mechatronics Vienna University of Technology
Getreidemarkt 9, 1060 Vienna, Austria
e-mail: {manfred.kaltenbacher,andreas.hueppe}@tuwien.ac.at

†Virtual Vehicle Research Center
Kompetenzzentrum - Das virtuelle Fahrzeug Forschungsgesellschaft mbH
Inffeldgasse 21/A, 8010 Graz, Austria
e-mail: aaron.reppenhagen@v2c2.at

† MAHLE Behr GmbH & Co. KG
Mauserstr. 3, 70469 Stuttgart, Germany
e-mail: wolfram.kuehnel@mahle.com

Key words: CFD, CAA, Arbitrary Mesh Interface, Multiphysics, Non-conforming Grids

Abstract. We present a recently developed computational scheme for the numerical simulation of flow induced sound for rotating systems. Thereby, the flow is fully resolved in time by utilizing a DES (Detached Eddy Simulation) turbulance model and using an arbitrary mesh interface scheme for connecting rotating and stationary domains. The acoustic field is modeled by a perturbation ansatz resulting in a convective wave equation based on the acoustic scalar potential and the substational time derivative of the incompressible flow pressure as a source term. We use the Finite-Element (FE) method for solving the convective wave equation and apply a Nitsche type mortaring at the interface between rotating and stationary domains. The whole scheme is applied to the numerical computation of a side channel blower.

1 INTRODUCTION

The cabin noise of modern ground vehicles is highly affected by flow related noise sources. This is especially the case, when the vehicle is not moving. Thereby, the fan-noise and outlet of the air-conditioning system are main acoustic sources and may reduce the comfort significantly. Rotating fans generate a highly turbulent flow field and can be identified as the main noise source in air conditioning units. Therefore, in addition to the aerodynamic efficiency of the fan, the acoustic signature of the fan is a main design criterion.

This contribution focuses on Computational Fluid Dynamics (CFD) simulations of rotating fans in air conditioning units using the Arbitrary Mesh Interface (AMI) which is implemented in OpenFOAM. For the computation of the acoustic sources, highly accurate unsteady CFD simulation data is needed. Therefore, the transient simulations are carried out by using a DES (Detached Eddy Simulation) turbulence model to accurately resolve the complex flow field. In addition, CAA (Computational AeroAcoustics) simulations with the Finite-Element (FE) research software CFS++ (Coupled Field Simulation) are performed, which uses a Nitsche type mortaring to couple the acoustic field between rotating and stationary parts [1]. By introducing on the interface between moving and quiescent grid an appropriate flux term in combination with a penalization term, the method retains symmetry, consistency and stability of the algebraic system of equations. Furthermore, to precisely approximate the acoustic far field condition, we apply our recently developed PML (Perfectly Matched Layer) technique [2]. To demonstrate the applicability of our overall computational scheme, we will present CFD and CAA computations of a side channel blower, as used in automotive air-conditioning systems.

2 Computational Aeroacoustics

The acoustic/viscous splitting technique for the prediction of flow induced sound was first introduced in [3], and afterwards many groups presented alternative and improved formulations for linear and non linear wave propagation [4, 5, 6, 7]. These formulations are all based on the idea, that the flow field quantities are split into compressible and incompressible parts.

We introduce a generic splitting of physical quantities to the Navier-Stokes equations. For this purpose we choose the following ansatz

$$p = \bar{p} + p^{\mathrm{ic}} + p^{\mathrm{c}} = \bar{p} + p^{\mathrm{ic}} + p^{\mathrm{a}} \tag{1}$$

$$\boldsymbol{v} = \bar{\boldsymbol{v}} + \boldsymbol{v}^{\mathrm{ic}} + \boldsymbol{v}^{\mathrm{c}} = \bar{\boldsymbol{v}} + \boldsymbol{v}^{\mathrm{ic}} + \boldsymbol{v}^{\mathrm{a}} \tag{2}$$

$$\rho = \bar{\rho} + \rho_1 + \rho^{\mathrm{a}} \,. \tag{3}$$

Thereby the field variables are split into mean ($\bar{p}$, $\bar{\boldsymbol{v}}$, $\bar{\rho}$) and fluctuating parts just like in the Linearized Euler Equations (LEE). In addition the fluctuating field variables are split into acoustic (p^{a}, $\boldsymbol{v}^{\mathrm{a}}$, ρ^{a}) and flow components (p^{ic}, $\boldsymbol{v}^{\mathrm{ic}}$). Finally, a density correction ρ_1 is build in according to (3). This choice is motivated by the following assumptions:

- The acoustic field is a fluctuating field.
- The acoustic field is irrotational, i.e. $\nabla \times \boldsymbol{v}^{\mathrm{a}} = 0$, and therefore may be expressed by the acoustic scalar potential ψ^{a} via

$$\boldsymbol{v}^{\mathrm{a}} = -\nabla \psi^{\mathrm{a}} \,. \tag{4}$$

- The acoustic field requires compressible media and an incompressible pressure fluctuation is not equivalent to an acoustic pressure fluctuation.

By doing so, we arrive for an incompressible flow at the following perturbation equations (for a detailed derivation of this perturbation equations we refer to [8])

$$\frac{\partial p^{\mathrm{a}}}{\partial t} + \overline{\boldsymbol{v}} \cdot \nabla p^{\mathrm{a}} + \rho_0 c^2 \nabla \cdot \boldsymbol{v}^{\mathrm{a}} = -\frac{\partial p^{\mathrm{ic}}}{\partial t} - \overline{\boldsymbol{v}} \cdot \nabla p^{\mathrm{ic}} \tag{5}$$

$$\rho_0 \frac{\partial \boldsymbol{v}^{\mathrm{a}}}{\partial t} + \rho_0 \nabla \left(\overline{\boldsymbol{v}} \cdot \boldsymbol{v}^{\mathrm{a}} \right) + \nabla p^{\mathrm{a}} = 0 \tag{6}$$

with the speed of sound c and mean density ρ_0. This system of partial differential equations corresponds to [5]. The source term is the substantial derivative of the incompressible flow pressure p^{ic}.

Using the acoustic scalar potential ψ^{a}, we may rewrite (6) by

$$\nabla \left(\rho_0 \frac{\partial \psi}{\partial t} + \rho_0 \overline{\boldsymbol{v}} \cdot \nabla \psi - p_{\mathrm{a}} \right) = 0 \,, \tag{7}$$

and arrive at

$$p^{\mathrm{a}} = \rho_0 \frac{\partial \psi^{\mathrm{a}}}{\partial t} + \rho_0 \overline{\boldsymbol{v}} \cdot \nabla \psi^{\mathrm{a}} \,. \tag{8}$$

Now, as shown in [9], we replace in the convective terms the material (Eulerian) velocity $\overline{\boldsymbol{v}}$ by the convective velocity $\overline{\boldsymbol{v}} - \boldsymbol{v}_{\mathrm{r}}$, where $\boldsymbol{v}_{\mathrm{r}}$ is the mechanical velocity of rotating parts. By doing so, we arrive at an ALE (Arbitrary Langrangian Eulerian) description for the acoustic pressure

$$p^{\mathrm{a}} = \rho_0 \frac{\partial \psi^{\mathrm{a}}}{\partial t} + \rho_0 \left(\overline{\boldsymbol{v}} - \boldsymbol{v}_{\mathrm{r}} \right) \cdot \nabla \psi^{\mathrm{a}} = \rho_0 \frac{D \psi^{\mathrm{a}}}{D t} \,; \quad \frac{D}{D t} = \frac{\partial}{\partial t} + \left(\overline{\boldsymbol{v}} - \boldsymbol{v}_{\mathrm{r}} \right) \cdot \nabla \,. \tag{9}$$

Furthermore, we transform (5) to the ALE framework

$$\frac{\partial p^{\mathrm{a}}}{\partial t} + \left(\overline{\boldsymbol{v}} - \boldsymbol{v}_{\mathrm{r}} \right) \cdot \nabla p^{\mathrm{a}} + \rho_0 c^2 \nabla \cdot \boldsymbol{v}^{\mathrm{a}} = -\frac{\partial p^{\mathrm{ic}}}{\partial t} - \left(\overline{\boldsymbol{v}} - \boldsymbol{v}_{\mathrm{r}} \right) \cdot \nabla p^{\mathrm{ic}} \,. \tag{10}$$

Now, we substitute (9) into (10) and arrive at

$$\frac{1}{c^2} \, \frac{D^2 \psi^{\mathrm{a}}}{D t^2} - \Delta \psi^{\mathrm{a}} = -\frac{1}{c^2} \, \frac{D p^{\mathrm{ic}}}{D t} \,. \tag{11}$$

This convective wave equation fully describes acoustic sources generated by incompressible flow structures and its wave propagation through flowing media. In addition, instead of the original unknowns p^{a} and $\boldsymbol{v}^{\mathrm{a}}$ we have know just the scalar unknown ψ^{a}. Furthermore, it allows to perform computations on stationary and rotating domains and results in the acoustic scalar potential, from which the acoustic pressure p^{a} is computed via (9).

3 Application

The setup of investigation consists of a side channel blower as used in todays automotive air conditioning units [10]. Figure 1 displays the geometry with the stationary and rotating domains.

The used work-flow is depicted in Fig. 2, which illustrates the forward coupling of the flow field to the acoustic field.

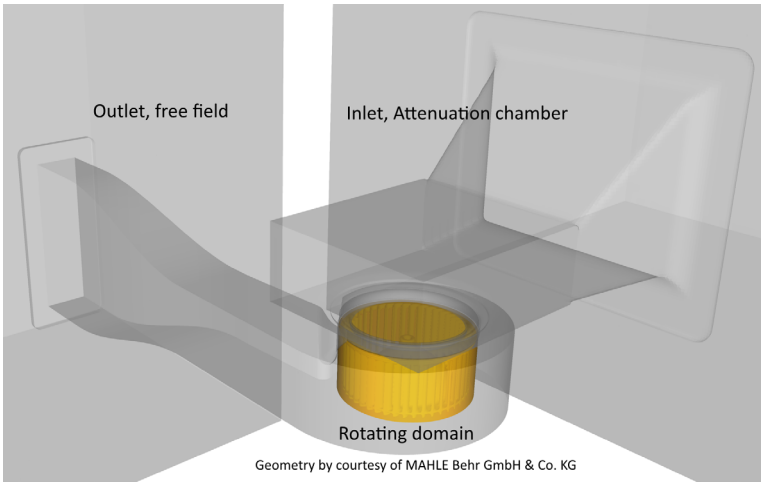


Figure 1: Geometry with rotating domain.

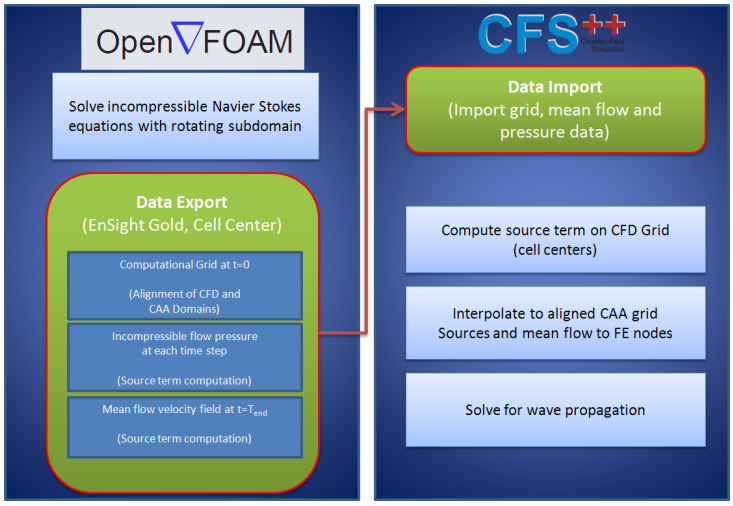


Figure 2: Work-flow for hybrid aeroacoustic computations in OpenFOAM and CFS++.

As discussed in the previous section, we compute the aeroacoustic sources based on the flow field and project them to the acoustic domain using a conservative mapping algorithm.

3.1 CFD

In this contribution the OpenFOAM (Open Field Operation and Manipulation) CFD Toolbox version 2.3.0 is used. OpenFOAM is an Open Source library of C++ routines to solve the Navier-Stokes equations based on the finite volume method. Since version 2.1.0 the arbitrary mesh interface (AMI) was implemented based on the algorithm described in [11]. The AMI allows simulation across disconnected, but adjacent mesh domains, which are especially required for rotating geometries.

Regarding the computational setup (see Fig. 1), we embed inside the quiescent domain (displayed in grey) a rotating subregion, marked in orange. The boundary of the rotating region is modeled through the AMI. The blower rotates with 1.860 rotations per minute giving a maximum velocity at the rotating interface of about 13 m/s, which corresponds to a Mach number of about 0.038. As the interior of the subregion is rotated around the z-axis, an air stream is generated owing towards the outlet with maximum velocities of about 35 m/s in the whole domain. Thereby we can safely assume the assumption of incompressible flow to be valid which enables the utilization of the given hybrid approach. The flow solution is computed using the pimpleDyMFoam solver implemented in OpenFOAM which can handle dynamic meshes with a time step size of $\Delta t = 10\,\mu$s. For the CFD computation a hex-dominant finite volume mesh consisting of 16.4 million cells was created with the automatic mesh generator HEXPRESS$^{\text{TM}}$ / Hybrid from Numeca. The transient simulation was carried out by using a detached-eddy simulation based on the Spalart-Allmaras turbulence model to accurately resolve the complex flow field [12]. The calculation was performed on the Vienna Scientific Cluster VSC2 with 144 Cores [13]. Figure 3 illustrates the highly unsteady flow field around the blower.

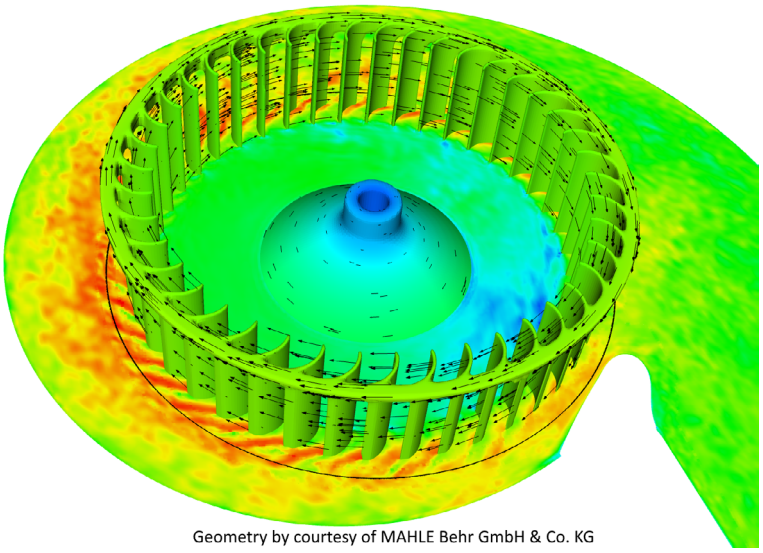
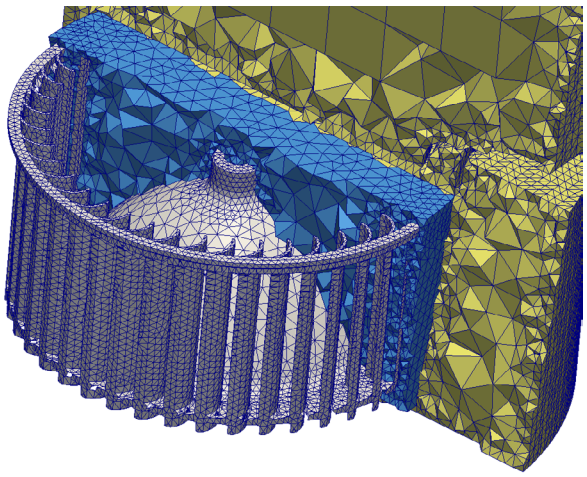


Figure 3: Cut at the z-plane colored by velocity, the black line represents the AMI.

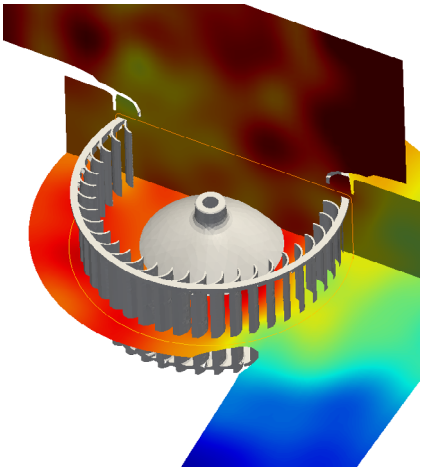
3.2 CAA

In accordance to the flow computation, the rotating domain is embedded into a quiescent propagation region. Furthermore, we add at the inflow and outflow boundaries of the CFD domain two additional regions, on which we apply an advanced *Perfectly-Matched-Layer* technique to effectively approximate acoustic free field conditions [2].

A close up of the computational grid is depicted in Fig. 4a. It can be seen, that a tetrahedral mesh is utilized around the more complex geometries and a coarser, hexahedra-dominant mesh is defined in the remainder of the domain. To create the grid, the Preprocessor ICEM is utilized and the maximum element size ($h_{\max} = 2\text{cm}$) is chosen to accurately compute acoustic wave propagation up to a frequency of approximately 3kHz. The resulting computational grid thereby contains 2.3 million elements and 682.952 nodes. To solve the resulting system of equations in each time step, a parallelized GMRES (Gen-



(a) Slice through computational grid for acoustic computations with rotating (blue) and quiescent (yellow) regions.

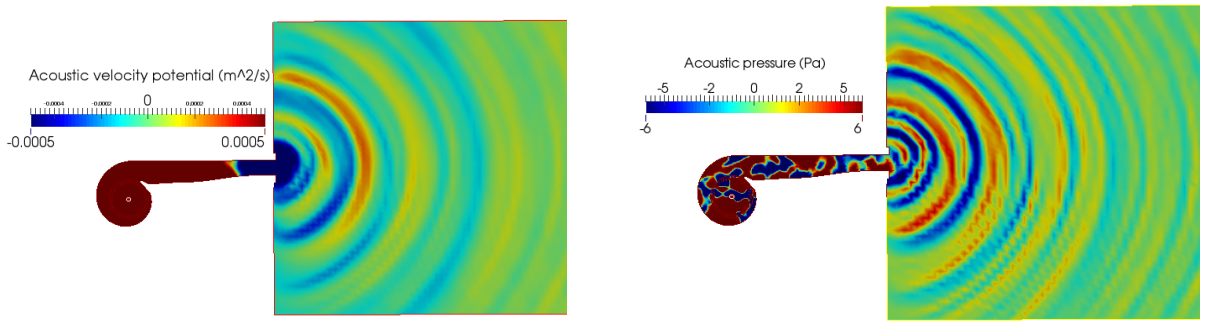


(b) Contours of acoustic velocity potential at the interface between computational regions. No numerical disturbances at the interface are visible.

Figure 4: Computational CAA grid and contours of acoustic velocity potential.

eralized Minimal Residual Method) implementation is utilized which converges to a minimum residual of 10^{-10} using an average of 108 iterations per time step. Utilizing a direct solution procedure is not necessarily beneficial in the context of rotating systems as the system matrix is altered in each time step due to the update of the rotating interface.

For a characteristic time step, we display in Fig. 4b the acoustic velocity potential at the interface. We can observe no visible numerical disturbances which demonstrates the correct physical computation of the acoustic field. Finally, Fig. 5 provides an impression of the acoustic wave propagation towards the exterior domain.



(a) Contours of velocity potential.

(b) Contours of acoustic pressure.

Figure 5: Acoustic waves in outlet domain.

4 Conclusion

A recently developed numerical scheme for computational aeroacoustics has been applied to calculate the noise generated by a side channel blower as used in automotive air-conditioning systems. Thereby, the flow is computed by a DES turbulence model and utilizing an arbitrary mesh interface between rotating and stationary domains. The acoustic field is modeled by a perturbation ansatz to separate flow and acoustic quantities, which results in a convective acoustic wave equation with the substantial derivative of the incompressible flow pressure as a source term. The equation is solved by an FE formulation with a Nitsche type mortaring coupling the acoustic field between the rotating and stationary domains.

Acknowledgments

The authors would like to acknowledge the financial support of the "COMET - Competence Centers for Excellent Technologies Programme" of the Austrian Federal Ministry for Transport, Innovation and Technology (bmvit), the Austrian Federal Ministry of Science, Research and Economy (bmwfw), the Austrian Research Promotion Agency (FFG), the Province of Styria and the Styrian Business Promotion Agency (SFG).

REFERENCES

[1] A. Hüppe, J. Grabinger, M. Kaltenbacher, A. Reppenhagen, G. Dutzler, and W. Kühnel. Non-Conforming Finite Element Method for Computational Aeroacoustics in Rotating Systems. In *20th AIAA/CEAS Aeroacoustics Conference*, 2014.

[2] B. Kaltenbacher, M. Kaltenbacher, and I. Sim. A modified and stable version of a perfectly matched layer technique for the 3-d second order wave equation in time domain with an application to aeroacoustics. *Journal of Computational Physics*, 235(0):407 – 422, 2013.

[3] J. C. Hardin and D. S. Pope. An acoustic/viscous splitting technique for computational aeroacoustics. *Theoretical and Computational Fluid Dynamics*, 6:323–340, 1994.

[4] W. Z. Shen, J. N. Sørensen. Aeroacoustic modelling of low-speed flows. *Theoretical and Computational Fluid Dynamics*, 13:271–289, 1999.

[5] R. Ewert and W. Schröder. Acoustic perturbation equations based on flow decomposition via source filtering. *J. Comp. Phys.*, 188:365–398, 2003.

[6] J.H. Seo and Y.J. Moon. Perturbed compressible equations for aeroacoustic noise prediction at low mach numbers. *AIAA Journal*, 43:1716–1724, 2005.

[7] C.D. Munz, M. Dumbser, and S. Roller. Linearized acoustic perturbation equations for low Mach number flow with variable density and temperature. *Journal of Computational Physics*, 224:352 – 364, 2007.

[8] A. Hüppe. *Spectral Finite Elements for Acoustic Field Computation.* PhD thesis, Alpen-Adria-Universität Klagenfurt, Austria, 2013.

[9] J. Donea, A. Huerta, J. Ph. Ponthot, and A. Rodriguez-Ferran. *Encyclopedia of Computational Mechanics*, chapter Arbitrary Lagrangian-Eulerian methods. Wiley, 2004.

[10] M. Weinmann. Validation of a transient blower simulation methodology. In *OpenFOAM User's Conference*, 2014.

[11] P. E. Farrell and J. R. Maddison. Conservative interpolation between volume meshes by local Galerkin projection. *Computer Methods in Applied Mechanics and Engineering*, 200:89–100, 2011.

[12] P. R. Spalart and S. R. Allmaras. A One-Equation Turbulence Model for Aerodynamic Flows. *Recherche Aerospatiale*, (1):5–21, 1994.

[13] Vienna scientific Cluster, http://vsc.ac.at/home/.

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis(Eds)

EXTENSION OF THE LS-STAG IMMERSED BOUNDARY METHOD FOR RANS-BASED TURBULENCE MODELS AND ITS APPLICATION FOR NUMERICAL SIMULATION IN COUPLED HYDROELASCTIC PROBLEMS

VALERIA V. PUZIKOVA AND ILIA K. MARCHEVSKY

Bauman Moscow State Technical University (BMSTU)
Applied Mathematics department
Russia, 105005 Moscow, 2nd Baumanskaya, 5
e-mail: valeria.puzikova@gmail.com, iliamarchevsky@mail.ru

Key words: Immersed Boundary Methods, The LS-STAG Method, Coupled Hydroelastic Problems, The Spalart-Allmaras RANS Model, Multigrid Methods

Abstract. The general approach to the application of the LS-STAG method for RANS simulation and for numerical simulation in coupled hydroelastic problems is suggested. According to the concept of the LS-STAG method normal Reynolds stress components are sampled on the base mesh (similar to pressure discretization) and shear ones are sampled in the upper right corners of the base mesh cells. Thus, for the shear Reynolds stresses an additional mesh (xy-mesh) is introduced. In this research the LS-STAG-discretization for convective and diffusive fluxes on the xy-mesh is developed. A software package is developed for numerical solution of the Navier-Stokes and Reynolds-averaged Navier-Stokes equations by using the LS-STAG method. To validate this approach and developed software package the flow past circular airfoil at the Reynolds numbers Re = 1000 and Re = 3900 was simulated. Some numerical results are also presented for simulation of a circular airfoil wind resonance phenomenon.

1 INTRODUCTION

Immersed boundary methods [1] have become popular in Computational Fluid Dynamics over recent years for simulating flows through complex solid geometries and in coupled hydroelasctic problems. The advantage of these methods over a method with a body-fitted mesh is their computational efficiency: they do not require regridding when domain shape changes in the simulation process due to hydroelastic body motion.

The LS-STAG method [2] for viscous incompressible flows simulation combines the advantages of immersed boundary methods, the marker and cells (MAC) method and level-set [3] method. The LS-STAG method allows to solve on the Cartesian meshes such

complicated problems of computational mechanics as coupled hydroelastic problems. In contrast to classical immersed boundary methods, the flow variables are computed in the cut-cells, and not interpolated. Moreover, the LS-STAG discretization in the cut-cells is performed by the same formulae as in rectangular cells. Numerical analogues of total mass, momentum and kinetic energy conservation laws in each cell is a base of LS-STAG discretization constructing, which allows to obtain physically realistic numerical solution.

However, the LS-STAG method, as all mesh methods has a significant limitation when simulating flows with high Reynolds number: it requires extremely small space and time steps. It leads to significant increase in computational cost. The traditional method of solving this problem is Reynolds averaged Navier—Stokes (RANS) simulation, Large Eddy Simulation (LES), Detached Eddy Simulation (DES) *etc.* [4] usage. However, as the analysis of the literature shows such modifications of the LS-STAG method currently do not exist. Therefore the purpose of this study is the constructing of the LS-STAG method extension for the numerical solution of two-dimensional RANS equations. This modification of the LS-STAG method is presented below.

2 GOVERNING EQUATIONS

Two-dimensional RANS equations is considered:

$$\nabla \cdot \mathbf{v} = 0, \quad \frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla)\mathbf{v} = \nabla p + \frac{1}{\mathrm{Re}}\Delta \mathbf{v} + \nabla \cdot \hat{\tau}^{\,t}. \tag{1}$$

Here $\mathbf{v} = \mathbf{v}(x, y, t) = u \cdot \mathbf{e}_x + v \cdot \mathbf{e}_y$ is the dimensionless Reynolds averaged velocity, $p = p(x, y, t)$ is the dimensionless Reynolds averaged pressure, $\hat{\tau}^t$ is the Reynolds stresses tensor. The relationship between $\hat{\tau}^t$ and flow Reynolds averaged variables is given by the turbulence model.

The boundary conditions on the computational domain are the following:

$$\mathbf{v}\big|_{\text{inlet}} = \mathbf{v}_\infty, \quad \frac{\partial \mathbf{v}}{\partial \mathbf{n}}\Big|_{\text{outlet}} = 0, \quad \frac{\partial p}{\partial \mathbf{n}}\Big|_{\text{inlet\&outlet}} = 0, \tag{2}$$

and boundary conditions on the airfoil are no-slip conditions:

$$\mathbf{v}\big|_{\text{airfoil}} = \mathbf{v}^{\text{ib}}, \quad \frac{\partial p}{\partial \mathbf{n}}\Big|_{\text{airfoil}} = 0. \tag{3}$$

Here $\mathbf{v}^{\text{ib}}$ is the velocity of the immersed boundary. The airfoil assumed to be rigid and it can oscillate with 1, 2 or 3 degrees of freedom. Its motion is described by dynamics equations:

$$\ddot{\vec{q}} = \vec{\Phi}(\vec{q}, \dot{\vec{q}}) + \vec{Q}^{\text{flow}} + \vec{Q}^{\text{ext}}. \tag{4}$$

Here $\vec{q}$ is the airfoil generalized coordinates vector, $\vec{\Phi}(\vec{q}, \dot{\vec{q}})$ is determined by elastic and viscous constraints imposed on the airfoil, $\vec{Q}^{\text{flow}}$ is the generalized aerodynamic force, $\vec{Q}^{\text{ext}}$ is external mass generalized forces vector.

In case of Reynolds Stress (RSM) RANS models the Reynolds stress transport equation is solved for simulating of $\hat{\tau}^{\,t}$. In case of Eddy Viscosity (EVM) RANS models the

eddy viscosity ν^t (and the turbulent kinetic energy k in case of two-equation models) is simulated and Reynolds stresses are evaluated using the Boussinesq assumption [4]:

$$\hat{\tau}^t = \nu^t \begin{pmatrix} 2\dfrac{\partial u}{\partial x} & \dfrac{\partial u}{\partial y} + \dfrac{\partial v}{\partial x} \\ \dfrac{\partial u}{\partial y} + \dfrac{\partial v}{\partial x} & 2\dfrac{\partial v}{\partial y} \end{pmatrix} + \frac{2}{3} k \hat{I}. \tag{5}$$

Here $\hat{I}$ is the second-order unit tensor.

In this study the details of the LS-STAG discretization constructing for EVM RANS approach are presented on example of the Spallart—Allmaras (S-A) turbulence model [5]. In this model the turbulent kinetic energy is assumed to be close to zero ($k \approx 0$). Thus, the following formulae for normal (6) and shear (7) Reynolds stresses are obtained from (5):

$$\tau^t_{xx} = 2\nu^t \frac{\partial u}{\partial x}, \quad \tau^t_{yy} = 2\nu^t \frac{\partial v}{\partial y}, \tag{6}$$

$$\tau^t_{xy} = \nu^t \left(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} \right). \tag{7}$$

The eddy viscosity in the Spallart—Allmaras turbulence model is given by the following equations:

$$\nu^t = \tilde{\nu} f_{\nu 1}, \quad \frac{\partial \tilde{\nu}}{\partial t} + (\mathbf{v} \cdot \nabla)\tilde{\nu} = P^\nu - D^\nu + \frac{1}{\sigma} \nabla \cdot [(\nu + \tilde{\nu}) \nabla \tilde{\nu}] + \frac{c_{b2}}{\sigma} (\nabla \tilde{\nu})^2. \tag{8}$$

Here $\tilde{\nu}$ is the S-A working variable, $D^\nu = \left(c_{w1} f_w - \dfrac{c_{b1}}{\varkappa^2} f_{t2} \right) \left(\dfrac{\tilde{\nu}}{d} \right)^2$ is the destruction term, $P^\nu = c_{b1}[1 - f_{t2}]\tilde{S}$ is the production term, d is the dimensionless distance from the field point to the nearest wall, ν is the dimensionless viscosity, $f_{t2} = c_{t3} \cdot e^{-c_{t4}\chi^2}$, $g = r + c_{w2}(r^6 - r)$, $c_{w1} = \dfrac{c_{b1}}{\varkappa^2} + \dfrac{1 + c_{b2}}{\sigma}$, $\tilde{S} = \left| \dfrac{\partial u}{\partial y} - \dfrac{\partial v}{\partial x} \right| + f_{\nu 2} \dfrac{\tilde{\nu}}{\varkappa^2 d^2}$, $f_{\nu 1} = \dfrac{\chi^3}{\chi^3 + c^3_{\nu 1}}$, $f_{\nu 2} = 1 - \dfrac{\chi}{1 + \chi f_{\nu 1}}$, $f_w = g \left(\dfrac{1 + c^6_{w3}}{g^6 + c^6_{w3}} \right)^{1/6}$, $\chi = \dfrac{\tilde{\nu}}{\nu}$, $r = \dfrac{\tilde{\nu}}{\tilde{S} \varkappa^2 d^2}$, $c_{t3} = 1.2$, $c_{t4} = 0.5$, $c_{w2} = 0.3$, $c_{w3} = 2$, $c_{\nu 1} = 7.1$, $c_{b1} = 0.1355$, $c_{b2} = 0.622$, $\varkappa = 0.41$, $\sigma = 2/3$.

The 'Trip-Less' (TL) approach [6] is used: firstly very high value of the S-A working variable is setted on inlet boundary and then once a separation zone is formed value of the S-A working variable on inlet boundary is greatly reduced. Thus, the boundary conditions for (8) take the following form:

$$\tilde{\nu}|_{\text{airfoil}} \equiv 0, \quad \left. \frac{\partial \tilde{\nu}}{\partial \mathbf{n}} \right|_{\text{outlet}} = 0, \quad \tilde{\nu}|_{\text{inlet}} = \tilde{\nu}_\infty(t) = \begin{cases} \nu, & t \le t^*, \\ 0.001\nu, & t > t^*. \end{cases} \tag{9}$$

3 MAIN IDEAS OF THE LS-STAG METHOD

The Cartesian mesh with cells $\Omega_{i,j} = (x_{i-1}, x_i) \times (y_{j-1}, y_j)$ is introduced in the rectangular computational domain Ω. It is denoted that $\Gamma_{i,j}$ is the face of $\Omega_{i,j}$ and $\mathbf{x}^c_{i,j} = (x^c_i, y^c_j)$ is the center of this cell. Unknown components $u_{i,j}$ and $v_{i,j}$ of velocity vector $\mathbf{v}$ are computed in the middle of fluid parts of the cell faces. These points are the centers of cells $\Omega^u_{i,j} = (x^c_i, x^c_{i+1}) \times (y_{j-1}, y_j)$ and $\Omega^v_{i,j} = (x_{i-1}, x_i) \times (y^c_j, y^c_{j+1})$ with faces $\Gamma^u_{i,j}$ and $\Gamma^v_{i,j}$ and areas M^x_{ij} and $M^y_{i,j}$ respectively. If $i = \overline{1, N}$, $j = \overline{1, M}$, the base mesh contains $E = N \cdot M$ cells, x-mesh contains $E_x = (N-1) \cdot M$ cells and y-mesh contains $E_y = N \cdot (M-1)$.

The level-set function $\varphi = \varphi(x, y)$ [3] is introduced for immersed boundary Γ^{ib} description [2]. The boundary Γ^{ib} is represented by a line segment on the cut-cell $\Omega_{i,j}$. Locations of this segment endpoints are defined by a linear interpolation of the variable $\varphi_{i,j} = \varphi(x_i, y_j)$. The cell-face fraction ratios $\vartheta^u_{i,j}$ and $\vartheta^v_{i,j}$ are introduced [2]. They take values in interval $[0, 1]$ and represent the fluid parts of the east and north faces of $\Gamma_{i,j}$ respectively. In 2D case, the cut-cells can be classified into trapezoidal, triangular and pentagonal cells. Examples of each type cut-cells are presented on fig. 1.

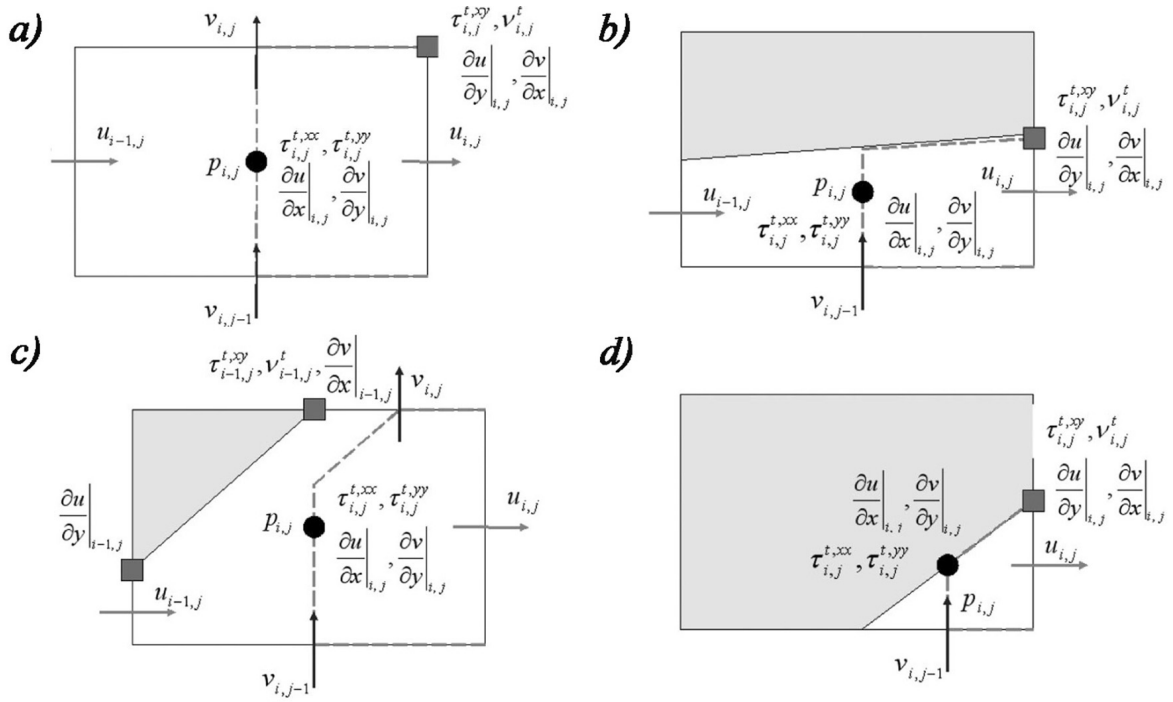


Figure 1: Location of the variables discretization points on the LS-STAG mesh: *(a)* Cartesian Fluid Cell; *(b)* North Trapezoidal Cell; *(c)* Northwest Pentagonal Cell; *(d)* Northwest Triangle Cell.

To preserve the five-point stencil structure of the MAC method we need to make distinction between the discretization of the normal and shear stresses (fig. 1). The normal stresses are sampled in the center of $\Omega_{i,j}$ cell while the shear stresses are sampled

in its corners. Normal and shear Reynolds stresses are sampled similarly. It is conveniently to sample the eddy viscosity and the S-A working variable at the same points as the shear stresses. Thus, in case of the LS-STAG method usage for RANS-based models we need the fourth mesh (xy-mesh with $E_{xy} = (N-1)\cdot(M-1)$ cells) with cells $\Omega^{xy}_{i,j} = (x^c_i, x^c_{i+1}) \times (y^c_j, y^c_{j+1})$. The faces of these cells are $\Gamma^{xy}_{i,j}$ and their areas are $M^{xy}_{i,j}$.

4 THE GENERAL FORM OF THE LS-STAG DISCRETIZATION FOR RANS EQUATIONS

According to the concept of the LS-STAG method equations (1) should be written in integral form for cell of the base mesh, cell of x-mesh and cell of y-mesh respectively:

$$\int\limits_{\Gamma_{i,j}} \mathbf{v}\cdot\mathbf{n}\,dS = 0, \tag{10}$$

$$\frac{d}{dt}\int\limits_{\Omega^u_{i,j}} u\,dV + \int\limits_{\Gamma^u_{i,j}} (\mathbf{v}\cdot\mathbf{n})u\,dS + \int\limits_{\Gamma^u_{i,j}} p\mathbf{e_x}\cdot\mathbf{n}\,dS - \int\limits_{\Gamma^u_{i,j}} \nu\nabla u\cdot\mathbf{n}\,dS - \underline{\int\limits_{\Gamma^u_{i,j}} \tau^t_{xx}\mathbf{e_x}\cdot\mathbf{n}\,dS} - \underline{\underline{\int\limits_{\Gamma^u_{i,j}} \tau^t_{xy}\mathbf{e_y}\cdot\mathbf{n}\,dS}} = 0,$$

$$\frac{d}{dt}\int\limits_{\Omega^v_{i,j}} v\,dV + \int\limits_{\Gamma^v_{i,j}} (\mathbf{v}\cdot\mathbf{n})v\,dS + \int\limits_{\Gamma^v_{i,j}} p\mathbf{e_y}\cdot\mathbf{n}\,dS - \int\limits_{\Gamma^v_{i,j}} \nu\nabla v\cdot\mathbf{n}\,dS - \underline{\int\limits_{\Gamma^v_{i,j}} \tau^t_{yy}\mathbf{e_y}\cdot\mathbf{n}\,dS} - \underline{\underline{\int\limits_{\Gamma^v_{i,j}} \tau^t_{xy}\mathbf{e_x}\cdot\mathbf{n}\,dS}} = 0.$$

Formally, the difference between the integral form of the RANS equations (10) and integral form of the Navier–Stokes equations, for which the LS-STAG discretization was constructed in [2], is only presence of the underlined terms.

The general form of the LS-STAG discretization for (10) can be written as the following:

$$D^x U_x + D^y U_y + \overline{U}^{ib} = 0, \tag{11}$$

$$\frac{d}{dt}(M^x U_x) + C^x U_x + G^x(P - \underline{T_{xx}}) - \underline{\underline{D^\tau_x T_{xy}}} - \nu K^x U_x + S^{ib,c}_x - \nu S^{ib,\nu}_x = 0,$$

$$\frac{d}{dt}(M^y U_y) + C^y U_y + G^y(P - \underline{T_{yy}}) - \underline{\underline{D^\tau_y T_{xy}}} - \nu K^y U_y + S^{ib,c}_y - \nu S^{ib,\nu}_y = 0.$$

Here $P\in\mathbb{R}^E$ is the discrete pressure, $U_x\in\mathbb{R}^{E_x}$ and $U_y\in\mathbb{R}^{E_y}$ are the discrete components of the velocity vector, $T_{xy}\in\mathbb{R}^{E_{xy}}$ is the discrete shear Reynolds stresses, $T_{xx}\in\mathbb{R}^E$ and $T_{yy}\in\mathbb{R}^E$ are the discrete normal Reynolds stresses; $S^{ib,c}_x\in\mathbb{R}^{E_x}$, $S^{ib,\nu}_x\in\mathbb{R}^{E_x}$, $S^{ib,c}_y\in\mathbb{R}^{E_y}$, $S^{ib,\nu}_y\in\mathbb{R}^{E_y}$ are source terms; $\overline{U}^{ib}\in\mathbb{R}^E$ is the mass flux; $D^x\in M(\mathbb{R})_{E\times E_x}$, $D^y\in M(\mathbb{R})_{E\times E_y}$, $D^\tau_x\in M(\mathbb{R})_{E_x\times E_{xy}}$, $D^\tau_y\in M(\mathbb{R})_{E_y\times E_{xy}}$ are the divergence discrete analogues; $K^x\in M(\mathbb{R})_{E_x\times E_x}$ and $K^y\in M(\mathbb{R})_{E_y\times E_y}$ represent the discretization of the diffusive terms; $C^x\in M(\mathbb{R})_{E_x\times E_x}$ and $C^y\in M(\mathbb{R})_{E_y\times E_y}$ represent the discretization of the convective terms; $G^x = -D^T_x$ and $G^y = -D^T_y$ are the gradient discrete analogues.

As in [2], matrices D^x and D^y have the following structure ($i = \overline{1,N}$, $j = \overline{1,M}$):

$$D^x_P(i,j)=\vartheta^u_{i,j}\Delta y_j,\ D^x_W(i,j)=-\vartheta^u_{i-1,j}\Delta y_j,\ D^y_P(i,j)=\vartheta^v_{i,j}\Delta x_i,\ D^y_S(i,j)=-\vartheta^v_{i,j-1}\Delta x_i. \quad (12)$$

We need to construct the LS-STAG discretization of terms with shear Reynolds stress underlined by two lines in (10), as they appear only in the RANS equations. Both for rectangular cells, and for all types of cut-cells the following unified formulae can be written down:

$$\int\limits_{\Gamma^u_{i,j}} \tau^t_{xy}\mathbf{e_y}\cdot\mathbf{n}\,dS \approx \frac{1}{2}(\vartheta^v_{i,j}\Delta x_i + \vartheta^v_{i+1,j}\Delta x_{i+1})\tau^t_x|_{i,j} - \frac{1}{2}(\vartheta^v_{i,j-1}\Delta x_i + \vartheta^v_{i+1,j-1}\Delta x_{i+1})\tau^t_{xy}|_{i,j-1}, \quad (13)$$

$$\int\limits_{\Gamma^v_{i,j}} \tau^t_{xy}\mathbf{e_x}\cdot\mathbf{n}\,dS \approx \frac{1}{2}(\vartheta^u_{i,j}\Delta y_j + \vartheta^u_{i,j+1}\Delta y_{j+1})\tau^t_{xy}|_{i,j} - \frac{1}{2}(\vartheta^u_{i-1,j}\Delta y_j + \vartheta^u_{i-1,j+1}\Delta y_{j+1})\tau^t_{xy}|_{i-1,j}.$$

Thus, matrices D^τ_x and D^τ_y in (11) can be defined as the following:

$$\begin{array}{l} i = \overline{1,N-1}: \left\{ \begin{array}{l} D^\tau_{x,P}(i,j) = (\vartheta^v_{i,j}\Delta x_i + \vartheta^v_{i+1,j}\Delta x_{i+1})/2,\ j = \overline{1,M-1}; \\ D^\tau_{x,S}(i,j) = -(\vartheta^v_{i,j-1}\Delta x_i + \vartheta^v_{i+1,j-1}\Delta x_{i+1})/2,\ j = \overline{2,M}; \end{array} \right. \\ j = \overline{1,M-1}: \left\{ \begin{array}{l} D^\tau_{y,P}(i,j) = (\vartheta^u_{i,j}\Delta y_j + \vartheta^u_{i,j+1}\Delta y_{j+1})/2,\ i = \overline{1,N-1}; \\ D^\tau_{y,W}(i,j) = -(\vartheta^u_{i-1,j}\Delta y_j + \vartheta^u_{i-1,j+1}\Delta y_{j+1})/2,\ i = \overline{2,N}. \end{array} \right. \end{array} \quad (14)$$

The time integration of the differential algebraic system (11) is performed with a semi-implicit Euler scheme. Predictor step leads to discrete analogues of the Helmholtz equation for velocities prediction $\widetilde{U}_x$, $\widetilde{U}_y$ at the time $t_{n+1} = (n+1)\Delta t$:

$$\frac{M^{n+1}_x\widetilde{U}_x - M^n_x U^n_x}{\Delta t} + C^n_x U^n_x + S^{ib,c,n}_x - D^{T,n}_x(P^n - \underline{T^n_{xx}}) - \underline{\underline{D^{\tau,n}_x T^n_{xy}}} - \nu K^{n+1}_x\widetilde{U}_x - \nu S^{ib,\nu,n+1}_x = 0, \quad (15)$$

$$\frac{M^{n+1}_y\widetilde{U}_y - M^n_y U^n_y}{\Delta t} + C^n_y U^n_y + S^{ib,c,n}_y - D^{T,n}_y(P^n - \underline{T^n_{yy}}) - \underline{\underline{D^{\tau,n}_y T^n_{xy}}} - \nu K^{n+1}_y\widetilde{U}_y - \nu S^{ib,\nu,n+1}_y = 0.$$

Here Δt is the constant time discretization step. Corrector step leads to the following discrete analogue of Poisson equation for $\Phi = \Delta t(P^{n+1} - P^n)$:

$$A^{n+1}\Phi = D^{n+1}_x\widetilde{U}_x + D^{n+1}_y\widetilde{U}_y + \overline{U}^{ib,n+1}, \quad (16)$$

$A = -D^x(M^x)^{-1}(D^x)^T - D^y(M^y)^{-1}(M^y)^T$, $A \in M(\mathbb{R})_{E\times E}$. Then flow variables at the time t_{n+1} are computed by the following formulae:

$$U^{n+1}_x = \widetilde{U}_x + (M^{n+1}_x)^{-1}D^{T,n+1}_x\Phi,\ U^{n+1}_y = \widetilde{U}_y + (M^{n+1}_y)^{-1}D^{T,n+1}_y\Phi,\ P^{n+1} = \frac{\Phi}{\Delta t} + P^n. \quad (17)$$

After this step in case of RANS equations, new values of Reynolds stresses T^{n+1}_{xx}, T^{n+1}_{yy}, T^{n+1}_{xy} are computed by solving the discrete analogues of the corresponding equations. It should be noted that linear systems are solved using the BiCGStab method [7] with the ILU- and multigrid [8] preconditioning. An original algorithm for the solver cost-coefficient estimation [9] is used for the optimal parameters choice for the multigrid preconditioner.

5 EXTENSION OF THE LS-STAG METHOD FOR THE SPALLART — ALLMARAS TURBULENCE MODEL

Since the eddy viscosity and shear Reynolds stresses sampling points are the same, according to (7) it is possible to write down:

$$\tau^t_{xy}|_{i,j} = \nu^t_{i,j}\left(\left.\frac{\partial u}{\partial y}\right|_{i,j} + \left.\frac{\partial v}{\partial x}\right|_{i,j}\right). \tag{18}$$

The eddy viscosity average on the cell $\Omega_{i,j}$ of the base mesh should be used for normal Reynolds stresses computing (6):

$$\tau^t_{xx}|_{i,j} = 2\overline{\nu^t}_{i,j}\left.\frac{\partial u}{\partial x}\right|_{i,j},\ \tau^t_{yy}|_{i,j} = 2\overline{\nu^t}_{i,j}\left.\frac{\partial v}{\partial y}\right|_{i,j},\ \overline{\nu^t}_{i,j} = \alpha_{i,j}(\nu^t_{i,j} + \nu^t_{i,j-1} + \nu^t_{i-1,j} + \nu^t_{i-1,j-1}), \tag{19}$$

$$\alpha_{i,j} = \begin{cases} 0, & \text{if } \Omega_{i,j} \text{ is a solid cell,} \\ 1/3, & \text{if } \Omega_{i,j} \text{ is a triangular cell,} \\ 1/4, & \text{otherwise.} \end{cases}$$

Moreover, the eddy viscosity reconstructing by S-A working variable (8) is performed in the obvious way, since ν^t, $\tilde{\nu}$ and $f_{\nu 1}$ are computed at the same points.

The transport equation (8) in integral form has the following form:

$$\frac{d}{dt}\int\limits_{\Omega^*}\tilde{\nu}\,dV + \int\limits_{\Gamma^*}(\mathbf{v}\cdot\mathbf{n})\tilde{\nu}\,dS = \int\limits_{\Omega^*}(P^\nu - D^\nu)\,dV + \frac{1}{\sigma}\int\limits_{\Gamma^*}[\nu+\tilde{\nu}]\nabla\tilde{\nu}\cdot\mathbf{n}\,dS + \frac{c_{b2}}{\sigma}\int\limits_{\Omega^*}(\nabla\tilde{\nu})^2\,dV. \tag{20}$$

It is obvious that $\int\limits_{\Omega^{xy}_{i,j}}(P^\nu - D^\nu)\,dV \cong M^{xy}_{i,j}(P^\nu_{i,j} - D^\nu_{i,j}) = M^{xy}_{i,j}S_{i,j}$ because production and destruction terms are sampled at the same points as the S-A working variable.

Convenience of $\tilde{\nu}$ discretization on the xy-mesh consists mainly in the fact that due to this $\partial\tilde{\nu}/\partial x$ and $\partial\tilde{\nu}/\partial y$ are sampled on the y-mesh and x-mesh respectively:

$$\left.\frac{\partial\tilde{\nu}}{\partial x}\right|_{i,j} = \frac{\tilde{\nu}_{i,j} - \tilde{\nu}_{i-1,j}}{\vartheta^v_{i,j}\Delta x_i}, \quad \left.\frac{\partial\tilde{\nu}}{\partial y}\right|_{i,j} = \frac{\tilde{\nu}_{i,j} - \tilde{\nu}_{i,j-1}}{\vartheta^u_{i,j}\Delta y_j}. \tag{21}$$

The discretization of the last term in (20) is the following:

$$\int\limits_{\Omega^{xy}_{i,j}}(\nabla\tilde{\nu})^2\,dV \cong (\beta_{i+1,j} + \beta_{i+1,j+1})\left(\left.\frac{\partial\tilde{\nu}}{\partial x}\right|_{i+1,j}\right)^2 + (\beta_{i,j} + \beta_{i,j+1})\left(\left.\frac{\partial\tilde{\nu}}{\partial x}\right|_{i,j}\right)^2 + \tag{22}$$

$$+(\beta_{i,j} + \beta_{i+1,j})\left(\left.\frac{\partial\tilde{\nu}}{\partial y}\right|_{i,j}\right)^2 + (\beta_{i,j+1} + \beta_{i+1,j+1})\left(\left.\frac{\partial\tilde{\nu}}{\partial y}\right|_{i,j+1}\right)^2 = S^G_{i,j}.$$

Here $\beta_{i,j} = \alpha_{i,j}V_{i,j}$, $V_{i,j}$ is the area of $\Omega_{i,j}$ cell.

And for the diffusive term in (20) we obtain

$$\int\limits_{\Gamma_{i,j}^{xy}} [\nu+\tilde{\nu}]\nabla\tilde{\nu}\cdot\mathbf{n}\,dS \cong \tag{23}$$

$$\cong \frac{1}{2}\left(\frac{\Delta y_j(\vartheta^u_{i,j}+\vartheta^u_{i+1,j})}{2}+\frac{\Delta y_{j+1}(\vartheta^u_{i,j+1}+\vartheta^u_{i+1,j+1})}{2}\right)\frac{\partial\tilde{\nu}}{\partial x}\bigg|_{i+1,j}\left(\nu+\frac{\tilde{\nu}_{i,j}+\tilde{\nu}_{i+1,j}}{2}\right)-$$

$$-\frac{1}{2}\left(\frac{\Delta y_j(\vartheta^u_{i,j}+\vartheta^u_{i-1,j})}{2}+\frac{\Delta y_{j+1}(\vartheta^u_{i,j+1}+\vartheta^u_{i-1,j+1})}{2}\right)\frac{\partial\tilde{\nu}}{\partial x}\bigg|_{i,j}\left(\nu+\frac{\tilde{\nu}_{i,j}+\tilde{\nu}_{i-1,j}}{2}\right)+$$

$$+\frac{1}{2}\left(\frac{\Delta x_i(\vartheta^v_{i,j}+\vartheta^v_{i,j+1})}{2}+\frac{\Delta x_{i+1}(\vartheta^v_{i+1,j}+\vartheta^v_{i+1,j+1})}{2}\right)\frac{\partial\tilde{\nu}}{\partial y}\bigg|_{i,j+1}\left(\nu+\frac{\tilde{\nu}_{i,j}+\tilde{\nu}_{i,j+1}}{2}\right)-$$

$$-\frac{1}{2}\left(\frac{\Delta x_i(\vartheta^v_{i,j}+\vartheta^v_{i,j-1})}{2}+\frac{\Delta x_{i+1}(\vartheta^v_{i+1,j}+\vartheta^v_{i+1,j-1})}{2}\right)\frac{\partial\tilde{\nu}}{\partial y}\bigg|_{i,j}\left(\nu+\frac{\tilde{\nu}_{i,j}+\tilde{\nu}_{i,j-1}}{2}\right).$$

Then $K^{xy}\in M(\mathbb{R})_{E_{xy}\times E_{xy}}$ and $S^{\nu}_{xy}\in\mathbb{R}^{E_{xy}}$ can be defined:

$$\int\limits_{\Gamma_{i,j}^{xy}} [\nu+\tilde{\nu}]\nabla\tilde{\nu}\cdot\mathbf{n}\,dS \cong K^{xy}_S(i,j)\tilde{\nu}_{i,j-1}+K^{xy}_W(i,j)\tilde{\nu}_{i-1,j}+K^{xy}_P(i,j)\tilde{\nu}_{i,j}+ \tag{24}$$

$$+K^{xy}_E(i,j)\tilde{\nu}_{i+1,j}+K^{xy}_N(i,j)\tilde{\nu}_{i,j+1}+S^{\nu}_{xy,\ i,j}.$$

Similarly, discretization for the convective term in (20) on the xy-mesh is obtained:

$$\int\limits_{\Gamma_{i,j}^{xy}} (\mathbf{v}\cdot\mathbf{n})\tilde{\nu}\,dS \cong \frac{[\bar{u}_{es}]_-+[\bar{u}_{en}]_-}{4}\cdot\tilde{\nu}_{i+1,j}-\frac{[\bar{u}_{ws}]_++[\bar{u}_{wn}]_+}{4}\cdot\tilde{\nu}_{i-1,j}+ \tag{25}$$

$$+\frac{([\bar{u}_{es}]_++[\bar{u}_{en}]_+)-([\bar{u}_{ws}]_-+[\bar{u}_{wn}]_-)+([\bar{v}_{nw}]_++[\bar{v}_{ne}]_+)-([\bar{v}_{sw}]_-+[\bar{v}_{se}]_-)}{4}\cdot\tilde{\nu}_{i,j}+$$

$$+\frac{[\bar{v}_{nw}]_-+[\bar{v}_{ne}]_-}{4}\cdot\tilde{\nu}_{i,j+1}-\frac{[\bar{v}_{sw}]_++[\bar{v}_{se}]_+}{4}\cdot\tilde{\nu}_{i,j-1},$$

$$\bar{u}_{es}=\bar{u}_{i,j}+\bar{u}_{i+1,j},\ \bar{u}_{en}=\bar{u}_{i,j+1}+\bar{u}_{i+1,j+1},\ \bar{u}_{ws}=\bar{u}_{i,j}+\bar{u}_{i-1,j},\ \bar{u}_{wn}=\bar{u}_{i,j+1}+\bar{u}_{i-1,j+1},$$

$$\bar{v}_{nw}=\bar{v}_{i,j}+\bar{v}_{i,j+1},\ \bar{v}_{ne}=\bar{v}_{i+1,j}+\bar{v}_{i+1,j+1},\ \bar{v}_{sw}=\bar{v}_{i,j}+\bar{v}_{i,j-1},\ \bar{v}_{se}=\bar{v}_{i+1,j}+\bar{v}_{i+1,j-1},$$

$$\bar{u}_{i,j}=\vartheta^u_{i,j}u_{i,j}\Delta y_j,\quad \bar{v}_{i,j}=\vartheta^v_{i,j}v_{i,j}\Delta x_i,\quad [c]_+=(c+|c|)/2,\quad [c]_-=(c-|c|)/2.$$

This leads to matrix $C^{xy}\in M(\mathbb{R})_{E_{xy}\times E_{xy}}$. Additionally, the $S^{c}_{xy}\in\mathbb{R}^{E_{xy}}$ contains a non-zero components, which correspond to the boundaries of the computational domain.

Thus, the following difference analogue of (20) is obtained:

$$\tilde{\nu}^{n+1}=(M^{xy,n+1})^{-1}\left[M^{xy,n}\{\tilde{\nu}^n+\Delta t\cdot S^n\}+\Delta t\left(\frac{K^{xy,n}\tilde{\nu}^n+S^{\nu,n}_{xy}+c_{b2}S^{G,n}}{\sigma}-C^{xy,n+1}\tilde{\nu}^n-S^{c,n}_{xy}\right)\right]. \tag{26}$$

6 NUMERICAL EXPERIMENTS

6.1 Flow past a fixed circular airfoil

The flow past circular airfoil was simulated using the developed modification of the LS-STAG method at the Reynolds numbers Re = 1000 (on non-uniform meshes 120×148 with $\Delta t = 5 \cdot 10^{-2}$ and 240×296 with $\Delta t = 10^{-3}$) and Re = 3900 (on non-uniform meshes 120×148 with $\Delta t = 10^{-3}$ and 240×296 with $\Delta t = 5 \cdot 10^{-4}$). These values of the Re were chosen because the experimental data [10] and results of other researchers [11, 12, 13] are known for them. The time averaged drag coefficient C_D and the Strouhal number St were computed. Computational results are shown in table 1. These results are in good agreement with experimental data for simulation on coarse meshes by using the proposed modification of the LS-STAG method.

Table 1: Comparison of C_D and St with established results from the literature.

Turbulence model	Number of cells	Re = 1000		Re = 3900	
		C_D	St	C_D	St
Experiment [10]		0.98	0.21	0.93	0.22
LES [11]	1 103 520	—	—	1.08	—
S-A, present study	17 760	1.12	0.26	0.86	0.18
S-A, present study	71 040	1.03	0.25	1.12	0.25
$k - \varepsilon$ [12]	46 304	1.00	0.15	1.00	0.15
Real $k - \varepsilon$ [12]	46 304	—	0.17	—	0.20
SST $k - \omega$ [12]	46 304	—	0.23	—	0.25
$k - \varepsilon$ [13], ANSYS	388 550	1.17	—	0.74	—
SST $k - \omega$ [13], ANSYS	388 550	0.99	—	0.62	—
LES [13], ANSYS	388 550	1.15	0.21	1.07	—

6.2 Circular airfoil wind resonance

Coupled aeroelastic problems appear when simulating autorotation and auto-oscillations phenomena, in particular, airfoil wind resonance. Such problems are complicated for numerical solution, since it is necessary to take into account interference between the flow and moving immersed body. In case of sufficiently massive body, coupled aeroelastic problems can be solved using step-by-step splitting numerical algorithm, firstly simulating flow around a body moving with known parameters and then computing the dynamics of the body with known hydrodynamic loads [14].

To simulate wind resonance phenomenon we have considered the motion of the circular airfoil with diameter D across the stream (with one degree of freedom). Airfoil's constrain assumed to be linear viscoelastic and its motion (4) is described by the following ordinary

differential equation:

$$m\ddot{y}_* + b\dot{y}_* + cy_* = F_y. \tag{27}$$

Here m is the airfoil's mass, b is the damping factor, c is the constraint's elasticity, F_y is lift force, y_* is the deviation from the equilibrium. The natural frequency of the system $\omega \approx \sqrt{c/m}$ (damping assumed to be small) can be set by varying of the coefficient c.

Number of computations have been performed on non-uniform grid 272×292 with time discretization step $\Delta t = 0.0001$ and the following dimensionless parameters: $\mathrm{Re} = 1000$, $V_\infty = 3.0$, $m = 39.15$, $b = 0.731$. The dimensionless natural frequency of the system is the following:

$$\mathrm{Sh}_\omega = \frac{\omega}{2\pi} \cdot \frac{D}{V_\infty} = 0.150 \ldots 0.280. \tag{28}$$

Computational results are in good agreement with the previous studies [15]. Maximum amplitude (fig. 2) is about $0.4D$ and it occures when the natural frequency of the system St_ω is close to the Strouhal number, calculated for a fixed airfoil $\mathrm{St} \approx 0.24$ [2].

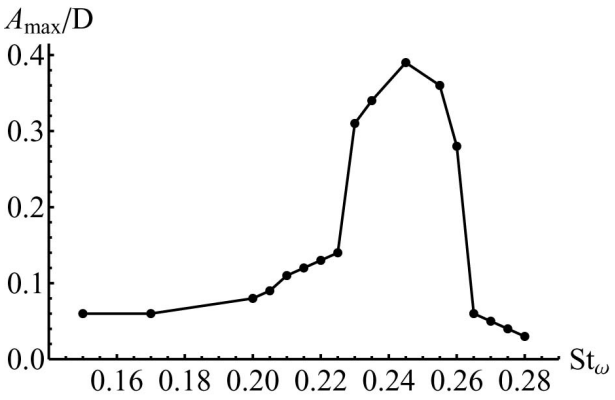


Figure 2: Maximum amplitude of the circular airfoil oscillations at $\mathrm{Re} = 1000$

7 CONCLUSIONS

- The key points of the LS-STAG method extension for RANS-based turbulence models are described. For the shear Reynolds stresses and for the eddy viscosity an additional mesh (xy-mesh) is introduced.

- The general approach to the construction of the LS-STAG discretization for differential equations of the EVM RANS models on the additional xy-mesh is shown on the example of the Spalarat — Allmaras model.

- A software package is developed for the numerical simulation of the bodies' motion in the viscous incompressible flow by using the LS-STAG method and its modifications.

- To validate this approach and the developed software package the flow past a circular airfoil at the Reynolds number values Re = 1000 and Re = 3900 was simulated. Computational results are in good agreement with established results from the literature.

- Simulation of a circular airfoil wind resonance phenomenon is considered. Computational results are in good qualitative agreement with the experimental data.

8 ACKNOWLEGEMENTS

The work was partially supported by Russian Federation President Grant for young scientists [proj. MK-3705.2014.8, MK-5357.2015.8].

REFERENCES

[1] Mittal, R. and Iaccarino, G. Immersed boundary methods. *Annu. Rev. Fluid Mech.* (2005) **37**:239–261.

[2] Cheny, Y. and Botella, O. The LS-STAG method: A new immersed boundary/level-set method for the computation of incompressible viscous flows in complex moving geometries with good conservation properties. *J. Comput. Phys.* (2010) **229**:1043–1076.

[3] Osher, S. and Fedkiw, R.P. *Level set methods and dynamic implicit surfaces.* Springer, (2003).

[4] Spalart, P.R. Strategies for turbulence modelling and simulations. *Int. J. Heat and Fluid Flow.* (2000). **21**: 252–263.

[5] Spalart, P.R. and Allmaras, S.R. A one-equation turbulence model for aerodynamic flows. *Recherche Aerospatiale.* (1994). **1**: 5–21.

[6] Shur, M., Spalart, P., Strelets, M. and Travin, A. Navier-Stokes simulation of shedding turbulent flow past a circular cylinder and a cylinder with backward splitter plate. *Comp. Fluid Dyn.* (1996). **96**: 676–682.

[7] Van der Vorst, H.A. Bi-CGSTAB: a fast and smoothly converging variant of Bi-CG for solution of non-symmetric linear systems. *SIAM J. Sci. Stat. Comp.* (1992). **2**:631–644.

[8] Wesseling, P. *An introduction to multigrid methods.* John Willey & Sons Ltd., (1991).

[9] Marchevsky, I.K. and Puzikova, V.V. OpenFOAM iterative methods efficiency analysis for linear systems solving. *Proceedings of the Institute for System Programming of RAS.* (2013) **24**:71–86. [in Russian]

[10] Zdravkovich, M.M. *Flow around circular cylinders.* Oxford University Press, (1997).

[11] Breuer, M. Large Eddy Simulation of the subcritical flow past a circular cylinder: numerical and modelling aspects. *Int. J. Numer. Meth. Fluids.* (1998). **28**: 1281–1302.

[12] Rahman, M.M., Karim, M.M. and Alim, M.A. Numerical investigation of unsteady flow past a circular cylinder using 2-D finite volume method. *J. Naval Arch. and Marine Eng.* (2007). **4**: 27–42.

[13] Patel, Y. *Numerical Investigation of flow past a circular cylinder and in a staggered tube bundle using various turbulence models. Master's thesis.* Lappeenranta University of Technology, (2010).

[14] Puzikova, V.V. and Marchevsky, I.K. Application of the LS-STAG immersed boundary method for numerical simulation in coupled aeroelastic problems. *Proceedings of 11th World Congress on Computational Mechanics (WCCM XI), 5th European Conference on Computational Mechanics (ECCM V), 6th European Conference on Computational Fluid Dynamics (ECFD VI).* (20–26 July 2014) :1995–2006.

[15] Klamo, J.T., Leonard, A. and Roshko, A. On the maximum amplitude for a freely vibrating cylinder in cross flow. *J. of Fluids and Struct.* (2005) **21**:429–434.

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis(Eds)

ON IN-SITU VISUALIZATION FOR STRONGLY COUPLED PARTITIONED FLUID-STRUCTURE INTERACTION

OLIVER FERNANDES*, DAVID S. BLOM†, STEFFEN FREY*, ALEXANDER H. VAN ZUIJLEN†, HESTER BIJL† AND THOMAS ERTL*

*VISUS
University of Stuttgart
Allmandring 19, 70569 Stuttgart, Germany
e-mail: info@visus.uni-stuttgart.de, web page: http://www.visus.uni-stuttgart.de

†Delft University of Technology, Faculty of Aerospace Engineering
Kluyverweg 2, 2629HT Delft, The Netherlands
e-mail: d.s.blom@tudelft.nl, web page: http://www.lr.tudelft.nl

Key words: in-situ visualization, fluid-structure interaction, multi-physics

Abstract. We present an integrated in-situ visualization approach for partitioned multi-physics simulation of fluid-structure interaction. The simulation itself is treated as a black box and only the information at the fluid-structure interface is considered, and communicated between the fluid and solid solvers with a separate coupling tool. The visualization of the interface data is performed in conjunction with the fluid solver. Furthermore, we present new visualization techniques for the analysis of the interrelation of the two solvers , with emphasis on the involved error due to discretization in space and time and the reconstruction. Our visualization approach also enables the investigation of these errors with respect of their mutual influence on the two simulation codes and their space-time discretization. For efficient interactive visualization, we employ the concept of explorable spatiotemporal images, which also enables finite-time temporal navigation in an in-situ context. We demonstrate our overall approach and its utility by means of a fluid-structure simulation using OpenFOAM that is coupled by the preCICE software layer.

1 INTRODUCTION

Scientific computing and parallel visualization are at the cusp of the new era of exascale computing. In-situ visualization plays a key role in this transition, since the involved data rates are expected to be beyond the capabilities of future storage systems [23]. A central difficulty with in-situ visualization in parallel computing is to conform to bandwidth constraints. On the one hand, parallel scientific computing requires very high bandwidth,

e.g., for the communication between different simulation processes, on the other hand, high-resolution visualization at high frame rates is a prerequisite for successful research and development. While many simulations are intrinsically global and require substantial parallelization traffic, there is large potential for optimization in the visualization stage.

Since the available bandwidth typically does not allow for the transmission of the entire data stream to dedicated visualization nodes, the data reduction has to take place directly at the simulation nodes. In this paper, we present an approach to minimize the bandwidth load exerted by producing a view-dependent reduced representation of the data, before sending it to the visualization nodes.

For the scalar fields associated with the fluid, a common approach is to generate images using volume raycasting. In more detail, rays are sent from the viewpoint through the volume (typically one per pixel in the final image). Along each ray, scalar sample values are taken, color and opacity are assigned to these values by means of a so-called transfer function, and the final color of a ray (or a pixel) is generated by compositing these color values (e.g., [16]). In a coupled fluid-structure simulation, several data sets are available for analysis, and need to be handled by the visualization. To be able to get a concise view of the phenomena arising from interactions between the fluid and the structure, data from both solvers need to be integrated into the same visualization. This provides the possibility to better explore the mutual influence of the two simulation domains.

In this paper, we introduce an integrated approach for the in-situ visualization of both the volume stemming from the fluid solver and the structure surface. We then calculate the deviation introduced by the relaxation step and use it to visualize the the magnitude of the displacement on the representation generated for the structure. This enables the examination of how the quasi-Newton step impacts the calculation of the next time step for the different locations of the fluid-structure interface.

2 RELATED WORK

In case an incompressible flow is considered, sub-cycling of the fluid and solid solvers is necessary in order to minimize the introduced error at the fluid-structure interface due to the partitioning. Different coupling techniques can be applied, such as Aitken's method [19], vector extrapolation [20], Interface-GMRES(R) [27, 26] , and the interface quasi-Newton inverse least squares (IQN-ILS) technique [9, 8]. The IQN-ILS technique [9] is an efficient [10, 8] and robust black-box coupling algorithm for which convergence theorems are available in [17], and is consequently used in this contribution.

For in-situ visualization, typically reduced representations of the original data are generated, usually resulting in both reduced cost for storage/transfer as well as for rendering in comparison to the original. For this purpose, image-based or view-dependent techniques are often used as they allow for high compression rates while still maintaining high quality for a certain range of view configurations. Image-based rendering infers new images from existing ones, e.g., with changed lighting or camera configuration [36]. A number of techniques has been proposed to construct such different representations from

multiple views, like view-dependent texture maps [7], warping [24], light fields [21], etc. Other techniques use multiple images to synthesize new surface-based views of volume data (e.g., [4, 3]). Such techniques allow the adaptation of color and lighting parameters [18], or transfer functions [35].

Shade et al. [30] introduced LDIs (Layered Depth Images) that represent one camera view with multiple pixels along each line of sight. Reconstruction of geometric surfaces from LDIs has been discussed by Frey et al. [14]. In volume rendering, layer-based representations have been used to defer operations such as lighting and volume classification [29]. Such representations have also proven effective to cache results [22] or certain volumetric properties along view rays that can be later reused for efficient transfer function exploration. Also for deferred transfer function exploration, Tikhanova et al. [32] convert a small number of volume renderings to a multi-layered image representation. In another work, [33] use an intermediate volume data representation which encodes the distribution of samples along each ray. Shareef et al. [31] use image-based modeling to render unstructured grids based on parallel sampling rays and 2D texture slicing.

Frey et al. [15] propose Volumetric Depth Images (VDI) for view-dependent volume visualization as an extension of the Layered Depth Image (LDI) approach. Instead of only saving for each view ray of one camera configuration the depth and color values for a set of surfaces, as in LDIs, VDIs store so-called supersegments, each consisting of a depth range as well as composited color and opacity. On this basis, Fernandes et al. [13] introduce space-time VDIs (STVDIs) that provide an efficient space-time representation of VDI streams by exploiting inter-ray and inter-frame coherence, while still maintaining spatiotemporal flexibility. STVDIs provide the technical foundation for the approach proposed in the paper at hand and is outlined in more detail in Sec. 4.

3 SIMULATION SETUP

3.1 Fluid solver

The flow is governed by the incompressible Navier-Stokes equations. Therefore, the governing equations for the fluid are given by the balance of momentum

$$\rho^f \frac{\partial \boldsymbol{v}^f}{\partial t} + \rho^f \left(\nabla \boldsymbol{v}^f\right) \boldsymbol{v}^f = \nabla \cdot \boldsymbol{\sigma}^f, \tag{1}$$

and the balance of mass, which reduces to a divergence free constraint on the velocity field due to incompressibility:

$$\nabla \cdot \boldsymbol{v}^f = 0 \qquad \text{in } \Omega^f, \tag{2}$$

where the velocity field is denoted by $\boldsymbol{v}^f$, the pressure field is denoted by p^f, and ρ^f represents the density. A Newtonian fluid is considered, which leads to the constitutive equation for the stress tensor $\boldsymbol{\sigma}^f$

$$\boldsymbol{\sigma}^f = -p^f \boldsymbol{I} + \rho^f \nu^f \left(\nabla \boldsymbol{v}^f + \nabla \boldsymbol{v}^{f^T}\right), \tag{3}$$

for a given kinematic viscosity ν^f.

The foam-extend-3.1 package[1] is used for the simulations, a fork of the well known OpenFOAM package[2]. A coupled solution algorithm [5] is employed, instead of the well known PISO pressure-velocity coupling technique. Here, the continuity and momentum equation are solved in a fully coupled implicit manner, instead of a segregated approach. A second order backward differencing scheme (BDF2) is used to integrate the governing equations in time. The fluid mesh is deformed with radial basis function interpolation [6].

3.2 Structure solver

The configuration of the structure domain is described by the displacement $\boldsymbol{u}^s$. An elastic and compressible structure is assumed, and the governing equation is given by the balance of momentum

$$\rho^s \frac{\partial \boldsymbol{v}^s}{\partial t} + \rho^s \left(\nabla \boldsymbol{v}^s\right) \boldsymbol{v}^s = \nabla \cdot \boldsymbol{\sigma}^s + \rho^s \boldsymbol{g} \qquad \text{in } \Omega^s. \tag{4}$$

Equation (4) is modified to use the total Lagrangian description, i.e. with respect to the initial reference state Γ^s, resulting in

$$\rho^s \frac{\partial^2 \boldsymbol{u}^s}{\partial t^2} = \nabla \cdot \left(J \boldsymbol{\sigma}^s \boldsymbol{F}^{-T}\right) + \rho^s \boldsymbol{g} \qquad \text{in } \Omega^s, \tag{5}$$

where the deformation gradient tensor $\boldsymbol{F}$ is defined as $\boldsymbol{F} = \boldsymbol{I} + \nabla \boldsymbol{u}^s$, and the Jacobian J is the determinant of the deformation gradient tensor $\boldsymbol{F}$. By applying the constitutive law for the St. Venant-Kirchhoff material, the Cauchy stress tensor $\boldsymbol{\sigma}^s$ is found by applying

$$\boldsymbol{\sigma}^s = \frac{1}{J} \boldsymbol{F} \left(\lambda^s \left(\operatorname{tr} \boldsymbol{E}\right) \boldsymbol{I} + 2\mu^s \boldsymbol{E}\right) \boldsymbol{F}^T, \tag{6}$$

with $\boldsymbol{E} = \frac{1}{2}\left(\boldsymbol{F}^T \boldsymbol{F} - \boldsymbol{I}\right)$, and the shear modulus μ^s [1].

The structure solver is also implemented within the foam-extend-3.1 framework for ease of implementation. Therefore, a finite volume discretization is used instead of a finite element approach.

3.3 Fluid-structure interaction

With regards to the fluid-structure interaction problem, the fluid solver and solid solver are considered as black boxes. In other words, only the input and output information is accessible. Whether a compressible or incompressible, viscous or inviscid flow is considered does not influence the used methodology described in this paper. Also, different models

[1] `http://www.extend-project.de/`

[2] `http://www.openfoam.org/`

for the solid domain can be applied, since only the input and output information from the fluid solver and solid solver is considered to be accessible.

Therefore, at each time step the response of the fluid solver F_f is defined as

$$\boldsymbol{y} = F_f\left(\boldsymbol{x}\right), \tag{7}$$

where $\boldsymbol{x}$ denotes the displacement of the fluid-structure interface, and $\boldsymbol{y}$ denotes the force acting on the fluid-structure interface. The response of the structure solver F_s is consequently defined as

$$\boldsymbol{x} = F_s\left(\boldsymbol{y}\right). \tag{8}$$

Typically, at every time step the fixed point equation

$$\boldsymbol{x} = F_s \circ F_f\left(\boldsymbol{x}\right) \tag{9}$$

must be satisfied, which can also be written as the interface residual $\boldsymbol{R}$

$$\boldsymbol{R}\left(\boldsymbol{x}\right) = F_s \circ F_f\left(\boldsymbol{x}\right) - \boldsymbol{x}, \tag{10}$$

which is solved with a minimization or optimization procedure aimed to find the optimal solution $\boldsymbol{x}^\star$ such that

$$\boldsymbol{x}^\star = \arg\min_{\boldsymbol{x}} \left|\left|\boldsymbol{R}\left(\boldsymbol{x}\right)\right|\right|_2. \tag{11}$$

At the fluid-structure interaction $\Gamma^{f\,s}$, the balance of stresses is enforced through

$$\boldsymbol{\sigma}^f \boldsymbol{n} = \boldsymbol{\sigma}^s \boldsymbol{n} \qquad \text{on } \Gamma^{f\,s}, \tag{12}$$

with the unit vector $\boldsymbol{n}$ normal to the fluid-structure interface $\Gamma^{f\,s}$, and the stress tensors $\boldsymbol{\sigma}^f$ and $\boldsymbol{\sigma}^s$. Also, the no-slip condition is imposed at the fluid-structure interface so that the velocities $\boldsymbol{v}^f$ and $\boldsymbol{v}^s$ must be equal:

$$\boldsymbol{v}^f = \boldsymbol{v}^s \qquad \text{on } \Gamma^{f\,s}. \tag{13}$$

As mentioned in the introduction, standard approaches used to solve the strongly coupled fluid-structure interaction problem are the Gauss-Seidel method [12], fixed under-relaxation [2], Aitken under-relaxation [28], and the IQN-ILS method [9, 11]. Here, the IQN-ILS method is used to solve the interface problem.

4 IN-SITU VISUALIZATION OF COUPLED FLUID-STRUCTURE INTERACTION

In this paper, we use a typical in-situ visualization setup consisting of a parallel simulation environment with integrated visualization. The fluid simulation nodes generate the data, employing domain decomposition for parallel execution. On each compute node a reduced representation of the density scalar field is produced and prepared to be sent to the visualization nodes. Finally, the visualization nodes generate images from these reduced representations interactively upon user request. In this section, we first briefly outline the underlying VDI and STVDI [15, 13] (Sec. 4.1), and on this basis outline our approach for the visualization of interaction properties (Sec. 4.2).

4.1 Space-Time Volumetric Depth Images

A volumetric depth image (VDI) is a representation of a three-dimensional scalar dataset after applying a color transfer function [15]. It is generated internally from the data volume as viewed from a specific vantage point, but can be actually explored in full 3D, for all parts of the simulation data which were in the original view frustum. Being generated from a specific view point allows the user to choose more intuitively which parts of the volume he wants to investigate, and restrict data retrieval to the chosen subset.

Volumetric Depth Images Data Structure. A VDI is produced by ray casting the dataset along rays aimed at the region of interest within the dataset. In the following, we assume that rays are uniformly seeded on a two dimensional grid with the size $X \times Y$. Along each ray r_{xy}, where xy denotes its position on the regular grid, a number of Z samples are taken (typically spaced equidistantly), from front to back. The color transfer function is then applied to associate a scalar value with a user-defined color and opacity. Strictly speaking, a sample depicts the value belonging to a value of a distinct point in space in the volume. However, in the context of raycasting and a limited sampling rate, it can also be seen as representing a certain volume around this point. Conceptually, this is the extension of a pixel in image space to object space.

Storing each segment individually would produce a resampled representation R of the size $X \times Y \times Z$ for the volume covered by the rays. Instead, to reduce the size of such a representation, samples having similar color are merged along a ray into so-called supersegments. This can be done efficiently on-the-fly during while tracing the ray along Z. For these supersegments, the starting and ending parameter value on the ray r_{xy} are stored, as well as the merged color. This produces a list L_{xy} of supersegments for each ray, each with variable length, depending on how the algorithm was able to find mergeable samples. As a result, the size of the VDI representation R has been reduced considerably with respect to the original data, with only minimal loss (depending on the similarity metric), to a set of $X \times Y$ lists each containing $|L_{xy}|$ elements.

Space-Time Clustering. VDIs already reduce the data required to represent the simulation data of interest considerably. This can further be improved upon by exploiting coherency of data across rays, and if considering additional consecutive timesteps, even time. To achieve this, the supersegments for all rays are grouped to form so called space-time regions. Similar to the merging of segments into supersegments for VDIs, this is done by comparing the color of neighboring supersegments ,where neighboring means in adjacent rays or the same ray from an adjacent timestep. In detail, adjacent supersegments that are considered similar by a user defined color metric are assigned the same region id. A representative average color is calculated for an entire region and used in the following.

At this point, already a huge data reduction has been achieved, since the color data, originally stored per supersegment, only needs to be stored per region. Note that the loss in quality due to this can be controlled by the user by changing the parameters in the similarity metric. To also reduce the data which needs to be stored to represent

the locations of supersegments, we employ an algorithm which recalculates and sorts the geometrical data such that it can efficiently be encoded to a stream by an entropy encoder such as the zlib scheme. For more details on this algorithm, refer to the related work [13].

Rendering. On the visualization node, the stream is again decoded using the reverse process explained in Sec. 4.1. All supersegments in a region receive the region's color. The reconstructed VDI representation is now used to define frustums. For each supersegment a frustum is constructed by using the starting and ending point in depth along the generating ray, and using the distance between rays as width and height. These frustums then cover the entire original viewing volume, and are a good approximation to the original simulation data. Unlike a ray cast, these new representation can be built from comparably sparse data and be rendered at much higher speeds (see [15] for details).

4.2 Structural Data and Visualization of Interaction Properties

Many nodes computing the fluid domain solution also produce data for the structure, in form of their boundary. This boundary can be directly converted into a renderable mesh. Being only a two dimensional quantity described as a set of points and and a connecting topology, the structure surface description needs far less data to be transferred, and hence does not any special processing to reduce bandwidth load. To display information concerning the interaction at the coupling boundary of the fluid and the structure solver, we use the fluid solver boundary to define a surface representation of the structure.
On this surface, boundary related data can be displayed, which becomes interesting if the data has a non-uniform distribution on the 2D manifold of the structure surface.

As an example of a data varying across the surface, the deformation of the structure could be examined. The deformation consists of several factors involving both solvers and the coupling mechanism. With many different factors involved, approximations have to be made to accommodate both the numerics and the physics.
As mentioned in Sec. 3.3, a coupling scheme needs to be applied to the structure solvers solution, in order to be able to simulate an incompressible flow without suffering from instabilities due to the added mass effect.
The displacement introduced has an impact on the fluid solvers next iteration. To judge the deviation introduced by relaxation, we now consider the displacement of the surface before and after the relaxation step, and display the magnitude of the displacement vector for the surface of the structure.

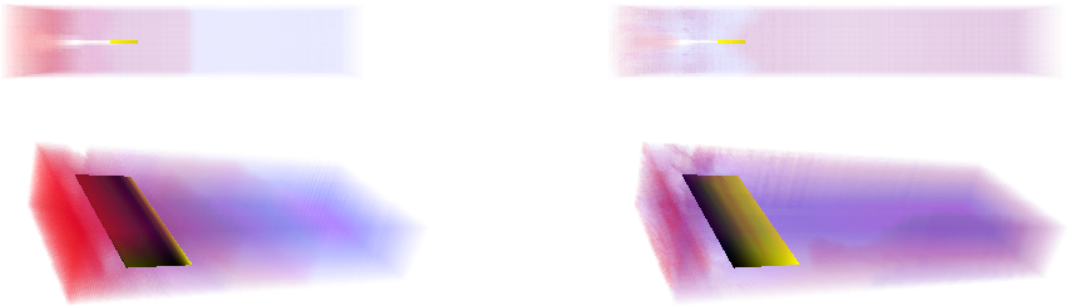
5 RESULTS

We demonstrate the integration of our technique in a distributed simulation setup, as already outlined briefly in 4 and 4.2.

The test case, originally proposed in [34], consists of a two-dimensional incompressible laminar flow around a fixed cylinder with an attached flexible cantilever. In [34], three different scenarios are presented with different fluid to structure density ratios. Here, the

fluid to structure ratio is set to $\rho^f/\rho^s = 1$, resulting in a strong coupling between the fluid and the structure. The reader is referred to [34] for further details on this fluid-structure interaction problem. An overview of the geometry is given by Figure 3.

In detail, the simulation outputs new time steps of the simulation in certain intervals (e.g., every 1000 time steps). This output is then directly fed to the VDI generation and spatiotemporal encoding/compression pipeline. This condensed representation is then communicated to our visualization node, that generates images from this representation. The table 2 shows the compression ratios determined during the simulation using the space-time VDI method. A good compression ratio of around 20% was achieved on average. The data sizes are compared to the output as written out for restarts by the solver.



(a) Timestep 0.1 (b) Timestep 3.3

Figure 1: Volume rendering of the pressure in cylinder flap testcase. The yellow color on the flap shows the distribution of the correction made by relaxation.

For the evaluation of the deviation visualization, we compared the output for the structure including the deviation, using several timesteps. The static parts of the structure (the cylinder) have not been displayed to reduce visual clutter. One can clearly see the spatial distrubution across the moveable surface, as shown in Fig. 1.

Timestep	Size (kBytes)	Ratio	CPU Time (s)
0.1	22	6.7%	4.2
1.4	53	16.3%	4.3
3.3	63	19.2%	4.3

Figure 2: The STVDI algorithm achieves a good compression ratio for different timesteps (simulation time). The timings refer to the slowest process performing the algorithm for a given timestep. Original data size: 326kByte.

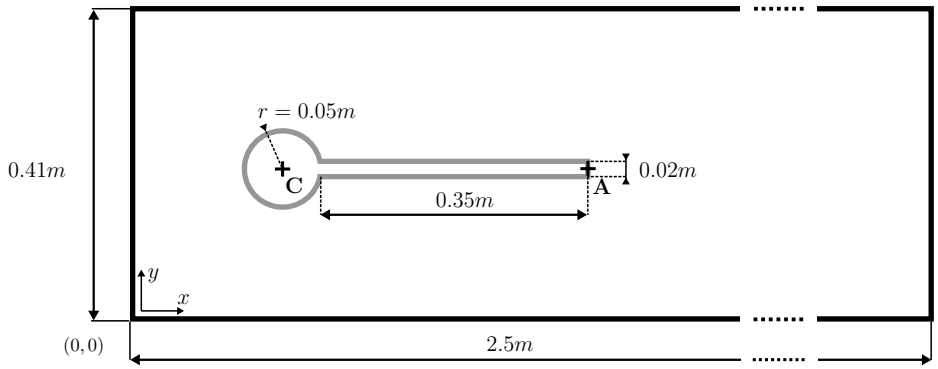


Figure 3: Geometry of the cylinder with an attached flap benchmark originally proposed in [34]. The figure shows a channel flow with a fixed cylinder and a flexible flap attached to the cylinder [25].

6 CONCLUSION

In this paper, we presented an integrated in-situ visualization approach for partitioned multi-physics simulation of fluid-structure interaction. In this context, we introduced new visualization techniques for the analysis of the interrelation of the two solvers, with emphasis on the involved error due to discretization in space and time. We showed that considerable compression can be gained by clustering in space-time and rearranging data for entropy encoding. For this, we employed the concept of explorable spatiotemporal images. Our visualization approach also enables the investigation of errors on the coupling interface due to a quasi-Newton relaxation step. We evaluated our overall approach with a fluid-structure simulation using an OpenFOAM implementation which allows for coupling by the preCICE software layer.

This method can also be extended to other metrics arising in the context of coupled systems, for example displaying coupling-layer internal interpolation error as part of the fluid-structure visualization. Further, the internal properties of the structure itself, as calculated in the structure solver, could be included to achieve a concise visualization.

REFERENCES

[1] P. Cardiff, A. Karač, and A. Ivanković. A large strain finite volume method for orthotropic bodies with general material orientations. *Computer Methods in Applied Mechanics and Engineering*, 268:318–335, January 2014.

[2] M. Cervera, R. Codina, and M. Galindo. On the computational efficiency and implementation of block-iterative algorithms for nonlinear coupled problems. *Engineering Computations*, 13(6):4–30, 1996.

[3] Baoquan Chen, Arie Kaufman, and Qingyu Tang. Image-based rendering of surfaces from volume data. In *Eurographics conference on Volume Graphics*, VG'01, pages 281–300. Eurographics Association, 2001.

[4] Jae-Jeong Choi and Yeong-Gil Shin. Efficient image-based rendering of volume data. In *Computer Graphics and Applications, 1998. Pacific Graphics '98. Sixth Pacific Conference on*, pages 70–78, 226, 1998.

[5] M. Darwish, I. Sraj, and F. Moukalled. A coupled finite volume solver for the solution of incompressible flows on unstructured grids. *Journal of Computational Physics*, 228(1):180–201, 2009.

[6] A. de Boer, A. H. van Zuijlen, and H. Bijl. *Radial Basis Functions for Interface Interpolation and Mesh deformation*, volume 71 of *Lecture Notes in Computational Science and Engineering*, chapter 6, pages 143–178. Springer Berlin Heidelberg, 2010.

[7] Paul E. Debevec, Camillo J. Taylor, and Jitendra Malik. Modeling and rendering architecture from photographs: a hybrid geometry- and image-based approach. In *23rd annual conference on Computer graphics and interactive techniques*, SIGGRAPH '96, pages 11–20, 1996.

[8] J. Degroote. Partitioned Simulation of Fluid-Structure Interaction. *Archives of Computational Methods in Engineering*, 20(3):185–238, September 2013.

[9] J. Degroote, K.-J. Bathe, and J. Vierendeels. Performance of a new partitioned procedure versus a monolithic procedure in fluid-structure interaction. *Computers and Structures*, 87(11-12):793–801, June 2009.

[10] J. Degroote and J. Vierendeels. Multi-solver algorithms for the partitioned simulation of fluid-structure interaction. *Computer Methods in Applied Mechanics and Engineering*, 200(25-28):2195–2210, June 2011.

[11] J. Degroote and J. Vierendeels. Multi-level quasi-Newton coupling algorithms for the partitioned simulation of fluid-structure interaction. *Computer Methods in Applied Mechanics and Engineering*, 225-228:14–27, 2012.

[12] C. Farhat and M. Lesoinne. Two efficient staggered algorithms for the serial and parallel solution of three-dimensional nonlinear transient aeroelastic problems. *Computer Methods in Applied Mechanics and Engineering*, 182(3-4):499–515, February 2000.

[13] Oliver Fernandes, Steffen Frey, Filip Sadlo, and Thomas Ertl. Space-time volumetric depth images for in-situ visualization. In *Large Data Analysis and Visualization (LDAV), 2014 IEEE 4th Symposium on*, pages 59–65. IEEE, 2014.

[14] S. Frey, S. Filip, and T. Ertl. Mesh Generation From Layered Depth Images Using Isosurface Raycasting. In *ISVC '13: Proceedings of the 9th International Symposium on Advances in Visual Computing*, pages 373–383, Berlin, Heidelberg, 2013. Springer-Verlag.

[15] S. Frey, F. Sadlo, and T. Ertl. Explorable volumetric depth images from raycasting. In *Proceedings of the Conference on Graphics, Patterns and Images*, pages 123–130, Aug 2013.

[16] Markus Hadwiger, Patric Ljung, Christof Rezk Salama, and Timo Ropinski. Advanced illumination techniques for gpu volume raycasting. In *ACM SIGGRAPH ASIA 2008 courses*, SIGGRAPH Asia '08, pages 1:1–1:166, New York, NY, USA, 2008. ACM.

[17] Rob Haelterman, J. Degroote, Dirk van Heule, and J. Vierendeels. The Quasi-Newton Least Squares Method: A New and Fast Secant Method Analyzed for Linear Systems. *SIAM Journal on Numerical Analysis*, 47(3):2347–2368, 2009.

[18] Taosong He, Lichan Hong, Arie Kaufman, and Hanspeter Pfister. Generation of transfer functions with stochastic search techniques. In *Visualization*, pages 227–234, 1996.

[19] Ulrich Küttler and Wolfgang A. Wall. Fixed-point fluid–structure interaction solvers with dynamic relaxation. *Computational Mechanics*, 43(1):61–72, 2008.

[20] Ulrich Küttler and Wolfgang A. Wall. Vector Extrapolation for Strong Coupling Fluid-Structure Interaction Solvers. *Journal of Applied Mechanics*, 76(2), J. 2009.

[21] Marc Levoy and Pat Hanrahan. Light field rendering. In *23rd annual conference on Computer graphics and interactive techniques*, SIGGRAPH '96, pages 31–42, 1996.

[22] Eric J. Luke and Charles D. Hansen. Semotus visum: a flexible remote visualization framework. In *Proceedings of IEEE Visualization '02*, pages 61–68, 2002.

[23] Kwan-Liu Ma. In situ visualization at extreme scale: Challenges and opportunities. *IEEE Comput. Graph. Appl.*, 29(6):14–19, November 2009.

[24] Leonard McMillan and Gary Bishop. Plenoptic modeling: an image-based rendering system. In *22nd annual conference on Computer graphics and interactive techniques*, SIGGRAPH '95, pages 39–46, 1995.

[25] Miriam Mehl, Benjamin Uekermann, H. Bijl, David Stephen Blom, Bernhard Gatzhammer, and A. H. van Zuijlen. Parallel coupling numerics for partitioned fluid-structure interaction simulations. *SIAM Scientific Computing*, submitted 2013.

[26] C. Michler, E. H. van Brummelen, and R. de Borst. An interface Newton–Krylov solver for fluid–structure interaction. *International Journal for Numerical Methods in Fluids*, 47(10-11):1189–1195, April 2005.

[27] Christian Michler, Harald van Brummelen, and René de Borst. An investigation of Interface-GMRES(R) for fluid–structure interaction problems with flutter and divergence. *Computational Mechanics*, 47(1):17–29, January 2011.

[28] D. P. Mok, W. A. Wall, and E. Ramm. Accelerated iterative substructure schemes for instationary fluid-structure interaction. In *First MIT Conference on Computational Fluid and Solid Mechanics*, pages 1325–1328, 2001.

[29] T. Ropinski, J. Prassni, F. Steinicke, and K. Hinrichs. Stroke-based transfer function design. In *IEEE/EG International Symposium on Volume and Point-Based Graphics*, pages 41–48, 2008.

[30] Jonathan Shade, Steven Gortler, Li-Wei He, and Richard Szeliski. Layered depth images. In *Proceedings on Computer Graphics and Interactive Techniques*, pages 231–242, 1998.

[31] Naeem Shareef, Teng-Yok Lee, Han-Wei Shen, and Klaus Mueller. An image-based modeling approach to GPU-based unstructured grid volume rendering. *Proceedings of Volume Graphics*, pages 31–38, 2006.

[32] A. Tikhonova, C. Correa, and Kwan-Liu Ma. Explorable images for visualizing volume data. In *IEEE Pacific Visualization Symposium*, pages 177–184, 2010.

[33] A. Tikhonova, C. D. Correa, and K.-L. Ma. An exploratory technique for coherent visualization of time-varying volume data. In *Proceedings of the Eurographics conference on Visualization*, pages 783–792, 2010.

[34] Stefan Turek and Jaroslav Hron. Proposal for Numerical Benchmarking of Fluid-Structure Interaction between an Elastic Object and Laminar Incompressible Flow. In Hans-Joachim Bungartz and Michael Schäfer, editors, *Fluid-Structure Interaction*, volume 53 of *Modelling, Simulation, Optimisation*, pages 371–385. Springer Berlin Heidelberg, 2006.

[35] Yingcai Wu and Huamin Qu. Interactive transfer function design based on editing direct volume rendered images. *Visualization and Computer Graphics, IEEE Trans. on*, 13(5):1027–1040, 2007.

[36] Heung yeung Shum and Sing Bing Kang. A survey of image-based rendering techniques. In *In Videometrics, SPIE*, pages 2–16, 1999.

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis(Eds)

ON PARALLEL SCALABILITY ASPECTS OF STRONGLY COUPLED PARTITIONED FLUID-STRUCTURE-ACOUSTICS INTERACTION

DAVID S. BLOM*, VERENA KRUPP†, ALEXANDER H. VAN ZUIJLEN*, HARALD KLIMACH†, SABINE ROLLER†, AND HESTER BIJL*

*Delft University of Technology, Faculty of Aerospace Engineering
Kluyverweg 2, 2629HT Delft, The Netherlands
e-mail: d.s.blom@tudelft.nl, web page: http://www.lr.tudelft.nl

†University Siegen, Simulation Techniques & Scientific Computing,
Hölderlinstr. 3, 57076 Siegen, Germany,
e-mail: verena.krupp@uni-siegen.de

Key words: fluid-structure-acoustics interaction, multi-physics, partitioned, parallel

Abstract. Multi-physics simulations, such as fluid-structure-acoustics interaction (FSA), require a high performance computing environment in order to perform the simulation in a reasonable amount of computation time. Currently used coupling methods use a staggered execution of the fluid and solid solver [6], which leads to inherent load imbalances.

In [12] a new coupling scheme based on a quasi-Newton method is proposed for fluid-structure interaction which coupled the fluid and solid solver in parallel. The quasi-Newton method requires approximately the same number of coupling iterations per time step compared to a staggered coupling approach, resulting in a better load balance when running in a parallel environment.

This contribution investigates the scalability limit and load-balancing for a strongly coupled fluid-structure interaction problem, and also for a fluid-structure-acoustics interaction problem. The acoustic far field of the fluid-structure-acoustics interaction problem is loosely coupled with the flow field.

1 INTRODUCTION

Simulation of fluid-structure-acoustics interaction will bring new insight into different applications, as, for example, the sound design of aircraft or wind energy plants. But such applications yield multi-scale problems including different length scales where solving large problems with a monolithic approach would be too expensive.

Also, highly sophisticated software codes are available for each single physical phenomena and it is desirable to reuse currently available codes for multi-physics simulations.

Since different phenomena typically appear spatially separated it is possible to decompose the overall simulation domain into non-overlapping partitions with distinctive treatment and link these partitioned via a surface coupling. This allows the usage of different numerical methods and tailored grid resolution for each individual partitioned. Moreover, each partitions can be solved by individual software codes and we refer to this approach as black-box approach. Additionally, this offers the possibility to benefit from prior experience on how to scale up each single simulation. We are aiming for such flexibility, but the coupling between each pair of physical solvers needs to be carried out carefully, to get a stable overall simulation, while not degrading the scalability.

Whereas for fluid-structure interaction the coupling approach is typically an implicit coupling, which yield same accuracy as a monolithic simulation, for the fluid-acoustic domain with sound generation and propagation several coupling approaches are known. For the flow domain, we are choosing a compressible flow solver, since sound generation implies small perturbations or changes in the density. Far away from the geometry, where only propagation of acoustic waves is relevant, the set of equations can be reduced to the linearized Euler equations. Additionally, the mesh, which needs to be very fine around the geometry, can be much coarser. Another approach would be to use an incompressible flow solver and extract the acoustic sources from the flow as proposed by Lighthill [11]. Note that the acoustic sources need to be extracted and propagated in the flow domain as well.

The immense computational cost of a three-dimensional simulation demands to make efficient use of todays massively parallel supercomputers. A scalable approach which uses the computational resources efficiently and avoid idle processor is indispensable. Load balancing inside one solver is understood very well, but load balancing over individual solvers which are coupled at the boundaries is a different approach.

To achieve such a computational setup, all included components should fulfill the scalability requirements. The high order discontinuous Galerkin solver Ateles, included in the APES framework [14], scales up to more than 100k CPU cores, the limit of currently available super computer facilities. It combines an octree-based data structure with a very local numerical scheme. The used coupling tool preCICE [9], which is responsible for the data exchange as well as for all interpolations between interfaces, is also executed in parallel and minimizes the amount of communication, since it works local on solver processes. OpenFOAM uses a standard domain decomposition approach, since an unstructured mesh is considered. The governing equations are solved in the separate domains and coupled on the boundaries with appropriate boundary conditions.

The presented coupling approach using scalable software allows solving large multi-scale problem on massively parallel machine where the numerical method as well as the computational resources can be perfectly tailored to the physics.

2 SIMULATION SETUP

The computational domain is partitioned into a fluid, structure and acoustics domain. In each domain different governing equations are solved, and also specialized solvers are developed in order to be able to easily switch between different models and software packages. At the fluid-structure interface, and at the fluid-acoustics interface a coupling tool is used to apply the appropriate boundary conditions.

2.1 Fluid solver

The flow is governed by the compressible Navier-Stokes equations. Therefore, the governing equations for the fluid are given by the mass continuity equation

$$\frac{\partial \rho^f}{\partial t} + \nabla \cdot \left(\rho^f \, \boldsymbol{v}\right) = 0, \tag{1}$$

the balance of momentum

$$\frac{\partial \left(\rho^f \, \boldsymbol{v}\right)}{\partial t} + \nabla \cdot \left(\rho^f \, \boldsymbol{v} \, \boldsymbol{v}\right) = -\nabla p^f + \nabla \cdot \boldsymbol{\tau} + \boldsymbol{B}, \tag{2}$$

and the balance of energy equation

$$\frac{\partial}{\partial t}\left(\rho^f \left(e + \frac{1}{2}\boldsymbol{v}\cdot\boldsymbol{v}\right)\right) + \nabla \cdot \left(\rho^f \, \boldsymbol{v} \left(e + \frac{1}{2}\boldsymbol{v}\cdot\boldsymbol{v} + \frac{p^f}{\rho^f}\right)\right) = -\nabla \cdot \boldsymbol{q} + \nabla \cdot \left(\boldsymbol{\tau}\cdot\boldsymbol{v}\right). \tag{3}$$

The velocity field is denoted by $\boldsymbol{v}^f$, the pressure field is denoted with p^f, and the density is given as ρ^f. The viscous stress tensor is given by $\boldsymbol{\tau}$, the internal energy by e, and the heat flux by $\boldsymbol{q}$. The body force is denoted by $\boldsymbol{B}$.

The foam-extend-3.1 package[1] is used for the simulations, a fork of the well known OpenFOAM package[2]. An unstructured mesh is used for the fluid domain, which gives the advantage to easily generate a grid for complex geometries, which is often the case for an engineering fluid-structure interaction test case. A second finite volume discretization is applied in space, and a second order backward differencing time integration scheme is used to propagate the solution in time. Regarding the pressure-velocity coupling, a coupled solution algorithm [4, 3] is employed. The continuity and momentum equations are solved in a fully coupled implicit manner. Thereafter, the energy equations is solved in a segregated manner.

Due to the use of an unstructured mesh, a mesh deformation technique based on radial basis function interpolation is used to interpolate the displacement of the fluid-structure interface into the complete flow field [5].

Coupling an implicit BDF2 time integration scheme for the flow with an explicit second or fourth order Runge Kutta scheme for the acoustics reduces to a first order time

[1] `http://www.extend-project.de/`

[2] `http://www.openfoam.org/`

integration for the acoustic domain. This poses extra restrictions on the used time step in order to have a stable integration.

At the fluid-acoustics interface, an interpolation is performed in order to transfer the pressure, density, and velocities from OpenFOAM to Ateles via the coupling tool preCICE. The ideal situation would be to use conforming or matching grids at the interface. However, due to the coupling of a finite volume solver with a discontinuous Galerkin solver, the meshes are not matching at the interface. Therefore, a radial basis function interpolation is utilized at the interface in order to reduce the introduced numerical errors due to the partitioning as much as possible.

2.2 Structure solver

The configuration of the structure domain is described by the displacement $\boldsymbol{u}^s$. An elastic and compressible structure is assumed, and the governing equation is given by the balance of momentum

$$\rho^s \frac{\partial \boldsymbol{v}^s}{\partial t} + \rho^s \left(\nabla \boldsymbol{v}^s\right) \boldsymbol{v}^s = \nabla \cdot \boldsymbol{\sigma}^s + \rho^s \boldsymbol{g} \qquad \text{in } \Omega^s. \tag{4}$$

Equation (4) is modified to use the total Lagrangian description, i.e. with respect to the initial reference state Γ^s, resulting in

$$\rho^s \frac{\partial^2 \boldsymbol{u}^s}{\partial t^2} = \nabla \cdot \left(J \boldsymbol{\sigma}^s \boldsymbol{F}^{-T}\right) + \rho^s \boldsymbol{g} \qquad \text{in } \Omega^s, \tag{5}$$

where the deformation gradient tensor $\boldsymbol{F}$ is defined as $\boldsymbol{F} = \boldsymbol{I} + \nabla \boldsymbol{u}^s$, and the Jacobian J is the determinant of the deformation gradient tensor $\boldsymbol{F}$. By applying the constitutive law for the St. Venant-Kirchhoff material, the Cauchy stress tensor $\boldsymbol{\sigma}^s$ is found by applying

$$\boldsymbol{\sigma}^s = \frac{1}{J} \boldsymbol{F} \left(\lambda^s \left(\operatorname{tr} \boldsymbol{E}\right) \boldsymbol{I} + 2\mu^s \boldsymbol{E}\right) \boldsymbol{F}^T, \tag{6}$$

with $\boldsymbol{E} = \frac{1}{2}\left(\boldsymbol{F}^T \boldsymbol{F} - \boldsymbol{I}\right)$, and the shear modulus μ^s [1].

The structure solver is also implemented within the foam-extend-3.1 framework for ease of implementation. Therefore, a finite volume discretization is used instead of a finite element approach.

2.3 Acoustics solver

The acoustic phenomena are also governed by the Navier-Stokes equations (1-3). A simplification can be obtained by neglecting the dissipation terms which leads to the inviscid compressible Euler equations. If there are only small changes in the flow, it can be linearized around the base flow. The base flow is denoted by the subscript 0 and the perturbation is denoted with the superscript a. The linearized variables describe the acoustic phenomena in the fluid. Using the assumption of ideal gas

$$p = \rho R T = (\gamma - 1)\left(e - \frac{\rho \mathbf{v} \cdot \mathbf{v}}{2}\right)$$

a relation between pressure p and energy e yields, where R is the ideal gas constant, T the temperature and γ the isentropic coefficient. In the following, we will treat only the primitive variables density ρ, velocity $\boldsymbol{v}$ and pressure p in the acoustic domain. The linearized Euler equations are given by the linearized equation of mass

$$\frac{\partial \rho^a}{\partial t} + \nabla \cdot (\boldsymbol{v}_0 \rho^a + \rho_0 \boldsymbol{v}^a) = 0, \tag{7}$$

the conservation of the velocity perturbation

$$\frac{\partial \boldsymbol{v}^a}{\partial t} + \nabla \cdot \left(\boldsymbol{v}_0 \boldsymbol{v}^a + \frac{1}{\rho_0} p^a \right) = 0 \tag{8}$$

and the conservation of the pressure perturbation

$$\frac{\partial p^a}{\partial t} + \nabla \cdot (\boldsymbol{v}_0 p^a + \gamma \; p_0 \; \boldsymbol{v}^a) = 0. \tag{9}$$

The acoustic partition is simulated using the high order Discontinuous Galerkin solver Ateles which is included in the APES framework. The Discontinuous Galerkin (DG) method is based on a polynomial representation within an element and flux calculation between elements. The choice of the polynomial degree controls the spatial discretization order. By choosing a high degree of the polynomial function a higher order method can be constructed.

A higher order scheme has different advantages. Firstly, it yields low numerical dissipation and dispersion errors, which is advantageous for approximating the wave propagation over long distances in the acoustic far field. Secondly, it shows high convergence rates in case the solution is smooth. Therefore, a high order scheme uses less elements while still maintaining the same accuracy compared to a second order finite volume method. An explicit second order Runge-Kutta method is used to propagate the solution in time.

The APES [14] framework, in which the solver in included, provides additional tools for pre- and post-processing on the basis of the common mesh library TreElM[3]. The TreElM library [10] relies on an octree representation of the mesh and provides the distributed neighborhood search within that mesh. The APES framework is designed to take advantage of the massively parallel systems available in supercomputing today. Using a space-filling curve for the domain composition of the octree mesh, gives hierarchically structured data and maintains locality. This locality can be perfectly exploited by the DG scheme, which is strongly coupled to data within one element and loosely coupled via element boundaries. Hence, the framework and solver are designed to take advantage of the massively parallel systems available in supercomputing today. By free choices of the spatial scheme order and the hybrid parallelism, the solver can be adapted to the executing machine.

[3] `https://bitbucket.org/apesteam/treelm`

2.4 Fluid-structure interaction

With regards to the fluid-structure interaction problem, the fluid solver and solid solver are considered as black boxes as mentioned in the introduction. In other words, only the input and output information is accessible. Note that the used model for the solid domain can be changed without changing the setup of the fluid-structure interaction problem, since only the input and output information from the fluid solver and solid solver is considered to be accessible.

Therefore, at each time step the response of the fluid solver F_f is defined as

$$\boldsymbol{y} = F_f\left(\boldsymbol{x}\right), \tag{10}$$

where $\boldsymbol{x}$ denotes the displacement of the fluid-structure interface, and $\boldsymbol{y}$ denotes the force acting on the fluid-structure interface. The response of the structure solver F_s is consequently defined as

$$\boldsymbol{x} = F_s\left(\boldsymbol{y}\right). \tag{11}$$

Typically, at every time step the fixed point equation

$$\boldsymbol{x} = F_s \circ F_f\left(\boldsymbol{x}\right) \tag{12}$$

must be satisfied, which can also be written as the interface residual $\boldsymbol{R}$

$$\boldsymbol{R}\left(\boldsymbol{x}\right) = F_s \circ F_f\left(\boldsymbol{x}\right) - \boldsymbol{x}, \tag{13}$$

which is solved with a minimization or optimization procedure aimed to find the optimal solution $\boldsymbol{x}^\star$ such that

$$\boldsymbol{x}^\star = \arg\min_{\boldsymbol{x}} \left|\left|\boldsymbol{R}\left(\boldsymbol{x}\right)\right|\right|_2. \tag{14}$$

At the fluid-structure interaction $\Gamma^{f\,s}$, the balance of stresses is enforced through

$$\boldsymbol{\sigma}^f \boldsymbol{n} = \boldsymbol{\sigma}^s \boldsymbol{n} \qquad \text{on } \Gamma^{f\,s}, \tag{15}$$

with the unit vector $\boldsymbol{n}$ normal to the fluid-structure interface $\Gamma^{f\,s}$, and the stress tensors $\boldsymbol{\sigma}^f$ and $\boldsymbol{\sigma}^s$. Also, the no-slip condition is imposed at the fluid-structure interface so that the velocities $\boldsymbol{v}^f$ and $\boldsymbol{v}^s$ must be equal:

$$\boldsymbol{v}^f = \boldsymbol{v}^s \qquad \text{on } \Gamma^{f\,s}. \tag{16}$$

Standard approaches used to solve the strongly coupled fluid-structure interaction problem are the Gauss-Seidel method [8], fixed under-relaxation [2], Aitken under-relaxation [13], and the IQN-ILS method [6, 7]. Here, the IQN-ILS method is used to solve the interface problem.

In case a compressible flow is simulated, an explicit coupling at the fluid-structure interface can also be applied. However, in this contribution the use of an implicit coupling

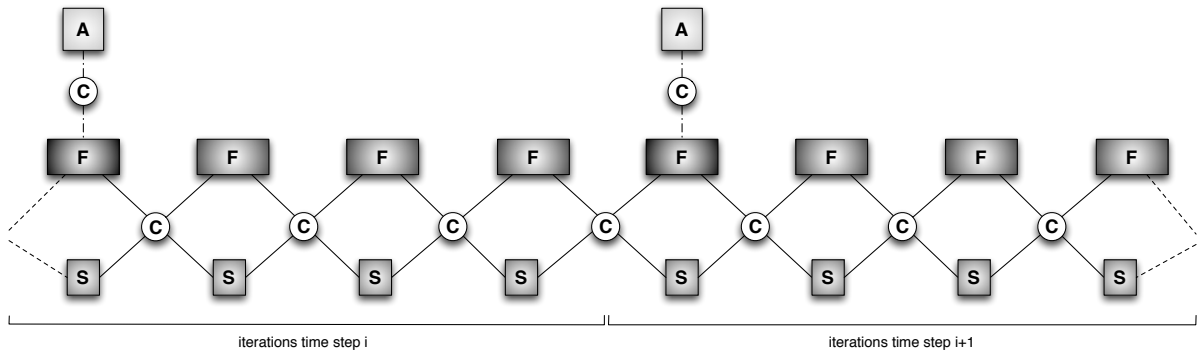


Figure 1: Overview of the execution of the fluid-structure-acoustics simulations. Multiple calls to the fluid and solid solvers are performed, since an implicit coupling is applied for the fluid-structure interaction problem. A good load balancing can be achieved with the proper number of cores for the acoustic domain. F: fluid, S: solid, A: acoustics.

technique is applied in order to remove the introduced partitioning error at the fluid-structure interface.

A distinction is made between the case where the solvers are coupled sequentially, as is generally the case, and in parallel. In [12] a new coupling scheme based on the IQN-ILS method is proposed which coupled the fluid and solid solver in parallel. Hence, a better load balancing is obtained giving the possibility to be able to scale a partitioned fluid-structure interaction further compared to the serially coupled IQN-ILS algorithm.

2.5 Fluid-structure-acoustics interaction

The coupling to the acoustic field is treated with an explicit coupling. In other words, in every time step the primitive variables (pressure, density, velocity) present at the fluid-acoustics interface are transferred via the coupling tool and applied with a boundary condition to the acoustic domain. Thereafter, the linearized Euler equations are solved in the acoustic field.

An overview of the used coupling technique is shown in Figure 1. As shown in the figure, the fluid and solid solvers are executed multiple times per time step. When the FSI iterations are converged, the acoustic solver is called. An optimal load balancing can be obtained by selecting the appropriate number of CPU cores for the acoustic domain, such that the computational time is approximately equal to the time required for the FSI iterations to converge.

3 NUMERICAL RESULTS

The used test case consists of a flexible cantilever in a cross flow. At the outer boundaries the fluid domain is coupled with the acoustic domain. Whereas are th structure-fluid interface is matching, the fluid-acoustic interface differs. In contrast to the fluid domain, which has 124 elements at the interface, the acoustic domain has only 24 elements. Hence,

Table 1: Parallel scalability test for a fluid-structure-acoustics simulation. The fluid, solid and acoustics solvers are coupled in a staggered sense, which leads to an non-optimal load balancing since processors of 2 solvers are always idle, which hurts particularly for the flow solver. The number of fluid cores varies for the computations shown here. The solid and acoustics solver both use one CPU core. 29 time steps are simulated.

No. fluid cores	Exec. time [s]	Timing fluid [s]	Timing solid [s]	Timing acoustics [s]
32	225	187.33	19.37	33.25
64	126	93.74	18.62	35.43
128	80	45.98	18.74	30.99
256	60	22.6	19.75	33.88

Table 2: Parallel scalability tests for a fluid-structure-acoustics simulation. The fluid, solid and acoustics solvers are coupled in parallel, leading to good load-balancing compared to the staggered coupling test case. The number of fluid cores varies for the computations shown here. The solid and acoustics solver both use one CPU core. 29 time steps are simulated.

No. fluid cores	Exec. time [s]	Timing fluid [s]	Timing solid [s]	Timing acoustics [s]
32	225	199.03	20.32	34.94
64	124	99.09	21.77	33.60
128	69	47.97	21.65	33.55
256	47	24.16	20.93	34.87

a fifth order discretization in space is used for the acoustic domain. The pressure field of the test case at $t = 0.2\,s$ is shown in Figure 2.

Table 1 and 2 show the scalability results comparing the use of a serial and parallel coupling scheme for the complete fluid-structure-acoustics interaction setup. The number of cores for the structure and acoustics solvers are kept at one. The number of cores for the fluid is increased from 32 to 256.

The results show the overhead of the communication between the different solvers, and show the efficiency of the parallel coupling schemes. With a proper load balancing, if the solvers are executed in parallel, smaller execution times are obtained in comparison to sequentially coupled solvers. More importantly, this parallel coupling combined with balanced work-load avoid idle processor and can make efficient use of today's supercomputers.

4 CONCLUSIONS

The used methodology has been described for a partitioned fluid-structure-acoustics setup. The fluid-structure interaction problem is coupled implicitly, whereas the fluid and acoustics domain are coupled explicitly. Scaling tests have been performed in order to examine the bottleneck of this approach. The use of a separate coupling tool between the different domains imposes extra communication costs. Further research is dedicated to use a asynchronous communication between the different solver. Also, the use of a

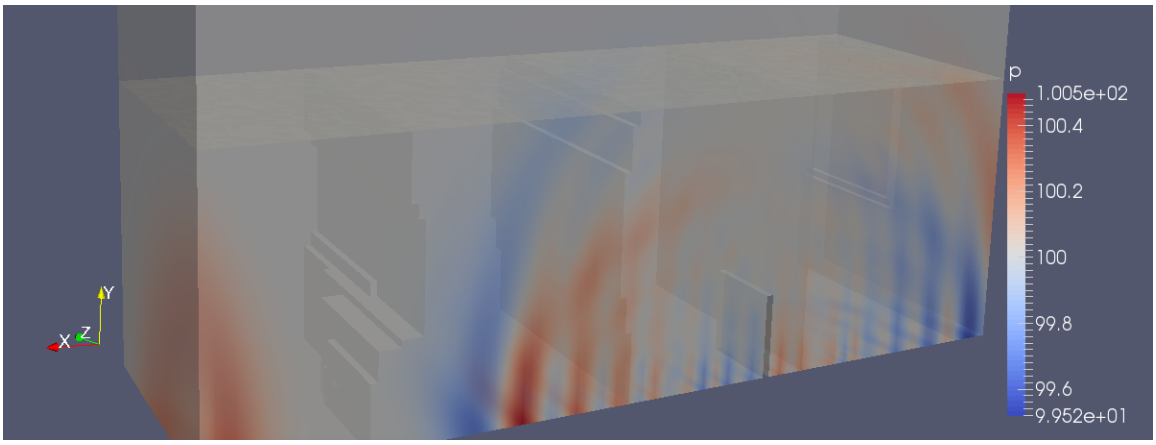


Figure 2: Pressure field for the fluid-structure-acoustics domain at t = 0.2 s. The domain decomposition of the flow is visible in the background. The fluid domain consists of 198 112 cells, the solid domain contains 288 cells, and the acoustic domain consists of 6 144 elements.

spectral deferred correction time integration scheme is considered in order to solve the complete fluid-structure-acoustics domain with a high order time integration scheme.

REFERENCES

[1] P. Cardiff, A. Karač, and A. Ivanković. A large strain finite volume method for orthotropic bodies with general material orientations. *Computer Methods in Applied Mechanics and Engineering*, 268:318–335, January 2014.

[2] M. Cervera, R. Codina, and M. Galindo. On the computational efficiency and implementation of block-iterative algorithms for nonlinear coupled problems. *Engineering Computations*, 13(6):4–30, 1996.

[3] M. Darwish and F. Moukalled. A Fully Coupled Navier-Stokes Solver for Fluid Flow at All Speeds. *Numerical Heat Transfer*, 65(5), 2014.

[4] M. Darwish, I. Sraj, and F. Moukalled. A coupled finite volume solver for the solution of incompressible flows on unstructured grids. *Journal of Computational Physics*, 228(1):180–201, 2009.

[5] A. de Boer, A. H. van Zuijlen, and H. Bijl. *Radial Basis Functions for Interface Interpolation and Mesh deformation*, volume 71 of *Lecture Notes in Computational Science and Engineering*, chapter 6, pages 143–178. Springer Berlin Heidelberg, 2010.

[6] J. Degroote, K.-J. Bathe, and J. Vierendeels. Performance of a new partitioned procedure versus a monolithic procedure in fluid-structure interaction. *Computers and Structures*, 87(11-12):793–801, June 2009.

[7] J. Degroote and J. Vierendeels. Multi-level quasi-Newton coupling algorithms for the partitioned simulation of fluid-structure interaction. *Computer Methods in Applied Mechanics and Engineering*, 225-228:14–27, 2012.

[8] C. Farhat and M. Lesoinne. Two efficient staggered algorithms for the serial and parallel solution of three-dimensional nonlinear transient aeroelastic problems. *Computer Methods in Applied Mechanics and Engineering*, 182(3-4):499–515, February 2000.

[9] B. Gatzhammer, M. Mehl, and T. Weinzierl. A coupling tool for the partitioned simulation of fluid-structure interactions. In T. Kvamsdal, B. Pettersen, P. Bergan, E. Onate, and J. Garcia, editors, *Computational Methods in Marine Engineering*, volume 3, pages 147–150, Trondheim, June 2009. International Center for Numerical Methods in Engineering (CIMNE), CIMNE.

[10] H. Klimach, M. Hasert, J. Zudrop, and S. Roller. Distributed Octree Mesh Infrastructure for Flow Simulations. *European Congress on Computational Methods in Applied Sciences and Engineering*, pages 1–15, 2012.

[11] M.J. Lighthill. On Sound Generated Aerodynamically. I. General Theorie. *Proceedings of The Royal Society London A*, 211:564–587, 1952.

[12] M. Mehl, B. Uekermann, H. Bijl, D. S. Blom, B. Gatzhammer, and A. H. van Zuijlen. Parallel coupling numerics for partitioned fluid-structure interaction simulations. *SIAM Scientific Computing*, submitted 2013.

[13] D. P. Mok, W. A. Wall, and E. Ramm. Accelerated iterative substructure schemes for instationary fluid-structure interaction. In *First MIT Conference on Computational Fluid and Solid Mechanics*, pages 1325–1328, 2001.

[14] J. Zudrop, H. Klimach, M. Hasert, K. Masilamani, and S. Roller. A fully distributed CFD framework for massively parallel systems. In *Cray User Group 2012*, Stuttgart, Germany, April 2012.

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis (Eds)

HOW THE MORPHOLOGIES OF PROGENITOR AND MATURE OSTEOCYTES CONTRIBUTES TO THEIR MECHANO-TRANSDUCTION

E. BUDYN[*], M. BENSIHOUM[†††], T. MARSAN[*], F. MANNEMARE[*], S. SASNOUSKI[†], P. TAUC[†], E. DEPREZ[†] AND H. PETITE[†††]

[*] Department of Mechanical Engineering, LMT CNRS UMR 8535
[†] Department of Biology, LBPA CNRS UMR 8113
Ecole Normale Superieure de Cachan
94230 Cachan, France
email: elisa.budyn@ens-cachan.fr, thibault.marsan@ens-cachan.fr, florian.mannemare@ens-cachan.fr, sergei.sasnouski@ens-cachan.fr, tauc@lbpa.ens-cachan.fr, deprez@lbpa.ens-cachan.fr, http://www.lmt.ens-cachan.fr

[††] Departments of Bioengineering and Orthopedic Surgery – CNRS UMR B2OA 7052
University Paris Diderot
University Paris-Sorbonne, France
email: morad.bensidhoum@paris7.jussieu.fr, herve.petite@univ-paris-diderot.fr, http://b2oa.eu/spip/

Key words: Osteocyte, bone, mechano-transduction, CFSE, h bm MSCs, MLOY4.

Abstract. A dual experimental and numerical top-down approach is applied to investigate the link between the osteocyte morphology and their mechanical perception of the environment at the progenitor and mature stages. The numerical model is based on explicit tissue morphology discretization to identify bone *in situ* diffuse damage at the cellular scale. The 3D morphology of a human mature osteocyte was reconstructed from deconvoluted confocal microscopy observations. The *in vitro* experimental model presents Live Allograft Bone Systems (LABS) where a patient progenitor (bm hMSC) or mature (MLOY4) osteocytes were reseeded into fresh human donor cortical bone tissues. The system was subjected to mechanical loading and simultaneously progenitor and mature cell specific possible calcium mediated cell signaling responses were measured by fluorescent flow cytometry using CFSE labeling.

1 INTRODUCTION

With increasing life expectancy, bone pathologies related to massive bone loss occur later in life and carry $5-$10 billion financial burden on the U.S. healthcare system. Human Haversian cortical bone is a complex hierarchical heterogeneous tissue resulting from continuous remodeling. Micro damage are therefore resorbed by osteoclasts cells before tubular lamellar structures called osteons are formed by osteoblast cells laying Type I collagen fibrils mineralized by hydroxyapatite nano-platelet crystals glued together with non-collagen proteins and proteoglycans. Trapped osteoblasts further differentiate into mechano-sensitive osteocytes that are able to sense stimulation produced by micro damage.

However bone healing ability declines with long term degeneration during aging, massive trauma or large tissue resections such as tumor removals. To promote bone growth in large defects, autograft bone offers the gold standard repair but is limited by suitable tissue quantities and donor site morbidity. Successful techniques for massive tissue regeneration can be however difficult to produce and often require addition of functional materials. Nonetheless, allograft bone can be stored but does not always perform as well as fresh allograft and the current tissue disinfection procedures such as supercritical carbon dioxide significantly modifies the tissue properties. Yet cleaned and decellularized tissue from a donor represents the ideal matrix for co-cultures of the recipient patient cells for fast tissue reintegration and functionalization [1].

Mature osteocytes have the particularity to bear 40 to 60 cytoplasmic processes extending into canaliculi to create a syncytial network with the neighboring cells with which they can transmit signals in a fashion similar to the nervous system. Because osteocytes regulate healthy bone turnover, it is essential to quantify the relationship between *in situ* mechanical stimulation and the cell biological response.

2 DESCRIPTION OF THE MODEL

2.1 Micro damage threshold identification

Dual experimental and numerical top-down investigations were applied through micro bending tests conducted on human femoral fresh cadaver samples to produce and image the growth of controlled nascent sub-microscopic damage near live osteocytes. The multi-scale local constitutive fracture mechanisms has been identified scale by scale after the balance of the energies at the global scale to evaluate the *in situ* stress field near bone cells shown in Figure 1 (a) and (B). The finite element model is based on explicit bone tissue morphology including the primary, secondary osteons and cement lines. The model Dirichlet boundary conditions are calculated by digital image correlation [2]. The osteon lamellae collagen orientations were modeled by cubic spline and crack bridging by the mineralized collagen fibers in the wake of the crack tips was modeled by a cohesive zone of 50 μm maximum length and 25 μm maximum opening from experimental measurements. The model must satisfy the macroscopic Dirichlet boundary conditions at each loading step, n. The micro damage threshold values and location were then identified by Newton-Raphson iteration algorithm on the local element elastic moduli at each macroscopic loading step and are described by the following constitutive relationship:

$$E_I^n = E_I^0(1 - d_I^n) \tag{1}$$

where E_I^n is the converged elastic moduli of element I at macroscopic loading step n, E_I^0 is the intact elastic moduli of element I prior to macroscopic loading and d_I^n is the converged microdamage parameter in element I at macroscopic load step n. The local micro damage threshold is calculated when the local maximum principal stress exceeds the identified local yield strengths, σ_y^n as follows:

$$d = \left(1 - \frac{\sigma_y^n}{\sigma^n}\right) \tag{2}$$

where σ^n is local máximum principal stress at loading step n in the intact material.

The numerical model shows nascent diffuse damage within osteon lamellae appearing in locations where visible micro cracks further appear at higher load in Figure 1 (C) and (D).

Multi-modal imaging techniques using UV and fluorescent microscopy shown in Figure 1 coupled to a hierarchical multi-level numerical simulations contributed to confirm the location of diffuse damage regions ahead of micro crack tips prior to their growth initiation [2]. Dual SEM and BSEM microscopy were applied to image the relationship between sub-micro cracks in the osteonal lamellae and the canaliculi [4].

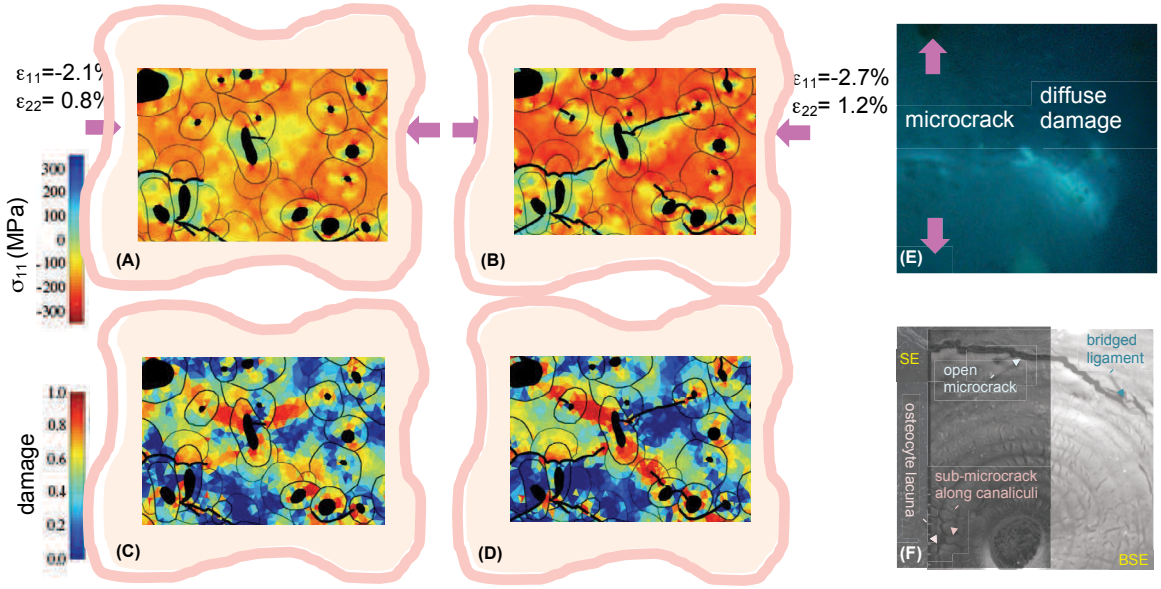


Figure 1: Diffuse damage numerical and experimental identifications in human Haversian cortical bone: (A) and (B) numerical identification of the compressive stress field before and after microcrack growth in a local FEM model including the osteon elastic moduli heterogeneity, the cohesive strength variation with the osteonal lamella orientations and mineralizations and the diffuse damage identification regularizing the local model response with the global sample response, (C) and (D) corresponding diffuse damage local identification, (E) diffuse damage visualization within the osteonal lamellae in the region of a microcrack tip using fluorescence microscopy of human cortical bone stained with calcein blue, (F) identification of sub-microcracks along osteocyte canaliculi in a region neighboring diffuse damage using superimposition of secondary electron and backscattered electron microscopy.

2.2 Three-dimensional numerical reconstruction of human osteocyte morphology

The osteocyte morphology was observed in fresh human cortical bone simple of 500 µm thickness. A osteocyte in a region containing intra-lamellar sub-micro damage is selected for confocal observation. A tissue depth of 16 µm was penetrated by two photon laser excitation in the 650 nm range that provided the best detection of the integrin proteins layer coating the osteocyte lacunae and the canaliculi in which the osteocytic processes run inside Haversian bone as shown in Figure 2-A. Resolution of 0.18 µm /px was achieved in segmentation planes

and 0.51 micron in the orthogonal direction to the planes. Due to the non-zero thickness of the confocal plane of observation, the images were deconvoluted to rebuild a cubic voxel of the 3D confocal observation of the osteocyte morphology.

The 3D morphology of the osteocytes was first reconstruction using medical image visualization software imageJ [5]. The cell body contours were digitally surveyed in the parallel segmentation planes and regularized by smoothing interpolation to remove non-physical noise perturbations. The canaliculi center paths were identified by finding the best path in 3D between points of the same canaliculus by bidirectional search A* on the reciprocal of the cost of moving to another point in the specific Fiji Plugin for cell processes tracking called Simple Neurite Tracer [6].

The complete osteocyte morphology was reconstructed in Matlab based the body contours from the 3D parallel plane segmentation and the regularized center lines of the cell processes using Tubeplot [7]. Previous osteocyte morphology reconstruction captured up to 9 cell processes [8, 9]. In the presented model, 23 processes of the osteocyte were rebuilt in Figure 2-B from the usual 40 bear by such cell.

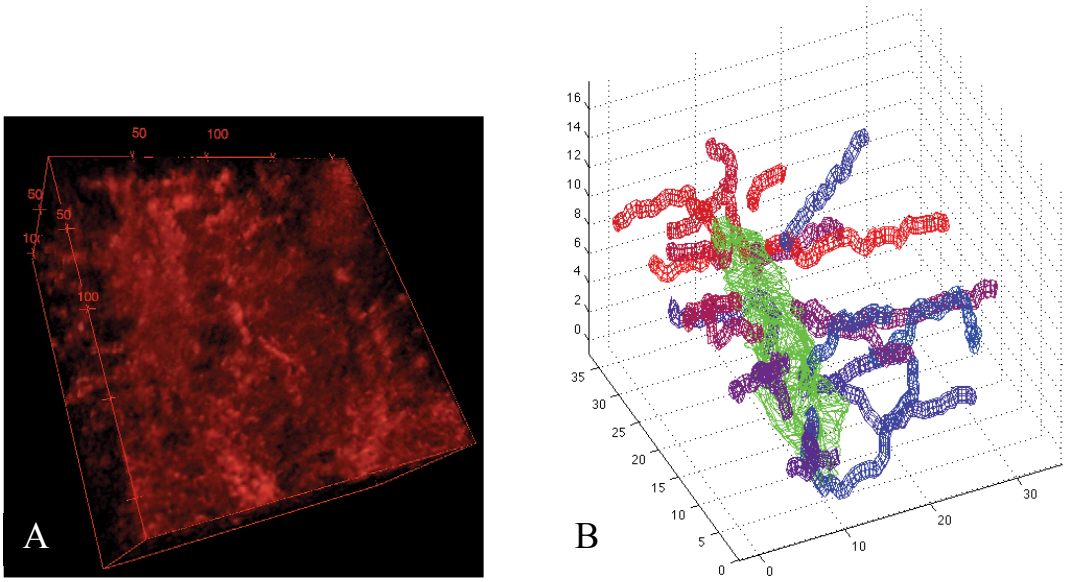


Figure 2: A- 3D confocal microscopy observation of a human osteocyte lacunae and canaliculi in fresh human Haversian cortical bone observed in the transverse direction to the osteons, B- 3D numerical reconstruction of a human osteocyte morphology from confocal microscopy observations.

2.3 Fluorescent flow cytometry in the osteocytes under mechanical loading

Osteocytes are known to respond to mechanical cues by releasing specific chemicals such as calcium [10, 11], PGE1 and NO [12] that can be observed under fluorescent microscopy by markers and are correlated to the cell biological reaction to mechanical stress. To quantify the cell morphology, displacements and *in situ* biological response, progenitor bone cells (MSC) and mature osteocytes (MLOY4) were reseeded in fresh human cortical bone tissues and observed in Nikon-Eclipse TE 200-U fluorescent microscope (B2OA Laboratory, Paris) in Figure 3 and 4 and a Nikon confocal microscope (LBPA Laboratory, Cachan) in Figure 3.

The fluorescence of the osteocytes was enhanced by labeling their cytoplasm with CFSE cell-tracker (Invitrogen). A protocol to clean fresh human bone without altering its mechanical properties was developed using a succession of baths with a solvent and a basic solution, antibiotics, trypsin and a detergent. Live cells were then reseeded in the human bone bathed in biological serum. Bone marrow human mesenchymal stem cells (bm hMSC) harvested on a patient were used for the progenitor osteocytes and mouse mature osteocyte-like cells (MLOY4) were used for mature osteocytes because of the difficulty to isolate primary human mature osteocytes. The cell suspensions (10^5 per ml) were then added to bathe the donor bone tissues. After 5 days the cells were evenly reseeded on the tissue. After 10 days progenitor and mature reseeded osteocytes reorganized as shown in Figure 3.

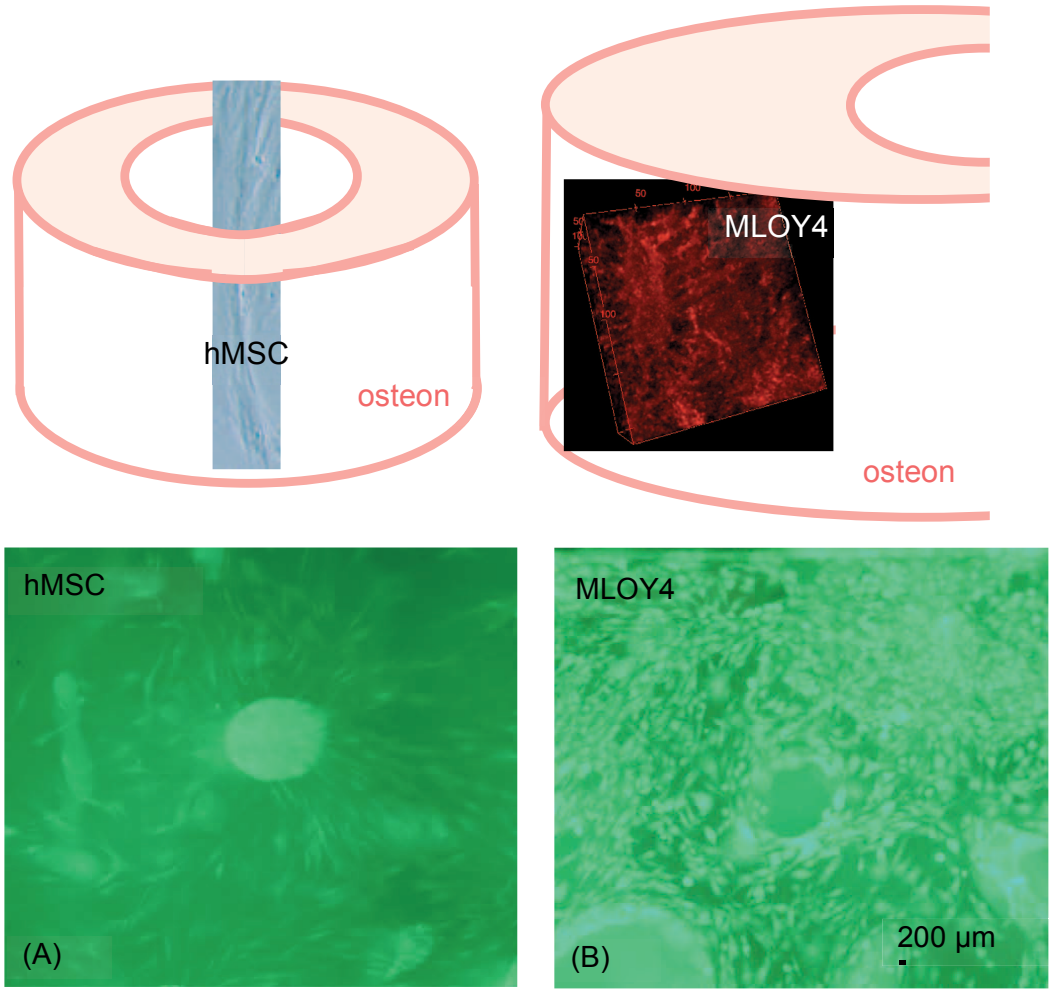


Figure 3: Cell reseeded in human Haversian cortical bone rearrangement after 10 days: (A) progenitor osteocytes (hMSC) moving into the Haversian canals and organized in radial hair-like arrays around the Haversian canals on the osteons transverse cross-section and schematic arrangement of hMSC in the Haversion canal *in vivo* presented above, (B) mature osteocytes (MLOY4) spread throughout the osteon cross-section and schematic arrangement of mature osteocyte observed under confocal microscopy in the osteonal lamellae *in vivo*.

3 RESULTS

3.1 Mechanical response of Live Allograft Bone Systems (LABS)

The numerical model identified nascent diffuse damage within osteon lamellae appearing prior to visible micro cracks as local stresses increase from yield values of 58/-77 MPa close to HAP (hydroxyapatite) strength to 108/-216 MPa when mineralized collagen fibers still bridge cracks [2]. These results confirm known brittle/ductile fracture behavior of bone through load bearing cooperation between HAP mineral phase and organic collagen fibers.

Multi-modal SEM and BSEM microscopic observations within osteon lamellae showed micro-damage patterns near osteocytic processes and canaliculi in Figure 1 (F) in regions neighboring a visible micro cracks and containing diffuse damage. Calcein blue staining of cortical bone reveals diffuse damage regions in Figure 1 (E) where calcium ions exposed on hydroxyapatite broken edges interact with the stain [13].

The load deflection curves of the recellularized bone tissues are shown in Figure 4 for the allograft bone systems containing either bm hMSC or MLOY4 and tested in the elastic region prior to visible micro crack initiation but containing diffuse damage [13] in the anti-plane longitudinal direction. The mechanical responses of the Live Allograft Bone Systems (LABS) were compared to the mechanical responses of fresh cadaver human bone tissues samples tested until complete failure in the three characteristic directions of the osteons.

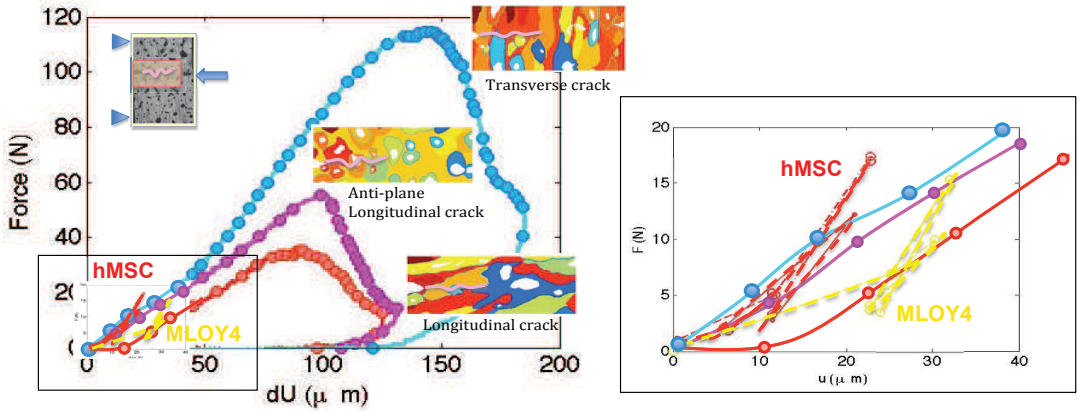


Figure 4: Load-deflection curves of micro 3-point bending tests of fresh cadaver human cortical bone samples in the longitudinal, anti-plane longitudinal and transverse directions until complete fracture compared to micro 3-point bending tests of anti-plane longitudinal human cortical bone samples that are recellularized with either progenitor or mature osteocytes and loaded in the elastic region prior to fracture.

3.2 Progenitor and mature osteocyte *in vitro* special reorganizations on bone tissue

bm hMSCs and MLOY4 cells reorganized *in vitro* as they would *in vivo* in fresh donor bone tissue as shown in Figure 3 after 10 days in an incubation chamber at 37°C. After 10 days incubation, reseeded MCSs relocated inside the Haversian canals in the human tissue samples. The elongated MSCs cells are located *in vivo* in the bone marrow and circulate inside the Haversian canals of cortical bone. MCSs also exhibit morphological features with strong affinities to the topography of their residing lodges. On the other had after 10 days, mature osteocytes seeded and proliferated on the solid osteonal lamellae structures with

characteristic spacing observed in bone osteocyte lacunae organization. The osteocytes also regained hedgehog-like shapes and spread osteocytic processes to recreate a syncytial network as *in vivo*. Progenitor and mature osteocytes reseeded in cortical bone tissues therefore reconstituted live allograft bone systems similar to *in vivo* systems.

3.3 Progenitor and mature osteocyte biological responses to mechanical loading

The osteocyte biological response to mechanical loading was visualized by the instantaneous variation of fluorescence intensity due to cytoplasm staining by CFSE cell tracker. CFSE is a small molecule of 0.473 kDa molecular mass, that is easily permeable to cell plasma membrane in the form of CFDA-SE that looses its diacetate group through an intracellular esterase to become CFSE that covalent binds to intracellular Lysine residues of intracellular proteins and amines. CFSE is a biochemical dye to track cell proliferation over multiple days and also indicates cell viability when used in short amount of time. CFSE and Fluo4-AM labeling have also been used together to track lymphocyte proliferation and measure cytosolic free calcium simultaneously [14]. Figure 4 shows that bm hMSCs lost suddenly their fluorescence under mechanical loading less than 21 N while osteocytes retains a significant amount of fluorescence in Figure 5 for larger mechanical loading-unloading increment cycles up to 147 N shown in Figure 4. However, at larger loading increment despite almost no fluorescence, high magnification reveal that the bm hMSCs contours are still perceivable which could indicate that the cell integrity are preserved and that the cells did not loose fluorescence by plasma membrane rupture but possible other cell signaling mechanisms.

CFSE binds to Lysine and Lysine is an essential α-amino acid that plays a role in different posttranslational modifications such as methylation of the ε-amino group into methyl-, dimethyl- and trimethyllysine that occurs in calmodulin. Calmodulin (CaM for Calcium-Modulated proteIn) is a calcium-binding multifunctional intermediate messenger protein that transduces calcium signals by binding calcium ions and modify its interactions to target proteins afterwards. Proteins that bind to CaM cannot bind to calcium and uses CaM as a calcium sensor and a signal transducer. Calmodulin 1 in particular is located in the cytosol and on the plasma membrane facing the cytosol, CALM1 interacts with cation channel receptors such as TRPV1 activated in the sensation of pain. Calmodulin also mediates processes such as inflammation and apoptosis. L-Lysine plays therefore a very important role in calcium absorption.

Lysine also plays a role in posttranslational modifications such as neddylation, biotinylation, pupylation and carboxylation but most importantly in acetylation, sumoylation, ubiquitination involved in apoptosis and cell response to stress. CFSE binds to lysine through an esterase that operates a deacetylation of CFDA-SE. Acetylation and deacetylation of a protein are significant posttranslational regulatory mechanisms that occur in the regulation of transcription factors, effector proteins, molecular chaperones and cytoskeletal proteins in an analogous manner to phosphorylation and dephosphorylation under the action of kinases and phosphatases. Acetylation and deacetylation has been suggested to crosstalk with key other biological mechanisms such as phosphorylation, methylation, ubiquitination, sumoylation [15] and different cell signaling such as the synthesis of tubulin of the microtubules of the cytoskeleton in which lysine take part and are solicited under mechanical load.

For all the above possible cell mechanisms' interactions and from previous experiments on cell CFSE labeling and calcium flow cytometry [14] and experiments on osteocytes and osteoblasts calcium different responses [16], it is hypothesized that CFSE labeling is indicative here of distinct calcium responses from bone stromal cells at their progenitor and mature stages of their differentiation shown in Figure 5.

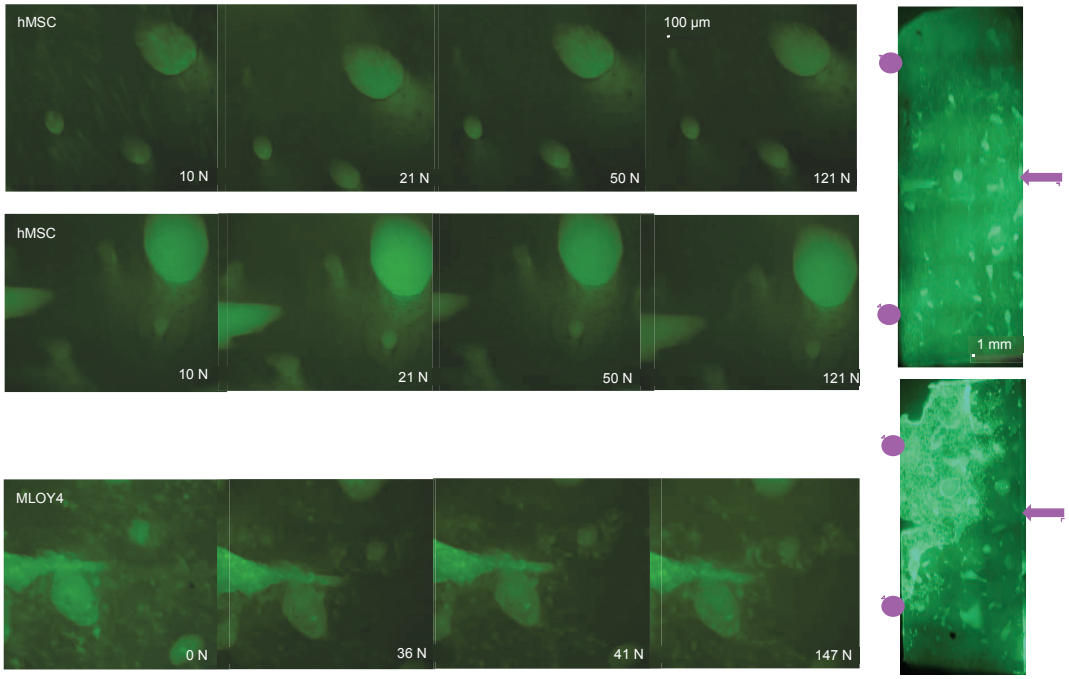


Figure 5: Fluorescence decreasing of the cytoplasm of hMSC and MLOY4 cells stained with CFSE cell tracker when subjected to increasing tensile loading produced by the tension in the human Haversian cortical bone environmental "substrate".

4 CONCLUSIONS

Hybrid experimental and numerical investigations quantified human osteocyte lacunae and canaliculi morphology alterations in bone diffuse damage areas that were correlated to nascent sub-micro cracks. Figure 4 showed the mechanical responses of live allograft bone systems containing live osteocytes that were consistent with the known mechanical response of fresh cadaver human bone tissue. Figure 3 reveals distinct spatial reorganizations of progenitor osteocytes in the Haversian canals after 10 days from the rearrangement of mature osteocytes regularly dispersed throughout the osteonal lamellae after 10 days of incubation. After 10 days the cells relocated in the allograft tissue microstructure studied here *in vitro* as they would be located *in vivo*: progenitor cells in the Haversian canals and mature osteocytes within the mineralized osteonal solid matrix. Human bone marrow mesenchymal stem cells also exhibited a typical elongated tubular shape and migrated into tubular Haversian canals while mature osteocytes displayed oblate-shaped bodies bearing multiple processes acting as hair-like structures shown in Figure 2 to detect mechanical loading in the mineralized extracellular matrix [17]. Loss of fluorescence shown in Figure 5 could be correlated to

possible calcium mediated cell signaling mechanisms involving lysine residues to which CFSE stain binds under the plasma membrane facing cytosolic free calcium participating into efflux through cytoplasmic membrane pumps or intracellular recycling. This aspect is currently under further investigations. Mesenchymal stem cells lost fluorescence under very low loading while a non-negligible percentage of the mature osteocytes retained some fluorescence at moderated load in Figure 5 that is in agreement with a modified calcium response of osteocytes through their differentiating stages observed in the literature [16].

ACKNOLEDGMENT

The authors are grateful to the support of the Farman Institute for the project OLA!. The first author is extremely grateful for the generous support from CMMI BMMB NSF program for Grant 1214816.

REFERENCES

[1] Nather, A., David, V., Teng, J.H.W., Lee, C.W., and Pereira, B.P. Effect of autologous mesenchymal stem cells on biological healing of allografts in critical-sized tibial defects simulated in adult rabbits. *Annals Academy of Medicine* (2010) **39**(8):599-606.

[2] Jonvaux, J., Hoc, T., and Budyn, E. Analysis of micro fracture in human Haversian cortical bone under compression. *International Journal for Numerical Method in Biomedical Engineering* (2012) **28**(9):974-998.

[3] E. Budyn and T. Hoc, Analysis of micro fracture in human Haversian cortical bone under transverse tension using extended physical imaging. *International Journal for Numerical Method in Engineering* (2010) **82**(8):940-965.

[4] Wang, Y., McNamara, L.M., Schafler, M.B., and Weinbaum, B. Strain amplification and integrin based signaling in osteocytes. *J. Musculoskelet. Neuronal Interact* (2008) **8**(4):332-334.

[5] Rasband, W. National Institue of Health. http://imagej.nih.gov/ij/

[6] Longair, M.H., Baker, D.A., and Armstrong, J.D. Simple Neurite Tracer: Open Source software for reconstruction, visualization and analysis of neuronal Processes. *Bioinformatics Advance Access*. (2011).

[7] Weisenberg, J.H. TubePlot. (2004) http://www.mathworks.com/matlabcentral/fileexchange/5562-tubeplot

[8] Stern-Rath A. and Nicolella D.P., Measurement and estimation of osteocyte mechanical strain. *Bone* (2013) **54**:191-195.

[9] Verbruggen, S.W., Vaughan, T.J., McNamara, L.M. Fluid flow in the osteocyte mechanical environment : a fluid–structure interaction approach. *Biomechanical Modelling in Mechanibiology* (2014) **13**:85-97.

[10] Rawlinson, S.C.F., Pitsillides, A.A. and Lanyon L.E. Involvement of different ion channels in oseoblasts' and osteocytes, early responses to mechanical strain, *Bone* (1996) **19**(6):609-614.

[11] Huo, B., Lu, X.L., Costa, K.D., Xu Q. And Guo X.E. An ATP-dependent mechanism mediates intracellular calcium signaling in bone cell network under single cell

nanoidentation. *Cell Calcium* (2010) **47**:234-241.

[12] Kitase, Y., Van Der Plas, A., Semeins, C.M., Ajubi, N.E., Frangos, J.A., Nijweide, P.J. and Burger, E.H.. Mechanical induction of PGe2 in osteocytes blocks glucocorticoid-induced apoptosis through both the beta-catenin and PKA pathways. *Journal of Bone Mineral Research* (2010) **25**(12):2657-2668.

[13] Sun, X., McLamore, E., Kishore, V., Slipchenko, M., Porterfield, D.M. and Akkus, O. Mechanical stretch induced calcium efflux from bone matrix stimulates osteoblasts. *Bone* (2012) **50**:581-591.

[14] Mukherjee, M.S., Giamberardino, C., Thomas, J., Evans, K., Goto, H., Ledford, J.G., Hsia, B., Pastva, A.M. and Wright J.R. Surfactant protein A integrates activation signal strength to differentially modulate T cell proliferation. *The Journal of Immunology* (2012) doi:10.4049/jimmunol.1100461.

[15] Yang, X.J. and Seto E. Lysine acetylation: codified crosstalk with other posttranslational modifications. *Molecular Cell* (2008) **31**:449-461.

[16] Lu, X.L., Huo, B., Chiang V. and Guo X.E. Osteocytic network is more responsive in calcium signaling that osteoblastic network under fluid flow. *Journal of Bone Mineral Research* (2012) **27**(3):563-574.

[17] Budyn, E., Tauc, P., Bensidhoum, M., Petite, H. and Deprez, E. Back to life: fresh osteocytes spreading their processes for optimum mechanotransduction near micro damage in dead bone. *Medical Engineering Centres Annual Meeting and Bioengineering14*, MECbioeng14 Imperial College London, England, UK, ISBN 978-0-9930390-0-3, (2014) p.63.

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis(Eds)

COUPLED RINGS IN RNA NANOTUBES AND PROPERTIES OF BIOLOGICAL NANOCLUSTERS

Shyam Badu, Roderick Melnik and Sanjay Prabhakar

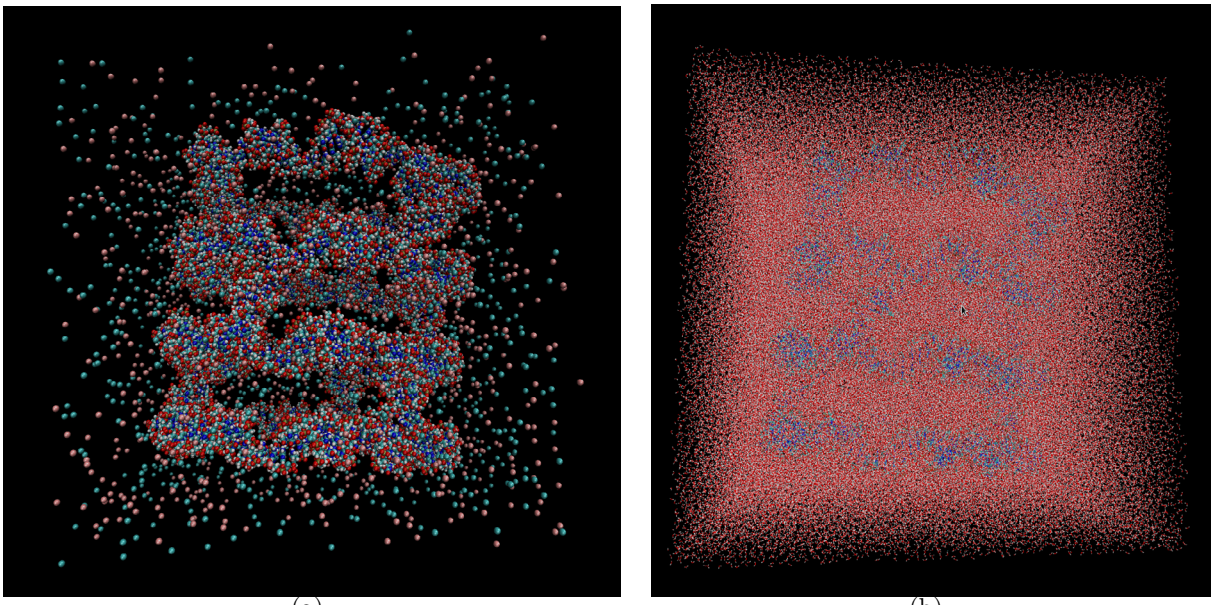
The MS2Discovery Interdisciplinary Research Institute, M2NeT Laboratory,
Wilfrid Laurier University, Waterloo Campus, 75 University Ave West,
Waterloo, ON, Canada N2L 3C5

Key words: RNA nanoclusters, Molecular Dynamics Simulation, Physiological Solutions

Abstract. In this contribution, we study biological nanoclusters where the coupling between their parts is essential. The main focus is on the RNA-based nanostructures. RNA molecules are very flexible in nature. This feature allows us to build various motifs which are indispensable in bio- nanotechnological applications. Based on the previous studies on RNA nanoclusters, in this contribution we analyze the structure and properties of RNA nanotubes, where we focus on nanotubes consisting of a series of coupled nanorings of around 20nm in diameter. We did a molecular dynamics (MD) simulation using CHARMM force field implemented in the NAMD package to study the structural and thermal properties of the nanotube in physiological solutions. Specifically, we have studied the behavior of these nanoclusters in solutions of NaCl, MgCl2 and KCl. We have analyzed such characteristics as the Root Mean Square Deviation (RMSD), the radius of gyration, the number of hydrogen bonds per base pairs, and the ionic distribution around the surface of the RNA nanocluster under study in different solutions. The variation of energy and temperature with simulation time have also been studied for all sets of simulations. The change in these features with the nature of solution has also been studied.

1 Introduction

Unlike most of the other biological molecular systems, RNA is structurally more flexible in nature which makes it an important biomolecular system. One of the most important application of the RNA nanotechnology research is in the field of biomedicine and bioengineering [1, 2]. For these applications it is very important to built the proper self assembly of the RNA building blocks. In our earlier studies, the modeling of RNA nanoclusters has been done by using the RNAI/II building blocks [3, 4, 5, 6]. The RNAI/II building blocks are the RNA strands which are taken from the protein data bank with the pdb code (2bj2.pdb) [7]. The elements of these RNA strands are the nucleobases Cytosine (C) , Guanine (G), Adenine (A), Thymine (T) and Uracil (U).



(a) (b)

Figure 1: VMD generated image of the RNA nanotube immersed into the MgCl2 Solution (a) Water molecules not displayed (b) With water molecules displayed.

The building blocks of these RNA nanoclusters i.e the RNAI and RNAII are defined as the sense and anti sense plasmids that control the replication of COLE1 [8, 9]. COLE1 is a DNA molecule separated from chromosomal DNA that is found in the cell of bacteria. The sequences for the RNAI is (GGCAACGGAUGGUUCGUUGCC) and that for the RNAII is (GCACCGAACCAUCCGGUGC) [7]. The six helical segments are constructed from RNAI and RNAII building blocks to model the RNA nanoring and the nanorings are connected via the links which are also built from the RNA strands of the size 22 nucleotides. The model structure of the RNA nanocluster is presented in Figure 1.

Experimentally the studies have been done for the calculation of free energy of RNA hairpin folding at different concentration of NaCl and KCl solutions [10] and found that the value of the free energy is depending on the logarithmic of the salt concentration as well as the stability of the RNA hairpin in NaCl solution was found more in comparison to the stability in in the KCl solution . The ionic concentration dependence studies on the binding of HIV -B virus has also been studied and it has been found that the binding of two human immunodeficiency viruses type-1 (HIV-1) is much more favourable at high concentration of salt in the presence of magnesium ion [11, 12]. Furthermore, the dependence of the self assembly of the tecto RNA has also attracted interest [11]. Recently, the several salt solutions has been used to study the detection of the MicroRNAs taking the DNA/silver nanocluster as the probe [13]. Having this in mind, in our current study the calculation of properties of RNA nanoclusters in different kind of solutions are performed for various type of ions.

The rest of this paper is organized as follows. In section 2 we describe the computational

methodologies used for calculations. The results are presented and discussed in Section 3. Concluding remarks and an outlook are then stated in Section 4.

2 Computational Details

In the modern development of the computational methods, the classical Molecular Dynamics simulation is one of the most versatile tools for the modeling of the biological systems [14, 15]. Here in our present contribution we have solvated the RNA nanocluster in a water box. To this solvated RNA nanocluster, we have added the K^+, $^{25}Mg^{2+}$and the $^{35}Cl^-$ions in order to make the RNA nanotube immersed in to the varieties of the salt solutions which can facilitate us to see the solution dependent behaviour of the RNA nanocluster The solvation and the ionization of the RNA nanocluster is done by using the VMD tool [16]. In order to do the molecular dynamics simulation of this RNA nanocluster at the constant temperature we have used the CHARMM27 force field [17] implemented by the software package NAMD [18].

In molecular dynamics simulation the classical equations of motion of a molecular system are solved by their time dependent integration. The potential of the system used during the molecular dynamics simulation using CHARMM force filed can be expressed as follows [17]:

$$V_{total} = \sum_{bond} K_b(r-r_0)^2 + \sum_{angle} K_\theta(\theta-\theta_0)^2 + \sum_{dihedral} K_\phi(1+cos(n\phi-\gamma)) + \sum_{Hbond}(\frac{C_{ij}}{r_{ij}^{12}} - \frac{D_{ij}}{r_{ij}^{10}}) + \sum_{Vanderwaals}(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{10}}) + \sum \frac{q_{ij}}{\varepsilon r_{ij}} \quad (1)$$

where the first term corresponds to bonds, second corresponding to angle parameters and so on as indicated in the Equation 1 defining the potential of the system.

The output of the molecular dynamics simulation has also been analysed by using VMD to calculate the radial distribution, ion distribution, RMSD and the radius of gyration at different concentrations of salt solution. The RNA-nanotube modeled from the RNA building blocks has been solvated in a water box. The size of the box is taken in such a way that the distance from the surface of nanocluster to the wall is slightly larger than the cut off radius used in the MD simulation. In order to make the system neutral we have added 1254 $^{23}Na^+$ to the nanotubes. On top of this, before the MD simulation the RNA nanotube with four nanorings is ionized with the 1254 $^{23}Na^+$ and $^{35}Cl^-$ ions added to the waterbox. In the case of the MgCl2 solution we have added 1254 $^{25}Mg^{2+}$and $^{35}Cl^-$ions to the solvated RNA nanocluster system prepared for the molecular dynamics simulation. In case of the MgCl2 solution the $^{25}Mg^{2+}$ions being doubly positive charged, the half of this will contribute to neutralization of the phosphate group's negative charges in the entire tube and the rest of this will contribute to the added $^{35}Cl^-$ions. These systems were first simulated at constant temperature and pressure using NAMD software. The temperature in the system has been controlled by using Langevin's method with damping parameter $\eta = 5$ ps^{-1}. For adding chemical bonds between the segments in the nanoclusters we have used the topotools available in the VMD.

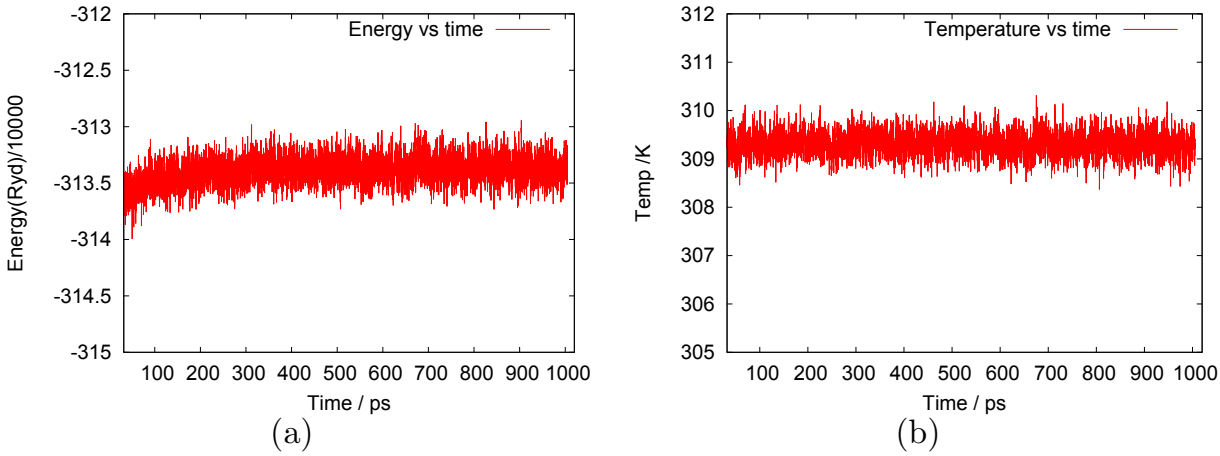


Figure 2: (a) Energy vs time and (b) Temperature vs time during the molecular dynamics simulation of 4 ring RNA nanotube in magnesium chloride solution.

3 Results and Discussion

In this contribution we focus on the behavior of RNA nanoclusters in the varieties of the metal chloride solutions including the physiological solution. The RNA nanoclusters are built from RNAI/II building blocks using molecular dynamics simulation. The construction of the RNA nanoclusters using these RNA building blocks have been provided in detail in our earlier work [4, 5]. The molecular dynamics simulations of the RNA nanotube have been performed in the simulation box for the time period 1ns at a constant temperature 310K. The energy and temperatures are saved in every 1ps. Variation of energy and the temperature during molecular dynamics simulation for the RNA nanotube with four RNA nanorings in the magnesium chloride solution is shown in Figure 2. From the energy plot it is clear that the energy at the beginning of the simulation goes on increasing but after about the simulation time 0.2ns the energy of the system becomes stable. The significant change in the energy is observed in the minimization region of the molecular dynamics simulation. The later part of the simulation is known as the production region around which we will do the calculations of the properties of the RNA nanocluster in the MgCl2 and NaCl solutions.

The number of $^{23}Na^{+}$, $^{25}Mg^{2+}$and $^{35}Cl^{-}$ions around the surface of the RNA nanotube within the distance 5 $\AA$ from its surface as a function of time in NaCl and MgCl2 solutions are presented in the Figure 3 (a). The trajectory of the molecular dynamics simulation is saved for 1ns time range in each of the solutions at 310K temperature. From these ionic distribution plots we see that the number of metal ions (i.e $^{23}Na^{+}$and $^{25}Mg^{2+}$) around the RNA nanotube surface are significantly higher than the number of $^{35}Cl^{-}$ions in both the solutions. The number of $^{25}Mg^{2+}$ions is less then the number of $^{23}Na^{+}$ions because the former is the divalent positive charged ion whereas the later is the monovalent singly charged ion. For a particular type of ions the number of ions at the beginning increases and approximately after 0.4ns the number of ions becomes consistent. The distribution of the $^{25}Mg^{2+}$and $^{35}Cl^{-}$ions around the surface of

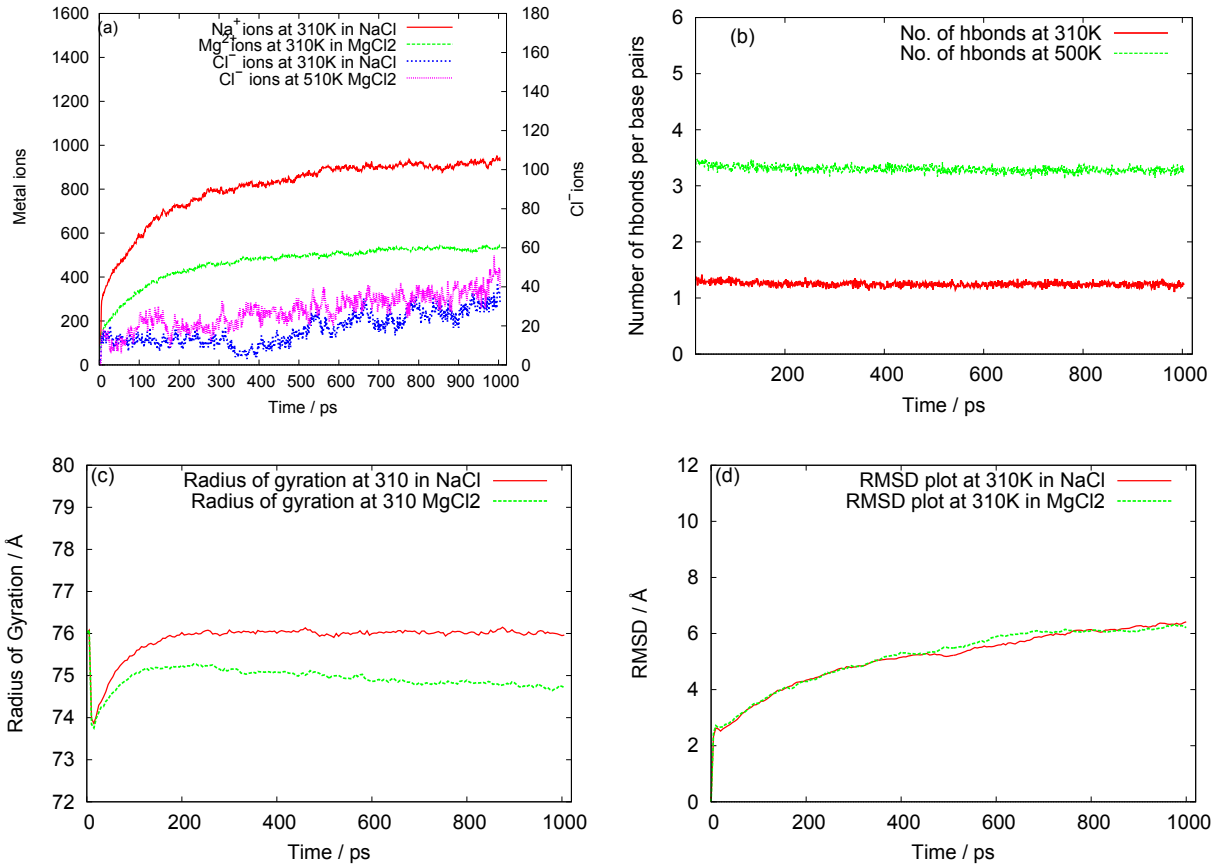


Figure 3: (a) Number of ions with the range of 5 $\AA$ (b) Number of bonds per base pairs (c) Radius of gyration and (d) RMSD of 4 ring RNA nanotube obtained from all atom molecular dynamics simulation.

the RNA nanocluster in MgCl2 solution is found similar to the distribution of the corresponding ions in NaCl solution which is ultimately similar to the characteristics obtained for other various RNA nanoclusters in physiological solution [4, 5]. Furthermore, the results for the radius of gyration, root mean square deviation and the number of basepairs in two different solutions are presented in Figure 3(b),(c) and (d) respectively. From the plots the radius of gyration of the system is increasing at the beginning which later becomes constant for all solutions. The value for the radius of gyration in the MgCl2 solution is higher than that in the NaCl solution. The root mean square deviation is not significantly changing on going from one to another salt solution.

In our earlier studies we have analysed the properties of the RNA nanocluster at different temperatures as well as the different concentration of the physiological solution. In current study the calculation of the properties in different kind of salt solution is done. The temperature and the concentration are taken constant during molecular dynamics simulation.

4 Conclusions

In this paper we have presented four ring RNA nanotube constructed from the RNA building blocks as an example of a coupled problem in studying complex RNA based biological systems. Its behaviour in different metal chloride solutions has been studied. The change in the behaviour of the RNA nanotube in these solutions are found similar to the results obtained in our earlier work those were performed for the RNA nanoclusters of the various size [3, 4, 5]. From our analysis, it is clear that the quality of the results are likely to be improved further by doing the MD simulation for longer period of time those are comparable to the time of the real biological process that occur in the human body. Furthermore, the study of the properties of the RNA nanocluster at the different concentration of $^{25}Mg^{2+}$ions in the solution will give a better understanding about the stability of the RNA nanocluster.

5 Acknowledgement

Authors are grateful to the NSERC and CRC Program for their support and Shared Hierarchical Academic Research Computing Network (SHARCNET: www.sharcnet.ca) for providing the computational facilities. Finally, we would like to thank Dr. P. J. Douglas Roberts for helping with technical SHARCNET computational aspects.

REFERENCES

[1] Guo, P., "RNA nanotechnology: Engineering, assembly and applications in detection, gene delivery and therapy," *J Nanosci Nanotechnol* **5**, 1964–1982 (Dec. 2005).

[2] Osada, E., Suzuki, Y., Hidaka, K., Ohno, H., Sugiyama, H., Endo, M., and Saito, H., "Engineering RNAprotein complexes with different shapes for imaging and therapeutic applications," *ACS Nano* **8**, 8130–8140 (Aug. 2014).

[3] Badu, S. R., Melnik, R., Paliy, M., Prabhakar, S., Sebetci, A., and Shapiro, B. A., "High performance computing studies of rna nanotubes," *Proceedings of IWBBIO-2014 (International Work-Conference on Bioinformatics and Biomedical Engineering)* , 601–607 (2014).

[4] Badu, S. R., Melnik, R., Paliy, M., Prabhakar, S., Sebetci, A., and Shapiro, B. A., "Modeling of RNA nanotubes using molecular dynamics simulation," *Eur Biophys J* **43**, 555–564 (Nov. 2014).

[5] Paliy, M., Melnik, R., and Shapiro, B. A., "Molecular dynamics study of the RNA ring nanostructure: a phenomenon of self-stabilization," *Phys. Biol.* **6**, 046003 (Dec. 2009).

[6] Yingling, Y. G. and Shapiro, B. A., "Computational design of an RNA hexagonal nanoring and an RNA nanotube," *Nano Lett.* **7**, 2328–2334 (Aug. 2007).

[7] Lee, A. J. and Crothers, D. M., "The solution structure of an RNA looploop complex: the ColE1 inverted loop sequence," *Structure* **6**, 993–1007 (Aug. 1998).

[8] Tomizawa, J. I., "Control of cole 1 plasmid replication: The process of binding of RNA i to the primer transcript," *Cell* **38**, 861–870 (Oct. 1984).

[9] Tomizawa, J. I., "Control of ColE1 plasmid replication: Binding of RNA i to RNA II and inhibition of primer formation," *Cell* **47**, 89–97 (Oct. 1986).

[10] Vieregg, J., Cheng, W., Bustamante, C., and Tinoco, I., "Measurement of the effect of monovalent cations on RNA hairpin stability," *J. Am. Chem. Soc.* **129**, 14966–14973 (Dec. 2007).

[11] Jaeger, L., Westhof, E., and Leontis, N. B., "TectoRNA: modular assembly units for the construction of RNA nano-objects," *Nucl. Acids Res.* **29**, 455–463 (Jan. 2001). PMID: 11139616.

[12] Kim, T. and Shapiro, B. A., "The role of salt concentration and magnesium binding in HIV-1 subtype-a and subtype-b kissing loop monomer structures," **31**(5), 495–510 (2013). PMID: 22812415.

[13] Shah, P., Cho, S. K., Thulstrup, P. W., Bhang, Y.-J., Ahn, J. C., Choi, S. W., Rrvig-Lund, A., and Yang, S. W., "Effect of salts, solvents and buffer on miRNA detection using DNA silver nanocluster (DNA/AgNCs) probes," *Nanotechnology* **25**, 045101 (Jan. 2014).

[14] Hansson, T., Oostenbrink, C., and van Gunsteren, W., "Molecular dynamics simulations," *Current Opinion in Structural Biology* **12**, 190–196 (Apr. 2002).

[15] Lindahl, E. R., "Molecular dynamics simulations," *Methods Mol. Biol.* **443**, 3–23 (2008).

[16] Humphrey, W., Dalke, A., and Schulten, K., "VMD: visual molecular dynamics," *Journal of Molecular Graphics* **14**, 33–38 (Feb. 1996).

[17] Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R. D., Kal, L., and Schulten, K., "Scalable molecular dynamics with NAMD," *J. Comput. Chem* **26**(16), 1781–1802 (2005).

[18] MacKerell, Bashford, D., Bellott, Dunbrack, Evanseck, J. D., Field, M. J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F. T. K., Mattos, C., Michnick, S., Ngo, T., Nguyen, D. T., Prodhom, B., Reiher, W. E., Roux, B., Schlenkrich, M., Smith, J. C., Stote, R., Straub, J., Watanabe, M., Wirkiewicz-Kuczera, J., Yin, D., and Karplus, M., "All-atom empirical potential for molecular modeling and dynamics studies of proteins," *J. Phys. Chem. B* **102**, 3586–3616 (Apr. 1998).

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis (Eds)

COUPLING FLUID AND SOLID DOMAINS IN MODELING DRUG TRANSPORT WITHIN TUMOR

MILOS KOJIC*†, MILJAN MILOSEVIC*, VLADIMIR SIMIC*, ARTURAS ZIEMYS†, MAURO FERRARI†

* Bioengineering Research and Development Center BIOIRC
Prvoslava Stojanovica 6, 34000 Kragujevac, Serbia
e-mail: mkojic42@gmail.com

† Houston Methodist Research Institute
6670 Bertner Avenue
Houston, TX 77030, USA

Key words: Tumor model, Pipe Finite Element, Fictitious 1D Finite Element, Coupled Problems

Abstract. Development of a feasible model for transport within complex vasculature network and tissue remains a challenge. Such a model is particularly important when considering drug transport within tumor environment. A drug used to cure the cancer is first transported through blood vessels, then it attaches to the vessel endothelium and faces biological barriers in the vessel wall to reach cancerous cells.
We have developed a model for convective-diffusive drug transport which is simple and computationally efficient. One of the challenges was to couple fluid domain within blood vessels and solid domain of the tumor microenvironment. We have introduced fictitious 1D finite elements which appropriately take into account transport characteristics of the vessel walls. These characteristics include leakage and permeability of the walls. In evaluating wall permeability of a drug, we implemented our hierarchical multiscale methodology which couples molecular dynamics (MD) and continuum FE model. A numerical homogenization procedure was employed to obtain equivalent continuum transport parameters which account for interaction on molecular level between drug and solid components of the wall microstructure. Also, a possibility of using equivalent continuum transport models for capillary beds is investigated in order to further simplify and increase efficiency for the overall model of tumor.
As a numerical example, we calculate transport through a capillary bed to illustrate applicability of our methodology.

1 INTRODUCTION

We here study transport of particles/molecules dissolved within blood, and their subsequent transport through tumor tissue. The tumor vasculature consists of an arterial and venous system, including larger vessels (arterioles and venules) - with diameters on the order of hundred micrometers, and capillary vessels with diameters as small as a few micrometers. We consider that transport occurs from the arterial network into tissue, which is composed of

cells and intercellular space, and back - from tissue to the venous networks. There are two coupled transport domains: a fluid domain consisting of blood and a solid (tissue) domain. Transport within tumors is complex due to irregular blood vessel branching and variability of vessel diameters and lengths. Also, blood flow is affected by presence of cells within blood plasma, so that the fluid has a colloidal character. In our model we approximate the blood by a Newtonian fluid and neglect additional resistance at branchings.

Two references are cited here which are the closest to our study. The most commonly used is the "network" method [1], where the network is represented by blood vessel segments with common edges (nodes) within the network. Pressure change along segments is governed by the Hagen-Poiseuille law, while the pressure is equal for all segments at a common node, and the total flux at interior nodes is equal to zero. A system of linear equations with respect to nodal pressures is formed and solved with the given boundary conditions, pressures and/or fluxes. A generalization of this concept is given in [2].

In our transport model for large vascular systems [3], it was necessary to resolve the coupling of the fluid and solid domains, and to incorporate blood vessel wall properties with respect to hydraulic and diffusive transport. We here give details about the methodology for this coupling.

In the next section we outline the main features of our tumor computational model and in Section 3 introduce the concept of the fluid-solid coupling. One numerical example is given in Section 4, followed by concluding remarks in the last section.

2 COMPUTATIONAL MODEL FOR MASS TRANSPORT WITHIN LARGE BLOOD VESSEL NETWORK AND TISSUE

Here is summarized the formulation of our model for transport within large blood vessels systems and tissue, according to reference [3].

The 1D continuity equation has the form

$$\frac{\partial v_x}{\partial x}+\frac{1}{r}\frac{\partial\left(rv_r\right)}{\partial r}=0 \tag{1}$$

where v_x and v_r are axial and radial velocity components, and x and r are axial and radial coordinates, respectively. Using boundary conditions at the (pipe) vessel internal wall R: v_x=0, $v_r=\partial R/\partial t$, this equation transforms into

$$\frac{\partial v}{\partial x}+2\frac{v}{R}\frac{\partial R}{\partial x}+\frac{2}{R}\frac{\partial R}{\partial t}=0 \tag{2}$$

where v is the mean velocity. We further assume that wall is incompressible and elastic with Young's modulus E, so that the radius change due to the change in pressure p (assuming cross-sectional uniformity), can be related by the equation:

$$\frac{\partial R}{\partial t}=k_E R^2\frac{\partial p}{\partial t} \tag{3}$$

where the elastic constant k_E is

$$k_E = \frac{3}{4\delta_0 E} \tag{4}$$

Here, we have neglected the change of the wall thickness δ_0 . The continuity equation (2) can be written in terms of the fluid flux Q ($Q=R^2\pi v$),

$$2R^3\pi k_E \frac{\partial p}{\partial t} + \frac{\partial Q}{\partial x} = 0 \tag{5}$$

The equation of the balance of linear momentum, the 1D Navier-Stokes equation, for a pipe cross-section with the radius *R*, is:

$$\frac{\partial v_x}{\partial t} + v_r \frac{\partial v_x}{\partial r} + v_x \frac{\partial v_x}{\partial x} + \frac{1}{\rho}\frac{\partial p}{\partial x} = \nu\left(\frac{\partial^2 v_x}{\partial r^2} + \frac{1}{r}\frac{\partial v_x}{\partial r} + \frac{\partial^2 v_x}{\partial x^2}\right) \tag{6}$$

where ν is kinematic viscosity, and ρ is fluid density. This equation can be written in terms of pipe flux Q as

$$\rho\frac{\partial Q}{\partial t} + 2\alpha\rho\frac{Q}{A}\frac{\partial Q}{\partial x} + \frac{1}{R}\left[-2\alpha\rho\frac{Q}{A}\frac{\partial R}{\partial x} + \frac{2\mu(\gamma+2)}{R}\right]\mathrm{Q} + A\frac{\partial p}{\partial x} = 0 \tag{7}$$

where A is the cross-sectional area, and μ is the fluid viscosity; α is a dimensionless parameter

$$\alpha = \frac{2}{R^2 v^2}\int_0^R r v_x^2 dr \tag{8}$$

which is a measure of variation of velocity with distance from the axis of symmetry (velocity profile); and γ is [4]:

$$v_x = \frac{\gamma+2}{\gamma}\left[1-\left(\frac{r}{R}\right)^{\gamma}\right]v \tag{9}$$

parameter which defines the profile shape, from parabolic to plug flow. The value γ=2 corresponds to a parabolic profile, while γ=9 can be used for oscillatory flow of one cardiac cycle. Parameters α and γ can be related as: $\alpha = (\gamma+2)/(\gamma+1)$. For a parabolic profile, which is commonly assumed, α=4/3.

Capillary wall is a multilayered composite structure with fenestrations (holes) allowing plasma leakage to the surrounding tissue [5]. Capillary walls can roughly be classified into non-fenestrated (continuous), fenestrated, and discontinuous (sinusoidal) [5]. The leakage can mathematically be described using the Starling hypothesis [6] (known also as Kedem-Katchalsky equation [7]):

$$Q = L_P A_w\left[(p_{ves} - p_{tis}) - \sigma(\pi_{ves} - \pi_{tis})\right] \tag{10}$$

where L_p is hydraulic conductivity [8], A_w is the surface area of the vessel-tissue interface, p_{ves} and p_{tis} are pressures within vessel and within tissue (interstitium pressure); π_{ves} and π_{tis} are osmotic pressures in vessel and tissue, and σ is the osmotic reflection coefficient. The leakage of fluid through vessel walls must be taken into account in the computational model.

Convective-diffusive transport within blood vessels can be simplified if blood is considered as a homogenous fluid, so that the governing equation of mass balance has the form [9]

$$-\frac{\partial c}{\partial t}-\frac{\partial c}{\partial x}v+\frac{\partial}{\partial x}\left(D\frac{\partial c}{\partial x}\right)+q=0 \tag{11}$$

where c is concentration, D is diffusion coefficient within fluid - which may depend on concentration, and q is a source term.

Particulate transport through the vessel wall is very complex due to various physical and biological effects, and can be expressed in the form [5]

$$Q_s = D_w A_w \left(c_{ves}-c_{tis}\right)+Q\left(1-\sigma\right)\Delta c_{lm} \tag{12}$$

where Q_s is the particulate flux, D_w is diffusion (or transport) coefficient of the wall (flux per unit area of the wall and unit concentration), c_{ves} and c_{tis} are concentrations within vessel and tissue; and Δc_{lm} is the mean logarithmic concentration,

$$\Delta c_{lm}=\frac{c_{ves}-c_{tis}}{\ln\left(c_{ves}/c_{tis}\right)} \tag{13}$$

which can also be taken as the arithmetic mean. Diffusion coefficient D_w can be determined experimentally or numerically.

Interaction between particles/molecules and the wall solid components (epithelium cells, fibers, etc.) can dominate the transport through the wall. This interaction on a molecular level may be incorporated into a continuum transport model by evaluation of the effective diffusion coefficients (or scaling functions) using MD procedures and a numerical homogenization (within a multiscale-hierarchical concept); such multiscale model has been developed and applied to various bioengineering problems [10-14].

We consider tissue as a porous medium where transport of fluid and particles/molecules (drugs, nutrients) occurs within intracellular space. The solid phase form a complex microstructure composed of cells, or cells and various types of fibers. The fundamental relations used for fluid flow and for diffusion are summarized in [15]. We cite here the basic Darcy's law (which can have additional terms [15]) and the continuity equation in case of neglecting deformations of the solid microstructure,

$$v_i=-k_{Di}\frac{\partial p}{\partial x_i},\ \text{no sum on } i; \qquad v_{i,i}+\mathrm{q}^v=0,\ \text{sum on } i\text{: } i=1,2,3 \tag{14}$$

where v_i are Darcy's velocities, k_{Di} are Darcy's coefficients, p is the interstitial pressure, and q^v is a source term; the second equation expresses the continuity equation. Diffusive transport equation has the form (11), now extended to three dimensions,

$$-\frac{\partial c}{\partial t}-v_i\frac{\partial c}{\partial x_i}+\frac{\partial}{\partial x_i}\left(D_i\frac{\partial c}{\partial x_i}\right)+q^V=0, \qquad \text{sum on } i\text{: } i=1,2,3 \tag{15}$$

and it also may be modified to account for specific biological effects [15]. Diffusion coefficients can also be calculated using the multiscale-hierarchical concept and numerical homogenization as for capillary wall: a reference volume is selected and equivalent diffusion parameters are evaluated from the mass release curves [14].

We next summarize the basic finite element equations for blood vessels and tissue, based on the above fundamental equations. First, the 2-node pipe 1D element for fluid flow is derived form (7) by a standard Galerkin procedure [9],

$$\left(\frac{1}{\Delta t}\mathbf{M}^{vv}+\mathbf{K}^{vv}\right)\mathbf{Q}+\mathbf{K}^{vp}\mathbf{P}=\mathbf{Q}^{ext}+\frac{1}{\Delta t}\mathbf{M}_{IJ}^{vv}\mathbf{Q}^t \tag{16}$$

where $\mathbf{Q}^{ext}$ is external nodal flux, and $\mathbf{Q}^t$ is the flux at start of time step of size Δt. The matrices are given in [3], with the matrix $\mathbf{K}^{\mathrm{vv}}$ being a function of the flux Q within the element. Further, we differentiate equation (7) with respect to axial coordinate x, and then write that equation in the weak form, as (for equilibrium iteration "i")

$$\left(\mathbf{M}^{p(i-1)}+\mathbf{K}^{p(i-1)}\right)\Delta\mathbf{P}^{(i)}=\mathbf{F}^{(i-1)}-\left(\mathbf{M}^{p(i-1)}+\mathbf{K}^{p(i-1)}\right)\mathbf{P}^{(i-1)}+\mathbf{M}^{p(i-1)}\mathbf{P}^t \tag{17}$$

where $\mathbf{P}^t$ is nodal pressure at start of time step. Details are given in [3].

The balance equation (17) represents the basic equation with pressures as the nodal variables in which the continuity equation (5) is incorporated. During solution process, equation (16) is used for updating fluxes in matrices of equation (17). The derived equations for 1D FE are applicable to deformable and rigid pipes. In case of a rigid pipe, equation (17) reduces to linear equation:

$$\mathbf{K}^p\mathbf{P}=\mathbf{F} \tag{18}$$

leading to a system of equations used in the "network" method [1], [2]. Details about the above derivations are given in [16].

In case of diffusion, there is additional system of equations, following from the balance equation (15),

$$\left(\frac{1}{\Delta t}\mathbf{M}^c+\mathbf{K}^c+\mathbf{K}^{cv}\right)^{(i-1)}\Delta\mathbf{C}^{(i)}=\mathbf{Q}_c^{ext}+\mathbf{Q}_c^V-\frac{1}{\Delta t}\mathbf{M}^{c(i-1)}\left(\mathbf{C}^{(i-1)}-\mathbf{C}^t\right)-\left(\mathbf{K}^c+\mathbf{K}^{cv}\right)^{(i-1)}\mathbf{C}^{(i-1)} \tag{19}$$

where the matrices and the source vector $\mathbf{Q}_c^V$ are evaluated at end of time step; $\mathbf{C}^{(i-1)}$ and $\mathbf{C}^t$ are nodal concentrations at the iteration (i-1) and start of time step, respectively; and $\mathbf{Q}_c^{ext}$ is

the external nodal flux vector. The matrix which couples convection and diffusion within the pipe is the matrix $\mathbf{K}^{cv}$. The matrices are given in [16].

The FE balance equation for fluid transport within tissue follows form continuity equation (14),

$$\mathbf{K}^t \Delta \mathbf{P}^{(i)} = \mathbf{Q}_t^{ext} + \mathbf{Q}_t^V - \mathbf{K}^t \mathbf{P}^{(i-1)} \tag{20}$$

while the convective-diffusive balance equation has the form (19); details are given in [3].

3 FINITE ELEMENT FOR COUPLING FLUID AND SOLID DOMAIN

We have above summarized the governing transport equations in differential and FE form. They differ for the fluid domain (within blood vessels) and tissue. The boundary between these two domains is represented by blood vessel walls, which have their own transport characteristics. Here, following [3] we introduce a 2-node 1D fictitious element to connect the two domains, with including the wall transport properties. The fictitious element AB (Fig. 1) have the node A connected to fluid, hence the node A is a node of 1D pipe element; and the node B is the node of the continuum (tissue) medium. Geometrically, the nodes A and B are at the same spatial position. There are gradients of pressure and concentration between nodes A and B. The characteristics of the AB element are such that they represent the particulate and fluid transport properties of the parts (L/2 in Figure 1) of pipe elements associated to the common node A.

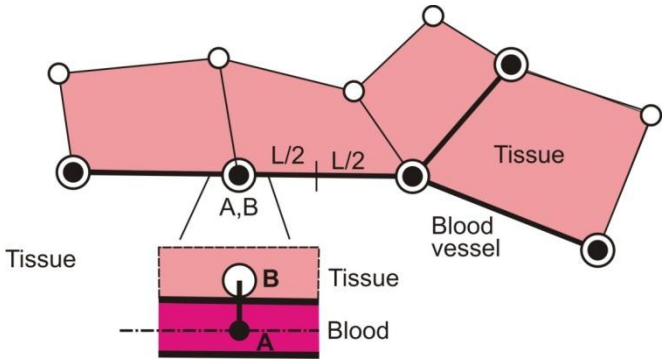


Figure 1: Fictitious 1D element for connection of blood vessel and tissue finite elements [3]. The node A belongs to a 1D vessel FE, while (at the same spatial position) the node B of the fictitious element belongs to a tissue continuum FE.

In order to formulate the fictitious FE, we express the fluid flux according to equation (10) as

$$Q_f = \bar{L}_p \bar{A}_p \left(P_A - P_B \right) \tag{21}$$

where $\bar{L}_p$ and $\bar{A}_p$ are the equivalent hydraulic conductivity and surface, respectively. We have neglected the osmotic part of the transport in this equation. The product $\bar{L}_p \bar{A}_p$ can be evaluated as

$$\bar{L}_p \bar{A}_p = \frac{1}{2} \sum_e L_p^e A^e \tag{22}$$

where summation is over elements *e* with the common nod A; the coefficient ½ indicates that half of the total element surface belongs to the common node A. Then, the incremental form of balance equation can be written as,

$$\bar{\mathbf{K}}^p \Delta \mathbf{P}^{(i)} = \mathbf{Q}^{ext(i-1)} - \bar{\mathbf{K}}^p \mathbf{P}^{(i-1)} \tag{23}$$

where

$$\bar{\mathbf{K}}^p = \bar{L}_p \bar{A}_p \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \tag{24}$$

$\mathbf{Q}^{ext(i-1)}$ are fluxes from the surrounding elements (for node A those are 1D blood vessel elements, and for node B are tissue elements); and the nodal pressure vector has the terms P_A and P_B.

We formulate the convective-diffusive FE starting from the expression for flux,

$$Q_c = \bar{D}_w \bar{A}_c \left(C_A - C_B \right) + \bar{Q} \left(1 - \sigma \right) \Delta C_{lm} \tag{25}$$

where $\bar{D}_w$ and $\bar{A}_c$ are the equivalent diffusive transport coefficient and equivalent diffusion surface, respectively; $\bar{Q}$ is the equivalent fluid flux, which can be determined from pipe flow equations; and ΔC_{lm} is the is the mean logarithmic concentration (according to (13)). The equation (25) can be transformed into the incremental form and then written into the FE format (23),

$$\bar{\mathbf{K}}^{c(i-1)} \Delta \mathbf{C}^{(i)} = \mathbf{Q}_c^{ext(i-1)} - \bar{\mathbf{K}}^{c(i-1)} \mathbf{C}^{(i-1)} \tag{26}$$

where $\bar{\mathbf{K}}^{c(i-1)}$ is the transport matrix, and $\mathbf{Q}_c^{ext(i-1)}$ are diffusion fluxes coming from connecting pipe and continuum elements. The matrix $\bar{\mathbf{K}}^{c(i-1)}$ is

$$\bar{\mathbf{K}}^{c(i-1)} = \left(\bar{D}_w \bar{A}_c + \frac{\bar{Q}^{(i-1)}}{\ln \left(C_A / C_B \right)^{(i-1)}} \left(1 - \sigma \right) \right) \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \tag{27}$$

The flux $\bar{Q}^{(i-1)}$ can be determined from equation (21).

Finally, we present a concept of the tumor model where the role of the fictitious 1D finite element is emphasized. Figure 2 shows a tumor model where larger vessels are represented by 1D pipe elements, while capillary bed is modeled by the equivalent continuum elements [3]. The pipe elements are connected to tissue by the fictitious elements. Also, the fictitious elements are employed in the numerical homogenization to obtain equivalent transport parameters of the continuum. With this concept, the model of tumor becomes very efficient and can be used in research and further in medical practice.

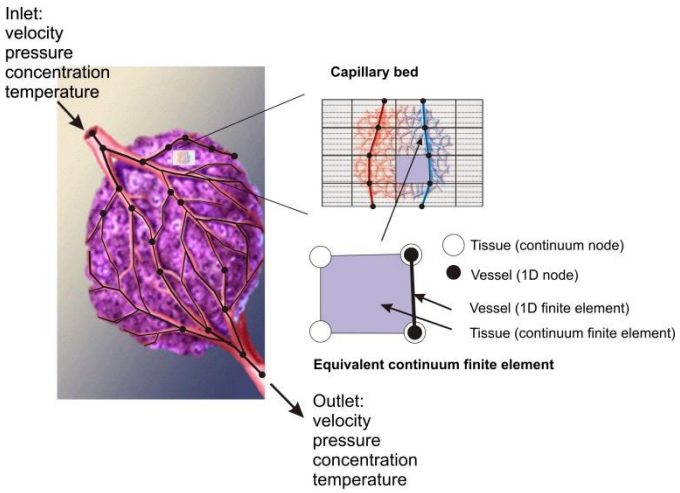


Figure 2: Concept of a tumor model with use of 1D fictitious elements for coupling fluid and solid domains

4 NUMERICAL EXAMPLE

Here we present application of our model on a capillary bed domain. The model consists of capillary network, located between the inlet artery (left) and outlet vein (right side of model). Blood vessels are modeled by 1D finite elements and surrounding tissue is modeled by 2D FEs. Boundary conditions include: inlet and outlet pressures and inlet and outlet concentration of a drug, with P_{IN} = 25mmHg, P_{OUT} = 10 mm Hg, C_{IN} = 25 M/l and C_{OUT} = 10 M/l.

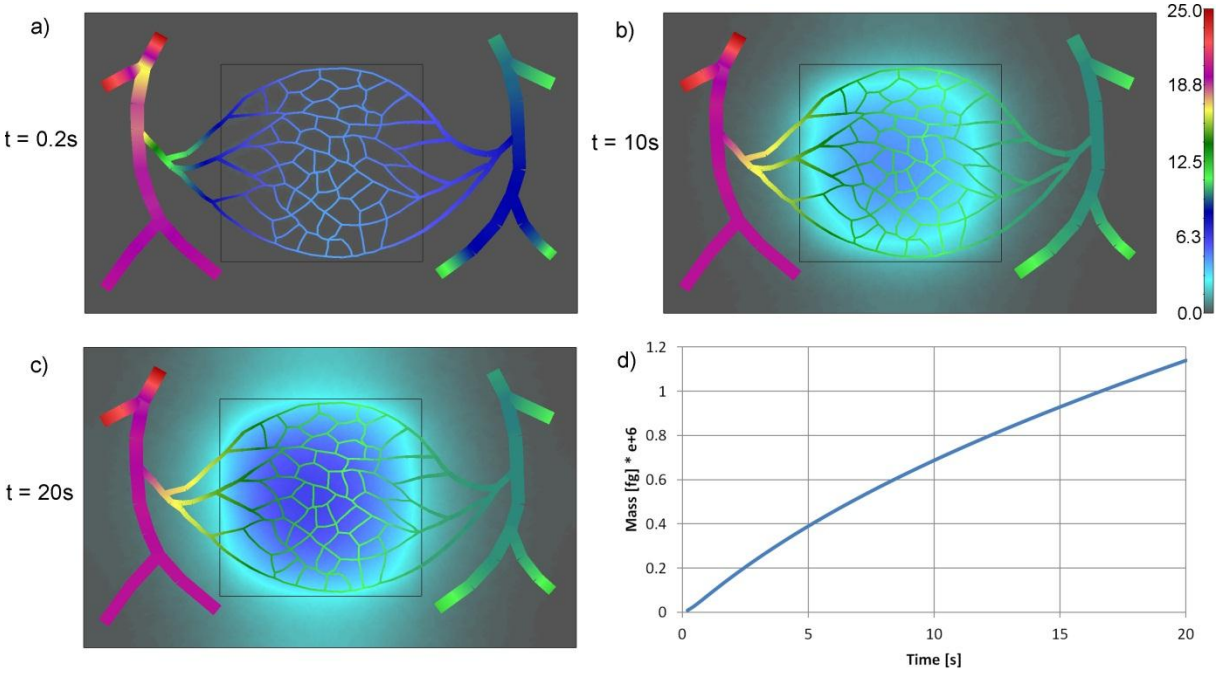


Figure 3: Concentration field within capillary bed domain for different times: a) t = 0.2s b) t = 10s and c) t = 20s; d) Mass change over time in the 2D surrounding tissue

It is taken that the tissue boundary, within the indicated square, is impermeable. Fluid viscosity is 7.5 x E-6 mmHgs, while diameters of vessels are in the range of 30 mm for artery and vain down to 4 mm for capillary network. Leakage coefficient of the wall is 1 x E-3 mm /s, permeability of capillary wall is 0.05 Mole / mm^2 s, diffusion coefficient in capillary is 5 x E+5 mm^2/s, diffusion coefficient in tissue is 1000 mm^2/s, and the Darcy coefficient in tissue is 1.0 x E-4 [mm^2 / mmHg s]. Concentration distributions of the drug within the capillary bed and surrounding tissue at times: t=0.2s, 10s and 20s is shown on Figures 3a,b,c respectively. Mass change during time in 2D surrounding tissue is shown on Fig 3d.

4 CONCLUSIONS

A 1D finite element is formulated for coupling fluid and solid domains in modeling transport within tumor. The element captures complex transport characteristics of capillary walls and provides an efficient way to simulate transport of drugs from blood vessels to tissue.

ACKNOWLEDGEMENTS

This study has been supported from funds of State of Texas Emerging Technology Fund, Alliance of NanoHealth (ANH), and the Ernest Cockrell Jr. Presidential Distinguished Chair (M.F.)

The authors acknowledge support from Ministry of Education and Science of Serbia, grants OI 174028 and III 41007, and City of Kragujevac.

REFERENCES

[1] Lipowsky, H. H and Zweifach B.W. Network analysis of microcirculation of cat mesentery. *Microvascular Research* (1974) 7: 73-83.

[2] Pries, A.R. ..... and Secomb T.W., Structural adaptation and heterogeneity of normal and tumor microvascular networks, *PLoS Computational Biology* (2009) 5 (5): e1000394.

[3] Kojic, M, Milosevic M., Kojic, N., Starosolski Z., Ghaghada K., Serda R., Annapragada,A., Ferrari M. and Ziemys, A. A multi-scale FE model for convective-diffusive drug transport within tumor and large vascular networks, *Comp. Meth. Appl. Mech. Engrg.* (under review).

[4] Smith, N.P., Pullan A.J. and Hunter J. An anatomically based model of transient coronary blood flow in the heart. *SIAM J. Appl. Math*., (2002) **62**: 990–1018.

[5] Jain R. Transport of molecules across tumor vasculature. *Cancer and Metastasis Reviews* (1987) **6**: 559-593.

[6] Kedem O. and Katchalsky A. Thermodynamic analysis of the permeability of biological membranes to non-electrolytes. *Biochim Biophys Acta*. (1958) **27**: 229-245.

[7] Kedem O. and Katchalsky A. A physical interpretation of the phenomenological coefficients of membrane permeability. *J. Gen. Physiology*. (1961) **145**: 143-179.

[8] Sevick E.M. and Jain R.K. Measurement of capillary filtration coefficient in a solid tumor. *Cancer Research* (1991) **51**: 1352-1355.

[9] Kojic M., Filipovic N., Stojanovic B. and Kojic, N. *Computer Modeling in Bioengineering - Theoretical Background, Examples and Software*, John Wiley and Sons,

Chichester, England, (2008).

[10] Ziemys A., Kojic M., Milosevic M., Kojic N., Hussain F., Ferrari M. and Grattoni A. Hierarchical modeling of diffusive transport through nanochannels by coupling molecular dynamics with finite element method. *J. Comp. Physics* (2011) **230**: 5722–5731.

[11] Kojic M., Milosevic M., Kojic N., Ferrari,M. and Ziemys A. On diffusion in nanospace, *J. Serbian Soc. Comp. Mechanics*, (2011) **5** : 84-109.

[12] Kojic M., Ziemys A., Milosevic M., Isailovic V., Kojic N., Rosic M., Filipovic N. and Ferrari M. Transport in biological systems. *J. Serbian Soc. Comp. Mechanics* (2011) **5**: 101-128.

[13] Ziemys, A., Kojic,M., Milosevic,M. and Ferrari, M. Interfacial effects on nanoconfined diffusive mass transport regimes. *Phys. Review Letters* (2012) 108: 236102-1-5.

[14] Kojic,M., Milosevic,M., Kojic,N., Kim, K., Ferrari, M. and Ziemys A. A multiscale MD–FE model of diffusion in composite media with internal surface interaction based on numerical homogenization procedure. *Comput. Methods Appl. Mech. Engrg*. (2014) **269**: 123–138.

[15] Khaled, A.-R.A. and Vafai, K. The role of porous media in modeling flow and heat transfer in biological tissues. *Int. J. Heat and Mass Transfer* (2003) **46**: 4989–5003.

[16] Kojic,M., Milosevic,M., Simic, V. and Ferrari M. A 1D pipe finite element with rigid and deformable walls. *J. Serbian Soc. Comp. Mechanics* (2014) **8**(2): 38-53.

.

MASS TRANSPORT PHENOMENA AT THE SOLID-LIQUID NANOSCALE INTERFACE IN BIOMEDICAL APPLICATION

ANNALISA CARDELLINI[*†], MATTEO FASANO[*], ELIODORO CHIAVAZZO[*] AND PIETRO ASINARI[*]

[*]Multi-scale-Modeling-Laboratory, Dipartimento di Energia (DENERG),
Politecnico di Torino,
Corso Duca degli Abruzzi 24, 10129 Torino, Italy
[*†]e-mail: annalisa.cardellini@polito.it, web site: http://www.polito.it/small

Key words: Nanomedicine, Nanoconfined Water, Water Transport Phenomena, Protein and Amino Acids, Molecular Dynamics.

Abstract. Understanding heat and mass transfer phenomena at solid-liquid nanoscale interface plays a crucial role for introducing novel and more rationally designed theranostic particles, drug with tailored features and for gaining new insight on the biomolecules functioning. For instance, the water transport properties in the proximity of Amyloid beta peptides can influence the formation of amyloid plaques found in the brains of Alzheimer patients.

In the present work, transport behavior of water molecules in nanoconfined conditions has been investigated. By means of equilibrium Molecular Dynamics (MD) simulations, characteristic length of water confinement has been evaluated in the proximity of several biomolecules such as proteins and amino acids. Moving from proteins to their building blocks (i.e. amino acids), a similarity in water behavior was initially expected; MD simulations results show, instead, a more complex picture revealing a difference between the potential of water nanoconfinement by either proteins or amino acids. Hence, the reduction of water mobility in the proximity of nanoscale interfaces does not rely only on the local physical and chemical properties of the biomolecules surface, but the effects of size and potentials overlap should be also taken into account.

1 INTRODUCTION

Nanomedicine is the science dealing with the design and development of therapeutic and/or diagnostic agents in the nanoscale range (i.e. diameters from 1 to 100 nm), including all the critical aspects linked to the transport and delivery of therapeutic molecules and drugs in the body. Several studies underline the importance of nanoscience in medicine. For instance, Nuclear Magnetic Resonance Imaging (NMRI), which is one of the most spread and powerful diagnostic techniques, is receiving great advantages thanks to the use of nanoparticles. Relaxivity parameters ($r_1$and r_2) of MRI contrast agents are enhanced by nanoconfined conditions: for examples, Ananta et al. experimentally found r_1 increases by confining contrast agents in mesoporous silica [1], whereas, Gizzatov A. et al. proved similar results for r_2 of ultra small paramagnetic iron oxides nanoparticles (USPIOs) [2]. Other examples of successful coupling of nanotechnology to medicine come from therapeutic

application. Nanotechnology-based approaches to hyperthermia seem to have great potentiality and a reduced invasiveness [3][4].

In all the previous examples, the understanding of heat and mass transport phenomena at the nanoscale takes on great importance for designing patient-tailored and more performing biomedical devices. In particular, the modulation and precise control of the water molecules diffusion in the vicinity of MRI contrast agents allow remarkable improvement in imaging performance, whereas an accurate study of heat conduction in the proximity of heated nanoparticles can be extremily useful to design their shape, size and coating for thermal ablation. Moreover, also proteins may modify their structure and thus their functionality according to the dynamics of the surrounding water environment [5][6]. Enzymatic activity, molecular recognition and folding process of proteins are strongly influenced by surrounding water mobility [7][8]. Furthermore the hydration layer has a key role in the formation of protein aggregates, such as those involved in Alzheimer's and Parkinson's diseases [9][10].

The properties of water nanoconfinement and the reduced water mobility at the solid-liquid interfaces have been studied and scaled by Chiavazzo et al. [13][14]. After having evaluated the self-diffusion coefficient of water in several nanoconfined geometries, authors formulated a scaling behaviour for water molecules diffusion in confined configurations: $D(\theta) \cong D_B(1-\theta)$ where D is the water self-diffusivity in the solvated system, D_B is the bulk self-diffusion coefficient of water and θ is a dimensionaless parameter which is related to confinement conditions. In particular, θ is given by the ratio between the water confined volume and the total volume of water in the system.

In this article, the nanoconfinement of water has been studied in the proximity of proteins and their building blocks: amino acids. Numerous experimental and theoretical studies have already demonstrated that the water molecules in the proximity of a protein surface are subjected to confined dynamics [15][16]. Moreover several works have shown the important role of amino acids in the protein structure stabilization [26][27], or their influence on water viscosity [26]. However, a broad physical understanding of the water mobility modification in the vicinity of any protein is still a subject of investigation, mainly because of the variety of amino acids physicochemical properties (hydrophilic vs. hydrophobic).

Here, molecular dynamics simulations are used for evaluating the different water confining capabilities of a large variety of proteins and amino acids. The main results show the hydration layer around poroteins, forming beacause of the attractive potential with interface, is almost 40% bigger than the amino acids. The reason of this descrepacy has been attributed to the few atoms of proteins building blocks involving in the non-bonded interaction with water molecules. This latter evidence sheds light on the difficulty in diriving water confinement around proteins starting from their building blocks: just 20 amino acids can join with several possible configurations providing different proteins characteristics and functionality. Finally, some points of discussions for a future work are presented.

2 METHODOLOGY

2.1 Molecular Dynamics simulations

In order to study the water transport properties at the solid-liquid nanoscale interface, Molecular Dynamics (MD) simulations have been performed by GROMACS (GROningen

MAchine for Chimical Simulations) software [17]. By solving classical Newton's laws, MD simulations are able to describe the motions of single atoms within a complex molecule or particle.

In this study, equilibrium MD simulations have been carried out for five proteins (B1 Immunoglobulin binding-domain, 1PGB; Ubiquitin, 1UBQ; Green Fluorescence Protein, 1QXT; Lysozime, 1AKI; Leptine, 1AX8) and 13 amino acids. The proteins sample covers a large variety of these biomolecules, with different structure and functionality, which allows to deduce water confinement behavior independently from the specific protein characteristics. Among the 20 amino acids in nature, just some examples per each group (polar, charged and neutral) have been studied (Figure 1). Hence, hydrophobic as well as hydrophilic behavior is investigated. In detail we considered arginine (ARG), aspartic acid (ASP), glutamic acid (GLU) and lysine (LYS) as charged amino acids; asparagine (ASN), glutamine (GLN), serine (SER), threonine (THR) and tyrosine (TYR) among polar amino acids and valine (VAL), isoleucine (ISO), leucine (LEU) and glycine (GLY) for the neutral family.

The geometries of all the studied cases have been downloaded from the RCSP Protein Data Bank [18] and the topologies have been completed by including CHARMM27 [19][20] force field. Each of these biomolecules has been solvated in a dodecahedral box of water where solvent molecules are described by spc/e model [20]. Before minimizing the energy, ions are added where needed, in order to achieve the neutrality of the system. Then two equilibration steps of 100 ps each have been performed: the former has been computed within a canonical ensemble (fixed Number of particles, Volume and Temperature-NVT) at 300 K by imposing a V-Rescale thermostat [21] on both water and biomolecule; the latter (fixed Number of particles, Pressure and Temperature -NPT), equilibrates the system to 300 K and 1 bar pressure, by means of Parinello-Rahman barostat [23][24]. Finally a 1 ns simulation has been performed for the equilibrated system.

Confining characteristics of the considered biological surfaces are studied by post-processing MD trajectories at steady state conditions.

The solvent mobility has been described by the self-diffusion coefficient related to the slope of the mean square displacement (MSD) of water by the Einstein relation:

$$\lim_{t\to\infty}\langle\|\boldsymbol{r}_i(t)-\boldsymbol{r}_i(0)\|^2\rangle_{i\in A}=6Dt$$

where $\boldsymbol{r}_i(t)$ and $\boldsymbol{r}_i(0)$ are the position vectors of the solvent molecules at the time t, and at the time t = 0, respectively and D the self diffusion coefficient of water. The considered simulation time is larger compared to the typical time of the velocity autocorrelation function, so that any dynamical coherence of the molecules motion has disappeared. The GROMACs tool to compute the solvent self-diffusion is *g_msd*. Instead by *g_sas*, the solvent accessible surface (SAS) [25] of each bio-structure can be calculated. Both the total S_{tot} and local solvent accessible surface related to the single atom ($S_{loc,i}$) are readily computed by GROMACS, once the geometry of the system is known (e.g. in the form of a *pdb* file). Moreover, the *g_sas* option, "*-q"* provides a *pdb* file containing the Cartesian coordinates of the points forming the solvent accessible surface.

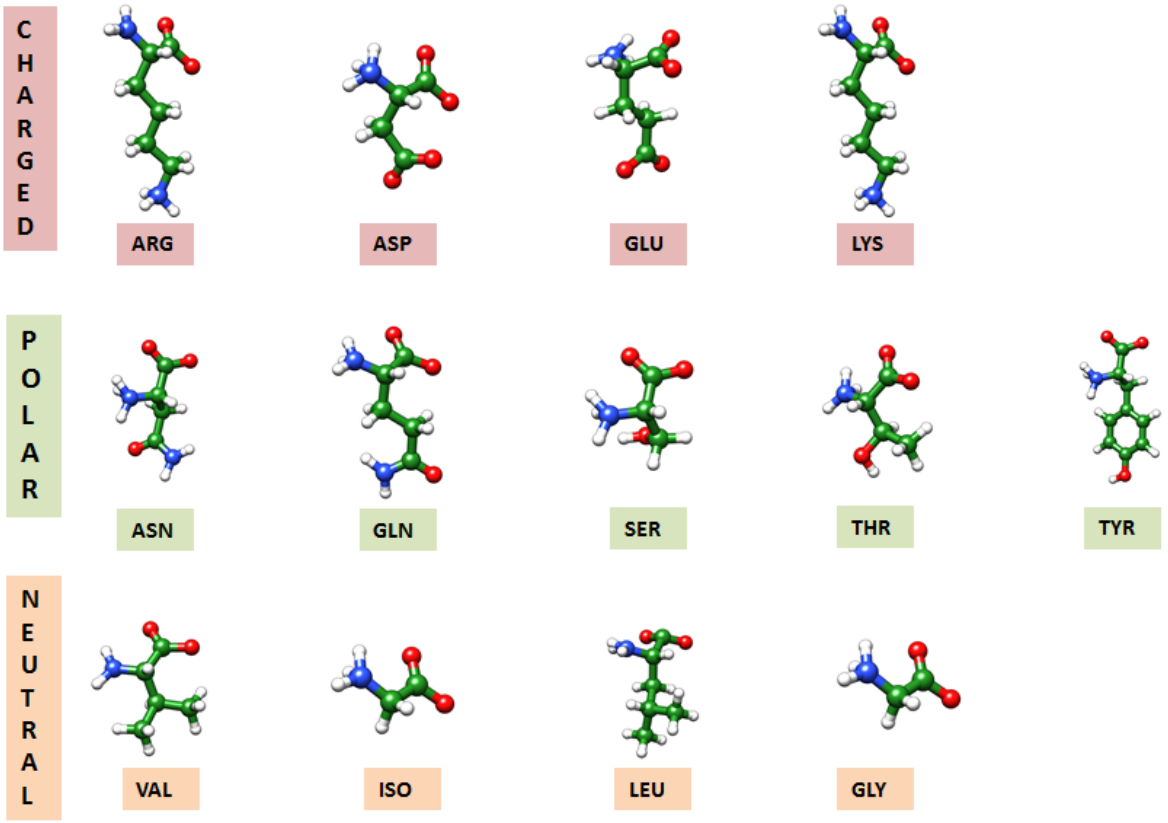


Figure 1: Structure of the 11 amino acids studied in the present work. Amino acids pictures are made with USCF Chimera [11].

2.2 Derivations of the characteristic length of water confinement

Moving from the bulk region to the proximity of solid-liquid surfaces, water molecules show a reduction of mobility and a more ordered structure. While in the bulk, fluid water molecules fluctuate with a kinetic energy proportional to $k_B T$, where k_B is the Boltzmann constant ($1.38 \times 10{-}23$ JK^{-1}) and T is the temperature, close to the solid surfaces they are subjected to Van der Waals (U_{vdw}) and Coulomb (U_c) interactions, which interfere with their state of agitation. This induces a layering of water molecules with reduced mobility at the soli-liquid interfaces, as already pointed out in other works [12].

Chiavazzo and colleagues [13] introduced a characteristic length, δ, to quantify the thickness of such confined water layer. The length δ is derived from the effective potential U_{eff} that water molecules feel at the Solvent Accessible Surfaces (SAS). The calculation takes into account just the interactions between water and the N atoms of the solid structure within a fixed cut-off radius (R_c), where R_c has been chosen large enough to include all atoms with a potential contribution different to zero, namely U_{eff} ~0. The corresponding effective potential energy U_{eff} is due to Van der Waals (U_{vdw}) and Coulomb (U_c) interactions and it has been computed as:

$$U_{eff}(n) = U_{vdw}(n) + \langle U_c \rangle(n), \tag{1}$$

where the *n*-axis is orthogonal to the SAS and it is passing through the center of the atom *i-th*. Van der Waals interactions are modelled by:

$$U_{vdw}(n) = \sum_{k=1}^{N_n} 4\varepsilon_k \left[\left(\frac{\sigma_k}{r_k}\right)^{12} - \left(\frac{\sigma_k}{r_k}\right)^{6} \right], \tag{2}$$

with ε_k, σ_k and r_k denoting the depth of the potential well, the distance where such potential becomes zero (both obtained with Lorentz-Berthelot rule), and the Euclidean distance between the generic line point with coordinate *n* and the center of *k−th* nearest neighbour, respectively. For the Coulomb interactions, the average potential energy between the N atoms and the water dipoles, at a fixed temperature T, is

$$\langle U_C \rangle(n) = -E\mu_w \Gamma\left(\frac{E\mu_w}{k_B T}\right), \tag{3}$$

where E, μ_w, and Γ denote the electrical field strength, water dipole moment (7.50×10^{-30} C m for SPC/E model) and the Langevin function. The $\langle U_C \rangle(n)$ formulation is derived from the Maxwell-Boltzmann distribution of water dipoles orientation, due to the thermal agitation. Knowing the effective potential $U_{eff}(n)$ for the atom *i*, a corresponding characteristic length δ_i can be estimated. According to [13], δ_i is evaluated considering the depth within which the effective potential is stronger than water molecules thermal energy. Therefore, based on the definition of δ_i, all the water molecules located within such a distance are significantly affected by the Van der Waals and Coulomb interactions, whereas all the water molecules beyond the characteristic length δ_i can escape the potential well generated by the solid wall. In general, the quantity δ_i varies per each atom *i*. The mean characteristic length $\bar{\delta}$ of the overall solid surface can be derived as:

$$\bar{\delta} = \frac{\sum_{i=1}^{N} \delta_i S_{loc,i}}{S_{tot}}, \tag{4}$$

with $S_{loc,i}$ and N being the specific (per-atom) SAS for the atom *i* and the total number of atoms, respectively. Note that the above formulation is general and applies to hydrophilic and hydrophobic surfaces, regardless of their electrostatic surface charge.

3 RESULTS

3.1 Water confinement at solvent-proteins interfaces

The computed characteristic lengths of water confinement for the five proteins previously described are shown in Table 1. Although the study cases present strong differences in structure, shape and functionality, the thickness of water molecules layer forming in the proximity of proteins surfaces is comparable and it oscillates in the range: 0.305±0.01 nm. Hence, results show that independently from the specific protein, just few layers of water molecules are significantly affected by the effective potential of the solid surface.

Regarding the solvent accessible surface, values cover a range from 36.04 nm^2 (B1

Immunoglobulin binding-domain) to 108.8 nm^2 (Ubiquitin).

The self-diffusion coefficient of water in the proximity of the studied proteins is reduced respect to the bulk value ($2.6 \cdot 10^{-9}$ m^2/s [26]) and it goes from $2.364 \pm 0.025 \cdot 10^{-9}$ m^2/s (Lysozime protein) to $2.434 \pm 0.014 \cdot 10^{-9}$ m^2/s (B1 Immunoglobulin binding-domain protein). The latter results are in good agreement with the prediction of the scaling law suggested by [13].

Table 1: Mean characteristic length $\bar{\delta}$ of water nanoconfinement, solvent accessible surface (SAS) and water self-diffusion coefficient for five solvated proteins

Proteins	$\bar{\delta}$ [nm]	SAS [nm^2]	D [m^2/s]
1 PGB	0.295	36.04	$(2.434 \pm 0.014) \cdot 10^{-9}$
1UBQ	0.309	48.02	$(2.411 \pm 0.063) \cdot 10^{-9}$
1QXT	0.302	108.8	$(2.425 \pm 0.016) \cdot 10^{-9}$
1AKI	0.306	67.4	$(2.364 \pm 0.025) \cdot 10^{-9}$
1AX8	0.315	68.9	$(2.372 \pm 0.016) \cdot 10^{-9}$

3.2 Water confinement at solvent-amino acids interfaces

As described in the previous paragraph, the relation between water nanoconfinement length and proteins surface properties appears not so trivial: although proteins present different fuatures and configurations, they show the same ability in confining water at their surface. Because of this complexity a deeper investigation has been carried out on proteins building blocks: amino acids in dilute aqueous solution.

The combined use of molecular dynamics simulations and computational methods has brought to the results in Table 2. Although just slight differences are evident among the characteristic length of polar, charged and neutral amino acids, it is interesting to note that those with similar solvent accessible surface present values of $\bar{\delta}$ increasing with the hydrophilicity. For example, the water molecules layer in the proximity of aspartic acid ($\bar{\delta}$ = 0.18 nm), is larger than the Valine characteristic length ($\bar{\delta}$ = 0.153 nm). This result can be considered in line with the definition of δ: because of the presence of charged and polar residues in the side chain of aspartic acid, water molecules are affected by a deeper potential well and thus by a stronger confinement at the solid-liquid interface.

Concerning the self-diffusion coefficients of water in the proximity of the latter amino acids, *D* values are comparable to those of proteins.

Table 2: Mean characteristic length $\bar{\delta}$ of water nanoconfinement and solvent accessible surface (SAS) for the 13 amino acids considered

Charged amino acids	$\bar{\delta}$ [nm]	SAS [nm^2]
ARG	0.208	3.06
ASP	0.180	2.26
GLU	0.185	2.52
LYS	0.184	2.77

Polar amino acids	$\bar{\delta}$ [nm]	SAS [nm^2]
ASN	0.168	2.30
GLN	0.182	2.55
SER	0.163	1.98
THR	0.164	2.19
TYR	0.213	3
Neutral amino acids	$\bar{\delta}$ [nm]	SAS [nm^2]
VAL	0.153	2.29
ISO	0.167	2.51
LEU	0.180	2.57
GLY	0.155	1.65

3.3 Interpreting the mismatch between δ proteins and amino acids

A comparison between Table 1 and Table 2 clearly leads to the consideration that the mean characteristic length of confinement evaluated for the amino acids is almost 40% less than the proteins one. This mismatch becomes even more evident from the Figure 2 where, the average value of $\bar{\delta}$ is plotted for the three amino acids groups (charged, polar and neutral) and for the studied proteins.

It is well established that the hydration layer forming at the solvent-particle interface is strongly dependent on the surface physicochemical properties. However, since amino acids are proteins building blocks, the significant differences in terms of the mean characteristic length could result unclear. The latter behavior may be due to a size-dependent effect. In fact, because of the reduced number of atoms forming the amino acids, a water molecule in their proximity is subjected to a weak effective potential.

Two further analysis have been followed to prove the size dependent feature of δ. First, a MD configuration made out of two identical amino acids (arginine in this case) has been set-up (Figure 3a). The initial distance, h, between arginines centers of mass has been fixed at 3 nm and 1 ns equilibrium molecular dynamics simulation has been performed while fixing the amino acids relative position. In this case, each amino acid was characterized by its own $\bar{\delta}$, which exactly corresponds to the specific value reported in Table 2. Then, h was progressively decreased to 2 nm, 1nm up to 0.5 nm values (Figure 3b-d). As shown in Figure 3, the solvent accessible volume in the region between the amino acids (h) is reduced as the amino acids were forced to approach each other. At one extreme, when h = 0.5 nm, water molecules are no more able to wet the entire surface of the two arginines, but they can just access to the external part of the coupled molecules. Comparing the $\bar{\delta}$, obtained for each set-up, it is interesting to note an increase of the characteristic length of water confinement as the inter amino acids distance reduces (Table 3). This behavior is strictly connected to the effective potential overlap: as the arginines approach, more atoms are located within the cut-off radius, R_c, and thus more intense non-bonded interactions can be established between water molecules and the atoms of the bio-structure. Similar analysis has been repeated increasing the number of amino acids within the water box. Configurations with four and nine

amino acids are then simulated. As evident from Table 3, the mean characteristic length of water confinement further increases, as the amino acids distance reduces. A maximum value of 0.279 nm is reached in case of 9 arginine amino acids arranged in a 3x3 matrix with inter distances of 0.5 nm.

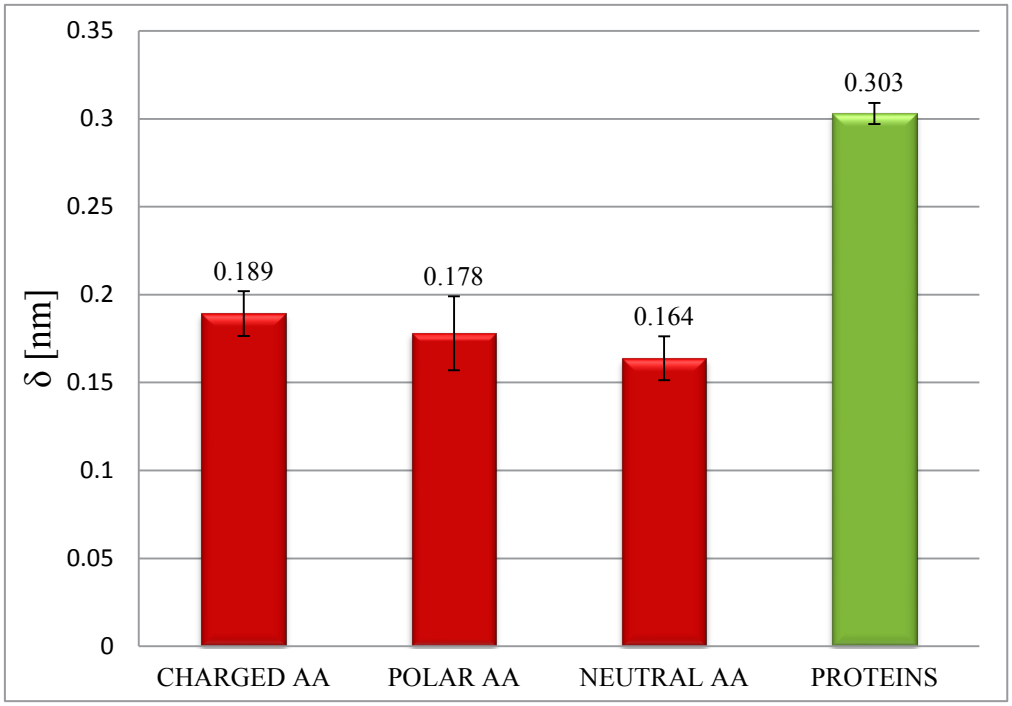


Figure 2: Comparison between amino acids' and protein' s characteristic length of water confinement. An avarage value of $\bar{\delta}$ is calculated among proteins and charged, polar, and neutral amino acids.

The latter value can be almost compared with the proteins hydration layer previously obtained, even if, probably the strong dipole of the peptide bonds (covalent chemical bonds formed when the amino group of one amino acid joins to the carboxyl group of its neighbor) can increase the effect of water confinement around a protein. Hence, in addition to the physicochemical properties of the solid surface, the water nanconfinement length is also influenced by the size of the considered molecule, and the potential overlapping assumes a relevant role to reduce the mobility of the polar solvent molecules close to the interfaces.

To confirm the argument of potential overlapping, a second analysis has been considered. The MD trajectories of the solvated Ubiquitin protein (1UBQ) have been performed. Then, just a small portion of 1UBQ solvent accessible surface was analyzed and the effective potential U_{eff} evaluated (Figure 4). Results in Figure 4 clearly show that the solvent felt the potential from several amino acids present in its vicinity. Hence, the water local confinement can be attributed to the complex overlapping of the non-bonded potentials arising from contiguous amino acids.

Table 3: Mean characteristic length $\bar{\delta}$ of water nanoconfinement and SAS for different configuration of coupled arginine amino acids.

n° Arginine molecules	h [nm]	$\bar{\delta}$ [nm]	SAS [nm^2]
1		0.208	3.06
2	3	0.198	6.21
2	2	0.209	6.23
2	1	0.216	6.09
2	0.5	0.255	5.11
4	0.5	0.256	9.10
9	0.5	0.279	16.15

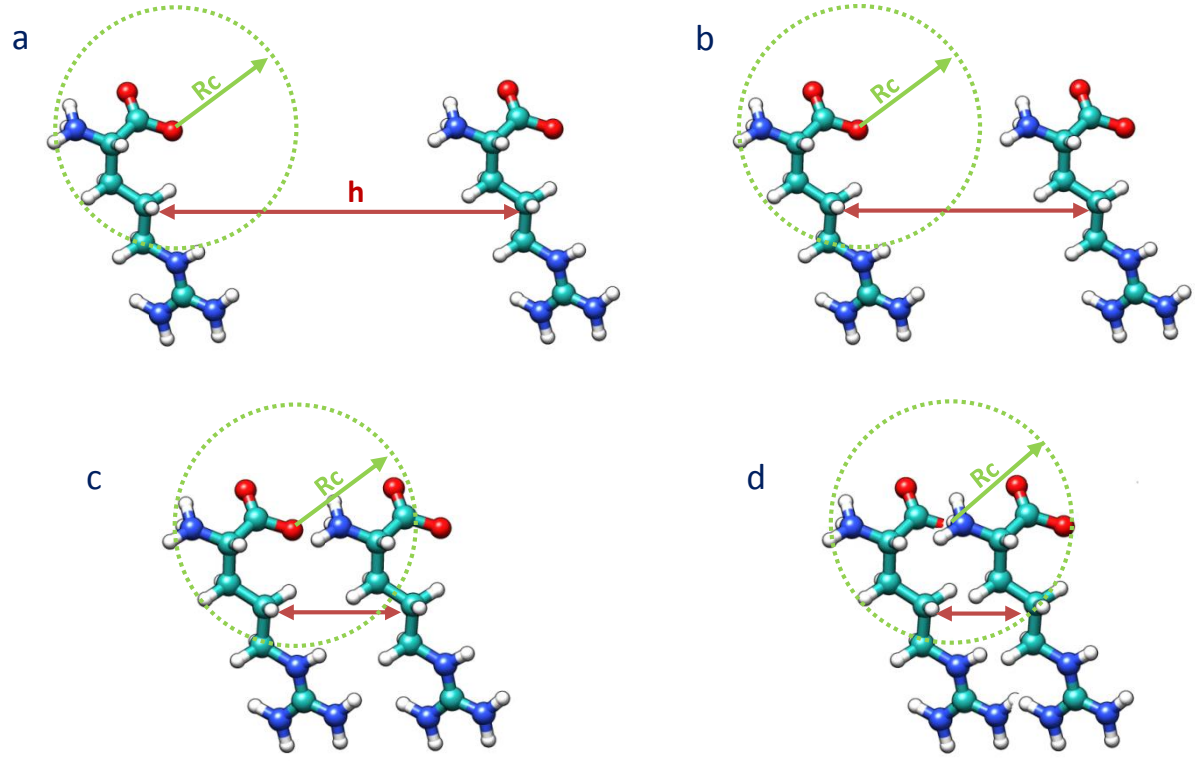


Figure 3: Two ARG amino acids at different distane h. a) h = 3nm b) h = 2nm. c) h = 1 nm. d) h = 0.5 nm. Atoms within the cut-off radius contribuite to the effective potential which influences the water mobility at the solid-liquid interfaces.

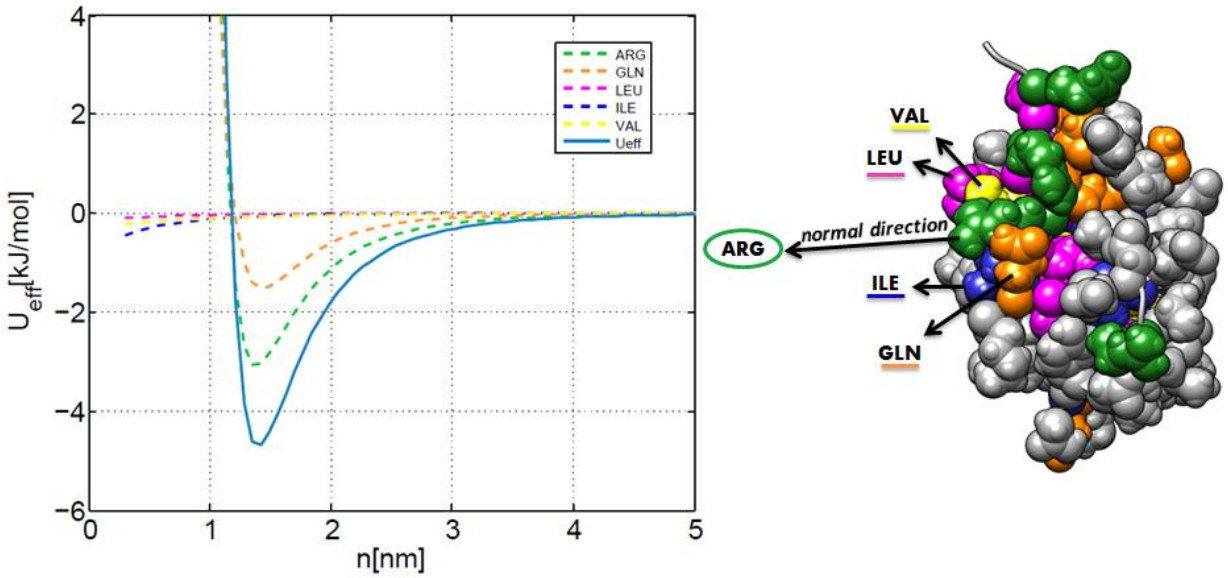


Figure 4: Local effective potential (U_{eff}) felt by water molecules located along one of the normal directions to Ubiquitin surface. Because of the precise position on the protein, the dominant contribution to the effective potential comes from arginine -ARG- and glutamic acid -GLN- atoms

4 CONCLUSIONS

The study of water dynamics under nanoconfined conditions has strong implication in nanomedicine, from the diagnostic to the therapeutic field.

In this work, the water transport behavior in the proximity of proteins and their building blocks, amino acids, has been investigated. By means of molecular dynamics simulations and matlab computational tools, the water confinement length $\bar{\delta}$ has been evaluated at some proteins and amino acids interface. The latter physical quantity is defined as the distance between a water molecule and the solid surface where the attractive non-bonded interactions of the solid prevail on the kinetic energy of the liquid, thus causing a reduced mobility for water. The analysis of several study cases has shown a consistent mismatch: the almost constant values of δ among different proteins ($\cong 0.3\,nm$) is significantly larger than the average one for amino acids ($\cong\ 0.19\,nm$). We have demonstrated that the hydration layer at the solid-liquid interface has a strong dependence on the biomolecules size, in addition to the physicochemical properties of the solid surface. Thus, the limited amount of atoms in the amino acids structure determines a weak attractive potential with the water molecules in their proximity and thus a reduced confinement length.

The latter evidence has been further proved by analyzing the $\bar{\delta}$ value in the light of the possible superposition principle of the solid-liquid non bonded potential. On one hand, $\bar{\delta}$ has been evaluated at the interface of two coupled amino acids at a variable distance, namely $h = 0.5 \div 3\,nm$. Moreover, also configurations with four and nine amino acids have been considered. On the other hand, the effective potential between water molecules and amino acids atoms has been calculated on a local portion of Ubiquitin surface. Results show that the enhanced water confinement capabilities of proteins do not depend on the average value of the amino acids one, but a cross interaction of the effective potential could occur. Therefore,

overall water confining properties of proteins are not strictly coupled to the ones of their building blocks, but they are enhanced by complex amino acids-amino acids interactions.

The study of the hydration layer around amino acids or very small peptides such as Alzheimer's amyloid-β(1–40) peptide could have implication in understaing the biomolecules aggragation, dynamics and functionality [29][30].

REFERENCES

[1] Ananta, J., Godin, B., Sethi, R., Moriggi, L., Liu, X., Serda, R.E., Krishnamurthy, R., Muthupillai, R., Bolskar, R.D., Helm, L., Ferrari, M., Wilson, L.J and Decuzzi, P. Geometrical confinement of gadolinium-based contrast agents in nanoporous particles enhances T1 contrast. *Nat. Nanotech.* (2010) **5**: 815.

[2] Gizzatov, A., Key, J., Aryal, S., Ananta, J., Cervadoro, A., Palange, A.L., Fasano, M., Stigliano, C., Zhong, M., Di Mascolo, D., Guven, A., Chiavazzo, E., Asinari, P., Liu, X., Ferrari, M., Wilson, L.J. and Decuzzi, P. Hierarchically structured magnetic nanoconstructs with enhanced relaxivity and cooperative tumor accumulation. *Adv. Funct. Mat.* (2014) **24**: 4584.

[3] Coss, R. and Linnemans W. The effects of hyperthermia of the cytoskeleton: a review. *International Journal of Hyperthermia* (1996) **12(2)**:173-196.

[4] Hildebraandt, B., Wust, P., Ahlers, O., Dieing, A., Sreenivasa, G., Kerner, T., Felix, R. and Riess, H. The cellular and molecular basis of hyperthermia. *Critical review in oncology/hematology* (2002) **43(1)**: 33-56.

[5] Zhou, H.-X. Macromolecular crowding and confinement: biochemical, biophysical, and potential physiological Consequences. *Annual review of biophysics* (2008) **37**:375.

[6] Ellis, R.J. and Minton, A.P. Cell biology: join the crowd. *Nature* (2003) **425(6953)**:27-28.

[7] Bizzarri, A.R. and Cannistraro, S. Molecular dynamics of water at the protein-solvent interface. *The Journal of Physical Chemistry B* (2002) **106(26)**:6617-6633.

[8] Best, R.B. and Hummer, G. Diffusive Model of proteins folding dynamics with Kramers turnover in rate. *Phys.Rev.Lett.* (2006) **96**:228104.

[9] Fernandez, A. and Scheraga, H.A. *Proc. Natl. Acad. Sci*. (2003) **100**:2391– 2396.

[10] Buchete, N.V. and Hummer, G. Structure and dynamics of parallel β-sheets, hydrophobic core and loops in Alzheimer's Aβ fibrils. *Biophys.J.* (2007) **92(9)**:3032-3039.

[11] Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C. and Ferrin, T.E. Ucsf chimera-a visualization system for exploratory research and analysis. *Journal of Computational Chemistry* (2004) **25(13)**:1605–1612.

[12] Castrillón, S.R.V., Giovambattista, N., Aksay, I. and Debenedetti, P.G. Effect of surface polarity on the structure and dynamics of water in nanoscale confinement. *J.Phys.Chem. B* (2009) **113**:1438-1446.

[13] Chiavazzo, E., Fasano, M., Asinari, P. and Decuzzi, P. Scaling behaviour for the water transport in nanoconfined geometry. *Nature Communications* (2014) **5**:4565.

[14] Fasano, M., Chiavazzo, E. and Asinari, P. Water transport control in carbon nanotube arrays. *Nanoscale Research letters* (2014) **9**:559.

[15] Denisov, V.P. and Halle, B. Protein hydration dynamics in aqueous solution. *Faraday Discussions* (1996) **103(0)**: 227–244.

[16] Tarek, M. and Tobias, D.J. The dynamics of protein hydration water: a quantitative comparison of molecular dynamics simulations and neutron-scattering experiments. *Biophysical journal* (2000) **79(6)**:3244–3257.

[17] Hess, B., Kutzner, C., Van der Spoel, D. and Lindahl, E. *J. Chem. Theory Comput.* (2008) **4**:435-447.

[18] www.rcsb.org

[19] MacKerell, AD.Jr, Feig, M. and Brooks, C.L. 3rd. Extending the treatment of backbone energetics in protein force fields: limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *J. Comp. Chem*. (2004) **25(11)**:1400–15.

[20] MacKerell et al., All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B*. (1998) **102(18)**:3586–3616.

[21] Berendsen, H.J.C., Grigera, J.R. and Strrtsma, T.P. The missing term in effective pair potentials. *J. Phys. Chem.* (1987) **91**:6269–6271.

[22] Bussi, G., Donadio, D. and Parrinello, M. *The Journal of Chemical Physics* (2007) **126**:014101.

[23] Parrinello, M. and Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys*. (1981) **52**:7182–7190.

[24] Nose, S. and Klein, M. L. Constant pressure molecular dynamics for molecular systems. *Mol. Phys*. (1983) **50**:1055–1076.

[25] Eisenhaber, F., Argos, P., Sander, C. and Scharf, M. *Comput. Chem*. (1995) **16**:273-284.

[26] Zhao, H. *Biophys. Chem.* (2006) **122**:157–183.

[27] Sterpone, F., Stirnemann, G., Hynes, J.T. and Laage, D. Water hydrogen-bond dynamics around amino acids: the key role of hydrophilic hydrogen-bond acceptor groups. (2010) **114**:2083-2089.

[28] Van der Spoel, D., van Maaren, P. J. and Berendsen, H. J. A systematic study of water models for molecular simulation: derivation of water models optimized for use with a reaction field. *J. Chem. Phys*. (1998) **108**:10220.

[29] Massi, F. and Straub, J.E. Structural and dynamical analysis of the hydration of the Alzheimer's beta-amyloid peptide. *J.Comput. Chem.* (2003) **24(2)**:143-53.

[30] Buchete, N.V., Tycko, R. and Hummer, G. Molecular dynamics simulations of Alzheimer's β-amyloid protofilaments. *J.Mol.Biol*. (2005) **363**:804-821.

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis(Eds)

3D PHASE-FIELD FOR PRESSURIZED FRACTURE PROPAGATION IN HETEROGENEOUS MEDIA

THOMAS WICK[*,,†], SANGHYUN LEE[†] AND MARY F. WHEELER[†]**

[*] Johann Radon Institute for Computational and Applied Mathematics,
Austrian Academy of Sciences,
4040 Linz, Austria
e-mail: thomas.wick@ricam.oeaw.ac.at,
web page: http://people.ricam.oeaw.ac.at/t.wick/

[**] Fakultät für Mathematik,
Technische Universität München,
85748 Garching bei München, Germany

[†] The Institute for Computational Engineering and Sciences,
The University of Texas at Austin,
Austin, Texas 78712, USA
e-mail: shlee@ices.utexas.edu
web page: http://users.ices.utexas.edu/∼shlee/

Key words: Phase-field, variational approach to fracture, pressurized cracks, porous media

Abstract. This work presents recent progress in phase-field-based fracture modeling in heterogeneous porous media. Existing algorithms that have been developed in the last years are extended to tackle three-dimensional configurations. Our solution technique is formulated in terms of a nested Newton loop that combines a primal-dual active set method (required for treating the crack irreversibility) and a Newton method to solve the nonlinear, fully-coupled PDE system. An advanced numerical test demonstrates the capabilities of our method.

1 INTRODUCTION

Crack propagation in elastic and porous media is currently one of the major research topics in mechanical, energy, and environmental engineering. In this paper, we concentrate specifically on pressurized-fracture propagation in heterogeneous media towards applications in porous media. Our method of choice is a variational approach for brittle fracture [9] that is formulated in terms of a phase-field technique. The methodology seems very promising as attested by several major publications in recent years [6, 7, 15, 14, 4, 2, 8].

Here, discontinuities in the displacement field across the lower-dimensional crack surface are approximated by an auxiliary phase-field function. This is an indicator function, which introduces a diffusive transition zone (brittle or mushy-zone are also common expressions depending on the discipline) between the broken and the unbroken material. Most importantly, fracture nucleation, propagation, kinking, and curvilinear paths are automatically included in the model; post-processing of stress intensity factors and remeshing resolving the crack path are avoided. Recent progress towards hydraulic fracturing scenarios has been made in [10] (using XFEM/GFEM), [13] (using peridynamics) [5, 1] (using variational fracture methods), and [18, 19] (phase-field methods that are related to variational techniques).

The crucial numerical aspects of a phase-field-based fracture propagation approach are techniques that include resolution of the length-scale parameter ε, the numerical solution of the forward problem and enforcement of the irreversibility of crack growth. The sum of these requirements leads to a variational inequality discretized in terms of an adaptive finite element scheme. A robust numerical framework has been proposed in [11] in which a primal-dual active set method [12] is combined with a Newton method for the forward problem and leading to a nested Newton loop.

The forward problem; namely, the coupled Euler-Lagrange PDE system is addressed monolithically in which both equations, elasticity and phase-field, are solved simultaneously. We propose a robust numerical scheme in terms of a Newton linearization to treat the non-convexity of the regularized energy functional and the consequences of non-convexity on the Euler-Lagrange system.

The **key aspect of this paper** is to extend these algorithms to three-dimensional applications and to demonstrate the potential of our method for two initially non-aligned cracks in a heterogeneous medium and their propagation leading two interaction through joining and branching. To the best of our knowledge such a result is novel for pressurized cracks computed with phase-field.

The outline of this work is as follows: We first state the governing equations in Section 2. Then, we briefly state our main algorithm in Section 3. A sophisticated 3D numerical example is presented in Section 4 in order to exploit the potential of our approach.

2 NOTATION AND EQUATIONS

In this section, we explain the governing equations. Here, we adopt an phase-field-based approach in terms of the Euler-Lagrange weak formulation for pressurized fractures that has been derived in [17, 18, 19]. The unknown solution variables are vector-valued displacements $\mathbf{u}$ and a smoothed indicator phase-field function φ with values in $[0, 1]$. For $\varphi = 0$, we denote a crack region and $\varphi = 1$ is the unbroken material. The intermediate values constitute a so-called transition zone that is dependent on a regularization (or length-scale) parameter ε. In particular, ε is the thickness of the transition zone. Due to the crack irreversibility condition,

$$\partial_t \varphi \leq 0,$$

the resulting system is a variational inequality.

Let $V := H_0^1(\Omega)$ and $W_{in} := \{w \in H^1(\Omega) | w \leq \varphi^{n-1} \leq 1 \text{ a.e. on } \Omega\}$ be the function spaces we work with; and for later purposes we also need $W := H^1(\Omega)$.

Formulation 1 (Variational inequality system for pressurized fractures) *Find* $(\mathbf{u}, \varphi) \in V \times W$ *with*

$$\Big(\big((1-\kappa)\varphi^2 + \kappa\big)\, \sigma(\mathbf{u}), e(\mathbf{w})\Big) - (\alpha - 1)(\varphi^2 p, div\ \mathbf{w}) = 0 \quad \forall \mathbf{w} \in V, \tag{1}$$

as well as

$$\begin{aligned} &(1-\kappa)(\varphi\, \sigma(\mathbf{u}) : e(\mathbf{u}), \psi - \varphi) - 2(\alpha - 1)(\varphi\, p\ div\ \mathbf{u}, \psi - \varphi) \\ &+ G_c\Big(-\frac{1}{\varepsilon}(1-\varphi, \psi - \varphi) + \varepsilon(\nabla\varphi, \nabla\psi - \varphi)\Big) \ \geq 0 \quad \forall\ \psi \in W_{in} \cap L^\infty(\Omega). \end{aligned} \tag{2}$$

where, $\sigma = \sigma(\mathbf{u})$ *is the stress tensor,*

$$\sigma := \sigma(\mathbf{u}) = 2\mu e(\mathbf{u}) + \lambda\ tr(e(\mathbf{u}))I,$$

and $e(\mathbf{u})$ *the symmetric strain tensor defined as*

$$e(\mathbf{u}) := \frac{1}{2}\left(\nabla\mathbf{u} + \nabla\mathbf{u}^T\right).$$

and μ *and* λ *are material parameters,* I *is the identity matrix,* G_c *is the energy release rate* $G_c > 0$*,* ε *is the previously mentioned crack surface regularization parameter, and finally* κ *is a positive regularization parameter for the elastic energy, with* $\kappa \ll \varepsilon$*.*

This system is based on a quasi-stationarity assumption in which the time t *(i.e. in a quasi-stationary setting* t *is more a load parameter rather than true 'time') enters through time-dependent boundary conditions, e.g.,* $\mathbf{u} = \mathbf{u}(t) = \mathbf{g}(t)$ *on* $\partial\Omega_D$ *with a prescribed boundary function* $g(t)$ *of Dirichlet-type or through time-dependent right hand side forces, e.g.,* $p := p(t)$ *as used in our numerical example.*

Remark 2.1 *In poroelastic applications, the pressure* p *is obtained from solving a pressure Darcy equation. This coupled system is solved with a fixed-stress iteration [21, 16] in which the pressure equation and elasticity equation are solved sequentially. Correct modeling of interface forces [17, 19]* $\varphi^2\nabla p$ *where the fluid pressure in the crack interacts with poroelastic fluid flow. However, for a given, piece-wise constant, pressure this term is zero and neglected in the rest of this paper. In addition, Biot's coefficient is chosen as* $\alpha = 0$*. Both terms appear in situations as investigated in [20, 22, 23].*

3 NUMERICAL APPROXIMATION

Our solution algorithm is based on a combined Newton iteration (to compute an update δU_k^h such that $U_{k+1}^h = U_k^h + \delta U_k^h$, where k is the Newton iteration number, and h the spatial discretization parameter) that assembles the crack irreversibility with the help of a semi-smooth Newton method and a second Newton iteration for the fully-coupled forward problem. The problem is quasi-static and temporal derivatives are approximated with a backward difference scheme. Spatial discretization is based on Galerkin finite elements using trilinear ansatz and test functions.

In the following, let the (discrete) residual be denoted by $R(U_k^h) = -A(U_k^h)(\Psi^h)$ with $U_k^h := \{\mathbf{u}_k^h, \varphi_k^h\} \in V^h \times W^h$ and (now neglecting the indices k and h):

$$\begin{aligned} A(U)(\Psi) = &\Big(\big((1-\kappa)\tilde{\varphi}^2 + \kappa\big)\,\sigma(\mathbf{u}), e(\mathbf{w})\Big) - (\alpha-1)(\tilde{\varphi}^2 p, \operatorname{div}\,\mathbf{w}) \\ &+ (1-\kappa)(\varphi\,\sigma(\mathbf{u}) : e(\mathbf{u}), \psi) - 2(\alpha-1)(\varphi\, p \operatorname{div}\,\mathbf{u}, \psi) \\ &+ G_c\Big(-\frac{1}{\varepsilon}(1-\varphi, \psi) + \varepsilon(\nabla\varphi, \nabla\psi)\Big) = 0 \quad \forall \Psi := \{\mathbf{w}, \psi\} \in V \times W. \end{aligned} \tag{3}$$

In this system, the well-known difficulty; namely the non-convexity of the first term on the energy level, $\Big(\big((1-\kappa)\varphi^2 + \kappa\big)\,\sigma(\mathbf{u}), e(\mathbf{w})\Big)$ is treated by replacing (the solution variable) φ by a (known) extrapolation $\tilde{\varphi}$ computed from two previous time steps. This heuristic procedure is very robust and yields excellent approximation properties with regard to the true φ. Verification in terms of benchmarks is provided in [11].

The corresponding Newton matrix is built by computing the directional derivative $A'(U_k^h)(\delta U_k^h, \Psi^h)$. Then, $\delta U_k^h := \{\delta\mathbf{u}_\mathbf{k}^\mathbf{h}, \delta\varphi_k^h\} \in V^h \times W^h$ such that

$$\begin{aligned} A'(U)(\delta U, \Psi) = &\Big(\big((1-\kappa)\tilde{\varphi}^2 + \kappa\big)\,\sigma(\delta\mathbf{u}), e(\mathbf{w})\Big) \\ &+ (1-\kappa)(\delta\varphi\sigma^+(\mathbf{u}) : e(\mathbf{u}) + 2\varphi\,\sigma(\delta\mathbf{u}) : e(\mathbf{u}), \psi) \\ &- 2(\alpha-1)p(\delta\varphi \operatorname{div}\,\mathbf{u} + \varphi\,\operatorname{div}\,\delta\mathbf{u}, \psi) \\ &+ G_c\Big(\frac{1}{\varepsilon}(\delta\varphi, \psi) + \varepsilon(\nabla\delta\varphi, \nabla\psi)\Big) = 0 \quad \forall \Psi := \{\mathbf{w}, \psi\} \in V \times W. \end{aligned} \tag{4}$$

Furthermore, $(B)_{ii}$ is a diagonal mass matrix [12].

The semi-smooth Newton method can be interpreted as a primal-dual active set method and vice versa [12]. With these preparations, we define our solution algorithm:

Algorithm 3.1 ([11]) *Repeat for $k = 0, 1, 2, \ldots$:*

1. *Assemble residual $R(U_k^h)$*
2. *Compute active set $\mathcal{A}_k = \{i \mid (B^{-1})_{ii}(R_k^h)_i + c(\delta U_k^h)_i > 0\}, \quad i: \text{ is a DoF}$*
3. *Assemble matrix $G = A'(U_k^h)(\cdot, \Psi^h)$ and right-hand side $F = -A(U_k^h)(\Psi^h)$*

4. *Eliminate rows and columns in* $\mathcal{A}_k$ *from* G *and* F *to obtain* $\widetilde{G}$ *and* $\widetilde{F}$

5. *Solve the linear system* $\widetilde{G}\delta U_k^h = \widetilde{F}$, *i.e,*

$$\widetilde{A}'(U_k^h)(\delta U_k^h, \Psi^h) = -\widetilde{A}(U_k^h)(\Psi^h) \quad \forall \Psi \in V^h \times W^h. \tag{5}$$

6. *Find a step size* $0 < \omega \leq 1$ *using line search to get*

$$U_{k+1}^h = U_k^h + \omega \delta U_k^h,$$

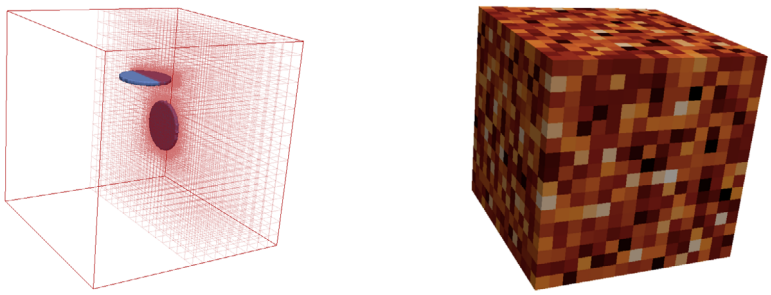
with $\widetilde{R}(U_{k+1}^h) < \widetilde{R}(U_k^h)$.

Finish if both stopping criteria are fulfilled simultaneously:

$$\mathcal{A}_{k+1} = \mathcal{A}_k \qquad \textit{and} \qquad \widetilde{R}(U_k^h) < \mathrm{TOL}\,.$$

4 NUMERICAL TEST: TWO PENNY-SHAPED FRACTURES IN A HETEROGENEOUS MEDIUM

In this final section, we present an advanced numerical example that demonstrates the capabilities for realistic scenarios of our method. This example is computed based on the finite element software deal.II [3] and based on an extension of the programming code developed for [11].



(a) Two initial fractures (b) Heterogeneity

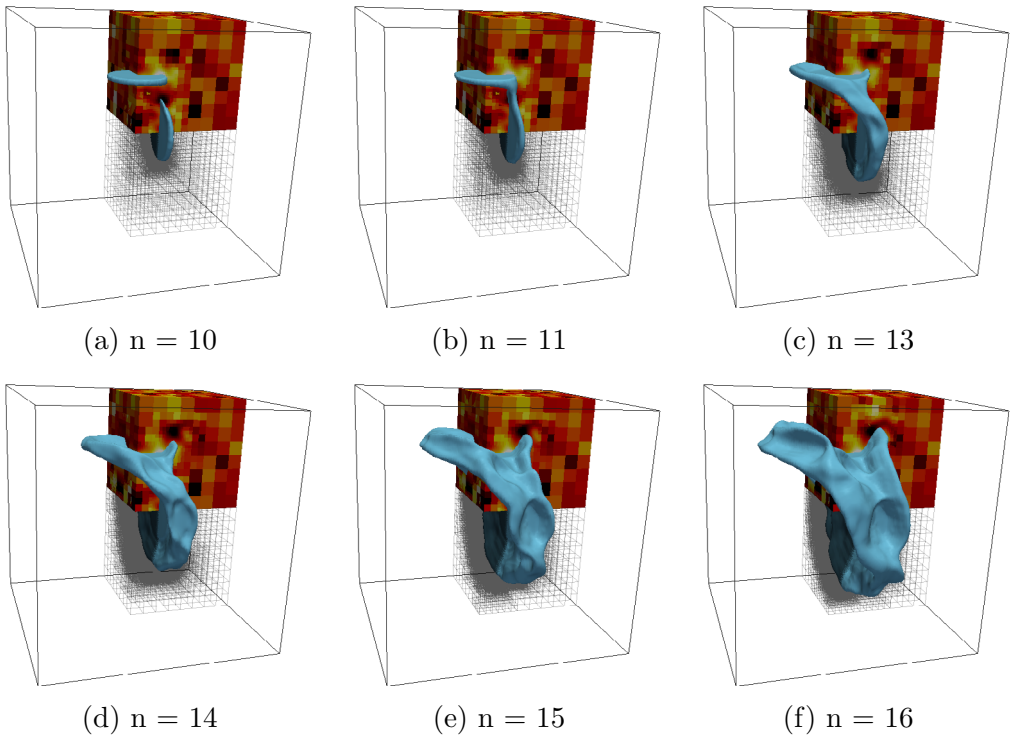
Figure 1: (a) Two initial penny shaped crack is position in the middle of the media. (b) Young's modulus(E) has the value range of the shale rock region; $E \in [1\,\mathrm{GPa}, 10\,\mathrm{GPa}]$ and $\nu = 0.15$.

We take a setting with two initial penny-shaped fractures in a 3D heterogeneous medium. The locally refined grid in the cube $\Omega = (0m, 4m)^3$ with hanging nodes is shown in Figure 1a. The cells around the crack interface are refined (if $\varphi(\mathbf{x}, t) < 0.8$) using the a predictor-corrector strategy [11]. In Figure 1a two initial fractures are described. The top fracture is centered at $(2, 3, 2)[m]$ with radius $r = 0.5m$ in $y = 3m$-plane and

the bottom fracture is centered at $(2.5, 2, 2)[m]$ with radius $r = 0.5m$ in $x = 2.5m$-plane. The crack is approximated as a volume by extending it with the spatial discretization parameter h. The thickness of both fractures are $2h_{\min}$. As outer boundary conditions we set the displacements zero on $\partial\Omega$.

The pressure is constant in space with $\alpha = 0$ in (1), but it is linearly increasing in time as $p = t \times 1\,\mathrm{MPa}[Pa]$, where t is the current time. The numerical constants are given as $k = 10^{-10} \times h_{\min}$ and $\varepsilon = 2h_{\min}$. Here $h_{\min} = 0.0541266m$ and the time steps are $dt = 0.01s$.

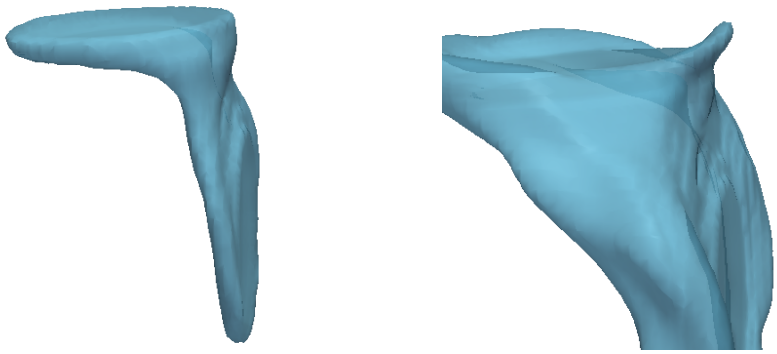
We observe joining and branching of non-planar fractures in heterogeneous media. Those are automatically captured by the phase field method. See Figure 2 and Figure 3 for the details. Finally, in Figure 3, the bulk and crack energies are observed. We clearly see that the crack energy remains constant while the cracks are not growing and this energy increases for growing fractures.



(a) n = 10 (b) n = 11 (c) n = 13

(d) n = 14 (e) n = 15 (f) n = 16

Figure 2: Sequence of snapshots of fractures propagating at each time step number n in 3D heterogeneous media. The blue regions corresponds to $\varphi < 0.1$. Both fractures grow non-planarly, then they join at $n = 11$ and start branching at $n = 13$.



(a) Joining of two fractures (b) Branching after joining

Figure 3: Detailed snapshots of the areas where the cracks are (a) joining at n=11 and (b) branching at n=13.

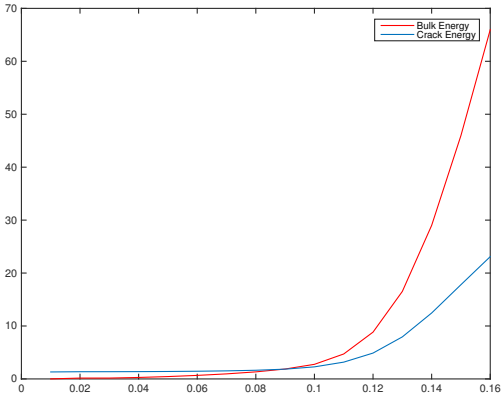


Figure 4: Evolution of bulk and crack energies. The crack energy starts increasing when the two fractures start growing at $n = 10$ (equal to $t = 0.1$). The bulk energy remains increasing even for propagating fractures since the applied pressure is still advanced.

5 CONCLUSIONS

In this work, we applied recently developed algorithms to compute three-dimensional pressurized-fracture propagation in heterogeneous media. The proposed method is efficient and robust. The novelty in comparison to existent literature is a detailed numerical simulation to show three-dimensional complex crack patterns, their forming, growth and joining.

REFERENCES

[1] Stefano Almi, Gianni Dal Maso, and Rodica Toader. Quasi-static crack growth in hydraulic fracture, 2014.

[2] M. Artina, M. Fornasier, S. Micheletti, and S. Perotto. Anisotropic mesh adaptation for crack detection in brittle materials. TUM Report, 2014.

[3] Wolfgang Bangerth, Timo Heister, Guido Kanschat, and Many others. *Differential Equations Analysis Library*, 2012.

[4] M. J. Borden, C. V. Verhoosel, M. A Scott, T. J. R. Hughes, and C. M. Landis. A phase-field description of dynamic brittle fracture. *Comput. Meth. Appl. Mech. Engrg.*, 217:77–95, 2012.

[5] B. Bourdin, C. Chukwudozie, and K. Yoshioka. A variational approach to the numerical simulation of hydraulic fracturing. SPE Journal, Conference Paper 159154-MS, 2012.

[6] B. Bourdin, G.A. Francfort, and J.-J. Marigo. Numerical experiments in revisited brittle fracture. *J. Mech. Phys. Solids*, 48(4):797–826, 2000.

[7] B. Bourdin, G.A. Francfort, and J.-J. Marigo. The variational approach to fracture. *J. Elasticity*, 91(1–3):1–148, 2008.

[8] S. Burke, Ch. Ortner, and E. Süli. An adaptive finite element approximation of a generalized Ambrosio-Tortorelli functional. *M3AS*, 23(9):1663–1697, 2013.

[9] G.A. Francfort and J.-J. Marigo. Revisiting brittle fracture as an energy minimization problem. *J. Mech. Phys. Solids*, 46(8):1319–1342, 1998.

[10] P. Gupta and C. A. Duarte. Simulation of non-planar three-dimensional hydraulic fracture propagation. *International Journal for Numerical and Analytical Methods in Geomechanics*, 38(13):1397–1430, 2014.

[11] T. Heister, M.F. Wheeler, and T. Wick. A primal-dual active set method and predictor-corrector mesh adaptivity for computing fracture propagation using a phase-field approach. ICES Report 14-27, Aug 2014.

[12] Michael Hintermüller, Kazufumi Ito, and Karl Kunisch. The primal-dual active set strategy as a semismooth newton method. *SIAM Journal on Optimization*, 13(3):865–888, 2002.

[13] Amit Katiyar, John T. Foster, Hisanao Ouchi, and Mukul M. Sharma. A peridynamic formulation of pressure driven convective fluid transport in porous media. *Journal of Computational Physics*, 261(0):209 – 229, 2014.

[14] C. Miehe, M. Hofacker, and F. Welschinger. A phase field model for rate-independent crack propagation: Robust algorithmic implementation based on operator splits. *Comput. Meth. Appl. Mech. Engrg.*, 199:2765–2778, 2010.

[15] C. Miehe, F. Welschinger, and M. Hofacker. Thermodynamically consistent phase-field models of fracture: variational principles and multi-field fe implementations. *International Journal of Numerical Methods in Engineering*, 83:1273–1311, 2010.

[16] A. Mikelić and M. F. Wheeler. Convergence of iterative coupling for coupled flow and geomechanics. *Comput Geosci*, 17(3):455–462, 2012.

[17] A. Mikelić, M.F. Wheeler, and T. Wick. A phase-field approach to the fluid filled fracture surrounded by a poroelastic medium. ICES Report 13-15, Jun 2013.

[18] A. Mikelić, M.F. Wheeler, and T. Wick. A quasi-static phase-field approach to the fluid filled fracture. ICES Report 13-22, submitted for publication, Aug 2013.

[19] A. Mikelić, M.F. Wheeler, and T. Wick. Phase-field modeling of pressurized fractures in a poroelastic medium. ICES-Preprint 14-18, Jul 2014.

[20] A. Mikelić, M.F. Wheeler, and T. Wick. A phase-field method for propagating fluid-filled fractures coupled to a surrounding porous medium. accepted for publication in SIAM Multiscale Modeling and Simulation, ICES Report 14-08, Jan 2015.

[21] A Settari and D A Walters. Advances in coupled geomechanical and reservoir modeling with applications to reservoir compaction. *SPE Journal*, 6(3):334–342, September 2001.

[22] T. Wick, G. Singh, and M.F. Wheeler. Pressurized fracture propagation using a phase-field approach coupled to a reservoir simulator. SPE 168597-MS, SPE Proc., 2013.

[23] T. Wick, G. Singh, and M.F. Wheeler. Fluid-filled fracture propagation using a phase-field approach and coupling to a reservoir simulator. ICES report 14-20, Aug 2014.

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis (Eds)

AN ELECTRO-CHEMO-MECHANICAL ANALYSIS OF SOLID OXIDE FUEL CELL CONSIDERING EVOLUTION OF MICROSTRUCTURE IN POROUS ELECTRODE USING PHASE-FIELD METHOD

M. MURAMATSU [*], S. TAKASE [**], K. YASHIRO [*], T. KAWADA [*] AND K. TERADA [***]

[*] Graduate School of Environmental Studies
Tohoku University
468-1-404, Aoba, Aramaki, Aoba-ku, Sendai 980-0845 Japan
e-mail: muramatsu@ee.mech.tohoku.ac.jp

[**] Department of Civil and Environmental Engineering
Tohoku University
468-1-404, Aoba, Aramaki, Aoba-ku, Sendai 980-0845 Japan
email: takase@irides.tohoku.ac.jp

[***] International Research Institute of Disaster Science
Tohoku University
468-1-404, Aoba, Aramaki, Aoba-ku, Sendai 980-0845 Japan
email: tei@irides.tohoku.ac.jp

Key words: Solid Oxide Fuel Cells, Potential Simulation, Stress Analysis, Microstructure.

Abstract. For the numerical simulation of oxygen potential distributions in Solid Oxide Fuel Cell (SOFC), the time-evolution of the anode microstructure is reflected in the macroscopic electrical conductivities and the amount of triple-phase boundaries. Once the oxygen potential distributions are determined, the time-variation of the reduction-induced strains due to nonstoichiometry of oxide materials is calculated along with the thermal strains. These strains cause the macroscopic stresses in mutually constrained components. Thus, the capability of the proposed method is demonstrated in characterizing the aging degradation of the macroscopic electro-chemo-mechanical behavior of SOFC that is caused by the Ni-sintering in cermet microstructures during long-period control.

1 INTRODUCTION

Components of Solid Oxide Fuel Cells (SOFCs) are always exposed to high temperature and large gas pressure under operation. Subjected to the starting and stopping control, the performance of SOFC is gradually degraded. One of the degradation factors can be the thermal expansive deformation due to temperature change [1], but the expansive deformation under reduction environment is distinguishing [1] for SOFC. These stress-free deformations

inevitably invoke the unexpectedly large stress due to the mutual constraints of the components, which sometimes causes the mechanical deterioration. At the same time, Ni particles in the porous anode electrode of SOFC are subjected to coursing due to sintering under high temperature environment. The temporal changes of the anode microstructures cause the degradation of not only the overall mechanical properties, but also the electrical performance.

In order to realize the macroscopic electro-chemo-mechanical coupling analyses of a SOFC under operation, we propose a characterization method of the time-varying overall or macroscopic electro-chemical and mechanical properties of anode electrodes by applying the phase-field method that enables us to capture the time-varying geometry of anode microstructures due to Ni-sintering. With this information at hand, the homogenization method is extensively applied to evaluate the temporal change of the macroscopic electro-chemo-mechanical properties that characterize the macroscopic inelastic mechanical behaviour and the oxygen potential distribution in SOFC.

For the numerical simulation of oxygen potential distributions, the time-evolution of the anode microstructure is reflected not only in the macroscopic electrical conductivities, but also the amount of triple-phase boundaries, which are the generation sites of electro-chemical reaction currents. Once the oxygen potential distributions are determined, the time-variation of the reduction-induced strains due to nonstoichiometry of oxide materials can be calculated along with the thermal strains, both of which cause the macroscopic stresses in mutually constrained components. Thus, the promise and capability of the proposed method can be demonstrated in characterizing the aging degradation of the macroscopic electro-chemo-mechanical behaviour of SOFC that is caused by the Ni-sintering in cermet microstructures during long-period control.

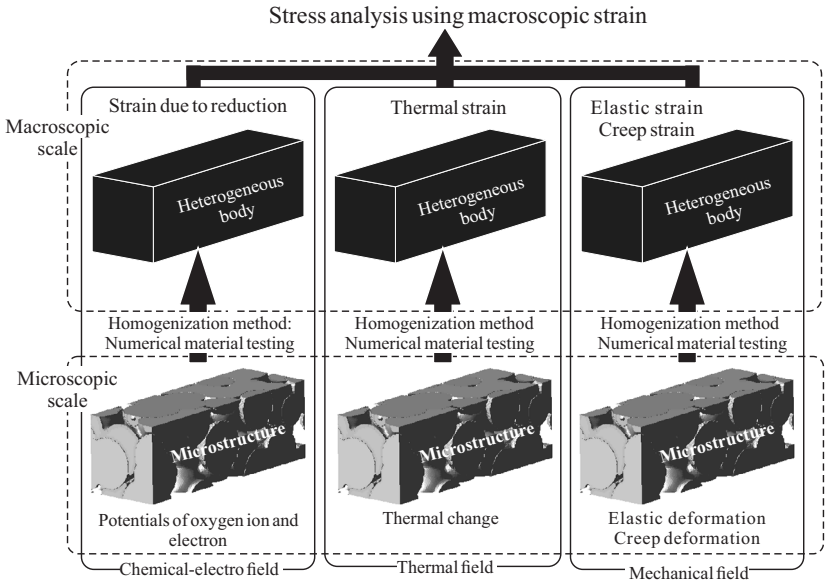


Figure 1: Framework of analysis

2 FORMULATION

2.1 Electro-chemical phenomenon

At the cathode electrode, oxygen O_2 is reduced by electron e- to produce oxygen ion O^{2-}, which can move through the electrolyte toward the anode electrode. The oxygen ion that reaches the anode electrode is oxidized by the reaction with H_2, which is provided as a fuel, to produce water H_2O and e^-.

$$\text{Cathode}: \frac{1}{2}O_2 + 2e^- \rightarrow O^{2-} \tag{1}$$

$$\text{Anode}: O^{2-} + H_2 = H_2O + 2e^- \tag{2}$$

Since the electrodes are porous materials, actual reactions are expected to occur at the interfaces between gas and solid phases. The latter electro-chemical reaction, i.e., oxidation with emission of electrons, occurs at the so-called triple phase boundaries (TPBs) of the anode electrode, which is generally a cermet composed of Ni particles and ion conductive ceramics. On the other hand, the cathode electrode, whose constituents are mixed ionic-electronic conductive ceramics, allows the reduction of O_2 at the so-called dual phase boundaries (DPBs), involving the ionization of the oxygen with absorption of electrons.

The electronic current density $i_{e^-}^{\text{Anode}}$ generated with the oxidation (2) is known to be determined by the Butler-Volmer relationship [3,4] as

$$i_{e^-}^{\text{Anode}} = i_0^{\text{Anode}}[\exp\{(2\alpha_a F)/(R\theta)\Delta E\} - \exp\{(2\alpha_c F)/(R\theta)\Delta E\}] \tag{3}$$

where $F = 96500\,\text{Cmol}^{-1}$ is the Faraday's constant, R is the gas constant, θ is the temperature and i_0^{Anode} is the exchange current density, α_a and α_c are the anodic and cathodic charge transfer coefficients. In addition, ΔE is the polarization voltage expressed by

$$\Delta E = (\mu_O - \mu_O^{\text{gas}})/2F \tag{4}$$

where μ_O and μ_O^{gas} are the oxygen potentials in material and gas, respectively. The oxygen potential is calculated with the partial pressure of oxygen p_{O_2} as

$$\mu_O = \mu_O^\circ + \frac{1}{2}R\theta\ln(p_{O_2}) \tag{5}$$

where $\mu_O^\circ = 0$ is the oxygen potential at standard condition. The normal component of current density i_{e^-} inside the Ni particles equals to the generated electric current density as

$$i_{e^-} = \boldsymbol{i}_{e^-} \cdot \boldsymbol{n} = i_{e^-}^{\text{Anode}} \tag{6}$$

where $\boldsymbol{n}$ is the outward unit vector on the TPB, directed to the normal to Ni surface. According to the electro-chemical reaction Eq. (2), the oxygen ionic current density generated at the same boundary surfaces can be evaluated as

$$i_{O^{2-}} = \boldsymbol{i}_{O^{2-}} \cdot \boldsymbol{n}' = i_{O^{2-}}^{\text{Anode}} = -i_{e^-}^{\text{Anode}} \tag{7}$$

where the outward unit vector $\boldsymbol{n}'$ is defined to the TPB, directed to the normal to the oxide's surface.

Table 1: Electric conductivities and oxygen vacancy

	σ_{e^-}	$\sigma_{O^{2-}}$	δ
LSCF (Cathode)	$354.45-2379.7\delta$	$\frac{4F^2}{V_m}\frac{D_v\delta}{R\theta}$ $D_v = 3.0\times10^{-1}\exp\left(-\frac{99280}{R\theta}\right)$	$\log p_{O_2} = \log\left(\frac{(3-\delta)^2([Sr]-2\delta)^4}{\delta^2\exp\left(\frac{-2(h_0-\theta s_0)}{R\theta}\right)}\exp\left(\frac{2(a_0k+\theta)}{R\theta}\right)\right)$ $h_0 = -79553$J/ mol, $s_0 = -48.54$J/ (mol·K), $a_0 = -610990$J/ mol, $k = 512.06$J/(mol·K), [Sr]=0.4
8YSZ (Electrolyte)	$1.31\times10^7\exp\left(-3.88\frac{F}{R\theta}\right)p_{O^2}^{-1/4}$	$1.63\times10^2\exp\left(-0.79\frac{F}{R\theta}\right)$	Almost no oxygen vacancy
Ni-YSZ (Anode)	$\sigma_{e^-}^{YSZ} = 6.6\times10^2\exp\left(-1.9\frac{e}{k\theta}\right)p_{O^2}^{-1/4}$ $+3.7\times10^6\exp\left(-3.7\frac{e}{k\theta}\right)p_{O^2}^{-1/4}$ $\sigma_{e^-}^{Ni} = 1.45\times10^4$	$\frac{9.7\times10^5}{\theta}\exp\left(-\frac{e}{k\theta}\right)$	Almost no oxygen vacancy
LCCr (Interconnect)	$p_{O_2} = \frac{\{(6+x)\mu F-\sigma_e V_m\}^2(\sigma_e V_m)^4}{K_{0x}^2(x\mu F-\sigma_e V_m)^2(\mu F-\sigma_e V_m)^4}$ $\mu = \frac{267.28086}{\theta}\exp\left(-\frac{11584.73}{R\theta}\right)$ $K_{0x} = \frac{0.000155}{\theta}\exp\left(-\frac{261026.3}{R\theta}\right)$	$\frac{4F^2}{V_m}\frac{D_v\delta}{R\theta}$ $D_v = 0.028883\exp\left(-\frac{76571.94}{R\theta}\right)$	$p_{O_2} = \left[\frac{1}{K}\left(\frac{3-\delta}{\delta}\right)\left(\frac{x-2\delta}{1-x+2\delta}\right)^2\exp\left(\frac{a\delta}{R\theta}\right)\right]^2$ $K = 4.227\times10^{-3}\exp\left(\frac{218410}{R\theta}\right)$, $a = 407.22\theta - 430048$, $x = [Ca] = 0.3$

The current density generated at the cathode with the reduction process expressed by Eq. (1) is also determined by the Butler-Volmer type equation as

$$i_{e^-}^{\text{Cathode}} = i_0^{\text{Cathode}}[\exp\{(2(1-\alpha)F)/(R\theta)\Delta E\} - \exp\{(2\alpha F)/(R\theta)\Delta E\}] \tag{8}$$

where the exchange current density i_0^{Cathode} and the charge transfer coefficient α are different from those for anode.. The relationships of this generated current density with the normal components of the current densities in the oxide are given as

$$i_{e^-} = \boldsymbol{i}_{e^-}\cdot\boldsymbol{n} = i_{e^-}^{\text{Cathode}} \tag{9}$$

$$i_{O^{2-}} = \boldsymbol{i}_{O^{2-}}\cdot\boldsymbol{n}' = i_{O^{2-}}^{\text{Cathode}} = -i_{e^-}^{\text{Cathode}} \tag{10}$$

2.2 Microscopic electronic and oxygen-ionic conduction problems

The electrical conduction in mixed ionic-electronic conductive ceramics can be characterized by the local transport of electro-chemical potentials and is governed by the following equations [5]:

$$-\frac{c}{2F}\dot{\mu}_O = -\nabla\cdot\boldsymbol{i}_{O^{2-}},\ \boldsymbol{i}_{O^{2-}} = -\frac{\sigma_{O^{2-}}}{-2F}\nabla\eta_{O^{2-}} \tag{11}$$

$$-\frac{c}{2F}(-\dot{\mu}_O) = -\nabla\cdot\boldsymbol{i}_{e^-},\ \boldsymbol{i}_{e^-} = -\frac{\sigma_{e^-}}{2F}(\nabla 2\eta_{e^-}) \tag{12}$$

$$\mu_O = \eta_{O^{2-}} - 2\eta_{e^-},\ \boldsymbol{i} = \boldsymbol{i}_{O^{2-}} + \boldsymbol{i}_{e^-} \tag{13}$$

where $\eta_{O^{2-}}$ is the electro-chemical potential of oxygen ion, η_{e^-} is the electro-chemical potential of electron, $c = c(\mu_O) = (-4F^2/V_m)(\partial\delta(\mu_O)/\partial\mu_O)$ is the electric capacitance, δ is the oxygen vacancy, is the molal volume, $\boldsymbol{i}$ is the total current density, and $\sigma_{O^{2-}}$ and σ_{e^-} are

the electric conductivities of oxygen ion and electron, respectively. It is noted that the source terms of reaction current densities are not appeared in Eqs. (11) and (12). The reaction currents are generated at the TPBs or DPBs. The standard Dirichlet boundary conditions with the values of electro-chemical potentials are used for the pore surface in the microscopic problem.

2.3 Macroscopic electronic and oxygen-ionic conduction problems

The local transport of macroscopic electro-chemical potentials is governed by the following equations:

$$-\frac{\tilde{c}}{2F}\dot{\tilde{\mu}}_{\mathrm{O}} = -\nabla\cdot\tilde{\boldsymbol{i}}_{\mathrm{O}^{2-}} + \tilde{j}_{\mathrm{O}^{2-}},\ \tilde{\boldsymbol{i}}_{\mathrm{O}^{2-}} = -\frac{\tilde{\sigma}_{\mathrm{O}^{2-}}}{-2F}\nabla\tilde{\eta}_{\mathrm{O}^{2-}} \tag{14}$$

$$-\frac{\tilde{c}}{2F}(-\dot{\tilde{\mu}}_{\mathrm{O}}) = -\nabla\cdot\tilde{\boldsymbol{i}}_{\mathrm{e}^-} + \tilde{j}_{\mathrm{e}^-},\ \tilde{\boldsymbol{i}}_{\mathrm{e}^-} = -\frac{\tilde{\sigma}_{\mathrm{e}^-}}{2F}(\nabla 2\tilde{\eta}_{\mathrm{e}^-}) \tag{15}$$

$$\tilde{\mu}_{\mathrm{O}} = \tilde{\eta}_{\mathrm{O}^{2-}} - 2\tilde{\eta}_{\mathrm{e}^-},\ \tilde{\boldsymbol{i}} = \tilde{\boldsymbol{i}}_{\mathrm{O}^{2-}} + \tilde{\boldsymbol{i}}_{\mathrm{e}^-} \tag{16}$$

where symbol $\tilde{\bullet}$ indicates the macroscopic (homogenized) quantity. The reaction current densities $\tilde{j}_{\mathrm{O}^{2-}}$ and $\tilde{j}_{\mathrm{e}^-}$ are appeared as source terms in Eqs. (14) and (15). Once the solutions of these equations $\tilde{\eta}_{\mathrm{O}^{2-}}$ and $\tilde{\eta}_{\mathrm{e}^-}$ are determined, the macroscopic oxygen potential $\tilde{\mu}_{\mathrm{O}}$ can be obtained.

2.4 Macroscopic deformation problem

The macroscopic (homogenized) governing equations of deformation problem are written in the following forms:

$$\nabla\cdot\tilde{\boldsymbol{T}} + \tilde{\rho}\tilde{\boldsymbol{b}} = 0 \tag{17}$$

$$\tilde{\boldsymbol{\varepsilon}} = \tilde{\boldsymbol{\varepsilon}}^{\mathrm{e}} + \tilde{\boldsymbol{\varepsilon}}^{\mathrm{c}} + \tilde{\boldsymbol{\varepsilon}}^{\mathrm{r}} + \tilde{\boldsymbol{\varepsilon}}^{\theta} \tag{18}$$

$$\tilde{\boldsymbol{T}} = \tilde{\boldsymbol{C}} : \boldsymbol{\varepsilon}^{\mathrm{e}} = \tilde{\kappa}(\mathrm{tr}\,\tilde{\boldsymbol{\varepsilon}}^{\mathrm{e}})\boldsymbol{I} + 2\tilde{\mu}\tilde{\boldsymbol{\varepsilon}}^{\mathrm{e}}_{\mathrm{dev}} \tag{19}$$

$$\tilde{\boldsymbol{\varepsilon}}^{\mathrm{c}} = \dot{\tilde{\gamma}}\tilde{\boldsymbol{T}}_{\mathrm{m}} / \| \tilde{\boldsymbol{T}}_{\mathrm{m}} \|,\ \dot{\tilde{\gamma}} = C_1 \| \tilde{\boldsymbol{T}} \|^{C_2} \exp(-C_3 / \tilde{\theta}) \tag{20}$$

$$\tilde{\boldsymbol{\varepsilon}}^{\theta} = \tilde{\boldsymbol{\alpha}}\Delta\theta \tag{21}$$

$$\tilde{\boldsymbol{\varepsilon}}^{\mathrm{r}} = \tilde{\boldsymbol{\beta}}\Delta\tilde{\delta} \tag{22}$$

where symbol $\tilde{\bullet}$ indicates the macroscopic (homogenized) quantity. The macroscopic reduction-induced strain ε^{r} is evaluated by the multiplication of the macroscopic vacancy $\tilde{\delta}$ from Table 1 by the coefficient β. The standard Dirichlet and Neumann boundary conditions are used.

2.5 Microstructures of anode

In this study, we assume that the main factor that deteriorates the anode made by Ni-YSZ is the sintering of Ni particles, which aggregate during the steady operation at 973-1173K. In particles touching with each other at lower temperature than the melting point, material transfer occurs so as to reduce the surface energy of the system. With the driving force of the

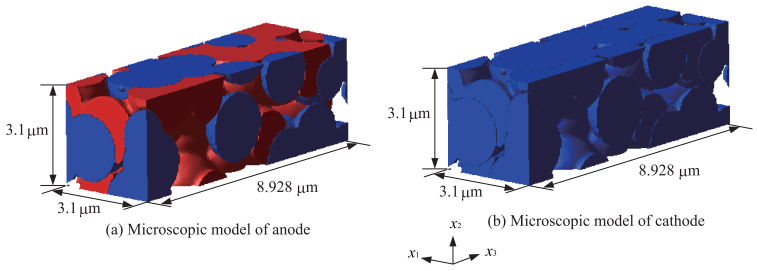


Figure 2: Model of microstructure of electrodes

surface energy, the shapes of particles change with time. To simulate such a time-variation, a phase-field model for sintering [6,7] is employed in this study. We employ the following Ginzburg-Landau type free energy [8]:

$$F = \int_{\mathscr{V}} [\sum_{\mathrm{i}} \frac{1}{2} \alpha_{\mathrm{i}} \| \nabla \phi_{\mathrm{i}} \|^2 + \sum_{\mathrm{i}} A_{\mathrm{i}} \phi_{\mathrm{i}}^2 (1-\phi_{\mathrm{i}})^2 + \sum_{i} B_{\mathrm{i}} \phi_{\mathrm{i}}^2 + \sum_{\mathrm{i}} (\sum_{\mathrm{i} \neq \mathrm{j}} \frac{\beta_{\mathrm{ij}}}{4} \phi_{\mathrm{i}}^2 \phi_{\mathrm{j}}^2) + \frac{\gamma}{2} \phi_{\mathrm{i}}^2 \phi_{\mathrm{j}}^2 \phi_{\mathrm{k}}^2] dv \tag{23}$$

where index i takes 1, 2 or 3, which indicates the phase of Ni, YSZ and void. Here, the first and second terms are the surface energies, the third term is the chemical free energy, and the fourth term is the interface energy of the TPB. Also, coefficients α_{i}, A_{i}, B_{i}, $\beta_{\mathrm{ij}} = \beta_{\mathrm{ji}}$ and γ are the material constants associated with the corresponding energies or potentials. The Cahn-Hilliard equations are used for the evolution equation because ϕ_{i} is a conservative quantity.

$$\frac{\partial \phi_{\mathrm{i}}}{\partial t} = \nabla \cdot \left\{ M_{\mathrm{i}}(\phi_{\mathrm{i}}, \theta) \left(\nabla \frac{\delta F}{\delta \phi_{\mathrm{i}}} \right) \right\} \tag{24}$$

where $M_{\mathrm{i}}(\phi_{\mathrm{i}}, \theta)$ is the mobility defined as $M_{\mathrm{i}}(\phi_{\mathrm{i}}, \theta) = D(\phi_{\mathrm{i}}) m_{\mathrm{i}}(\theta)$ with the diffusivity function $D(\phi_{\mathrm{i}})$ and the mobility function $m_{\mathrm{i}}(\theta)$, i.e., $D(\phi_{\mathrm{i}}) = D_{\mathrm{vol}} f(\phi_{\mathrm{i}}) + D_{\mathrm{vap}}(1 - f(\phi_{\mathrm{i}})) + D_{\mathrm{surf}} \phi_{\mathrm{i}}(1-\phi_{\mathrm{i}})$ and $m_{\mathrm{i}}(\theta) = 10^8 / 2[1 + \tanh\{(\theta - \alpha_{\mathrm{i}}) / b_{\mathrm{i}}\}]$ where D_{vol}, D_{vap} and D_{surf} are the diffusivities for the volumetric diffusion, the gas diffusion and the surface diffusion, respectively, and a_{i} and b_{i} are the mobility parameters and $f(\phi_{\mathrm{i}})$ takes the form of $f(\phi_{\mathrm{i}}) = \phi_{\mathrm{i}}^3 (10 - 15\phi_{\mathrm{i}} + 6\phi_{\mathrm{i}}^2)$. Substitution of $M_{\mathrm{i}}(\phi_{\mathrm{i}}, \theta)$ into Eq. (24) yields the following equation:

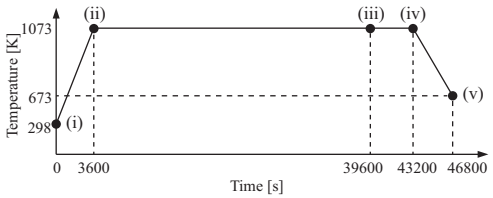
$$\frac{\partial \phi_{\mathrm{i}}}{\partial t} = m_{\mathrm{i}}(\theta) \left\{ (\nabla D_{\mathrm{i}}(\phi_{\mathrm{i}})) \cdot \left(\nabla \frac{\delta F}{\delta \phi_{\mathrm{i}}} \right) + D_{\mathrm{i}}(\phi_{\mathrm{i}}) \left(\nabla^2 \frac{\delta F}{\delta \phi_{\mathrm{i}}} \right) \right\} \tag{25}$$

Discretizing Eq. (25) with finite differential method, we simulate the development of order parameters.

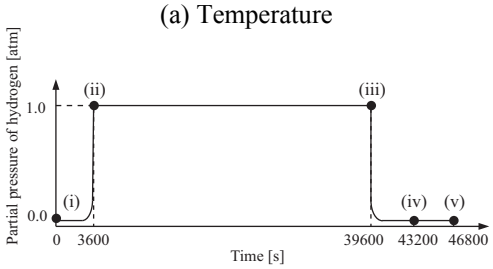
3 ANALYSIS CONDITION

3.1 Material constants

We employ $La_{0.6}Sr_{0.4}Co_{0.2}Fe_{0.8}O_{3-\delta}$ (LSCF) for the cathode, 8mol% yttria-stabilized zirconia (8YSZ) for the electrolyte, Ni-YSZ cermet for the anode La(Ca)CrO_3 (LCCr) for the



(a) Temperature



(b) Hydrogen gas

Figure 3: Control condition of the cell

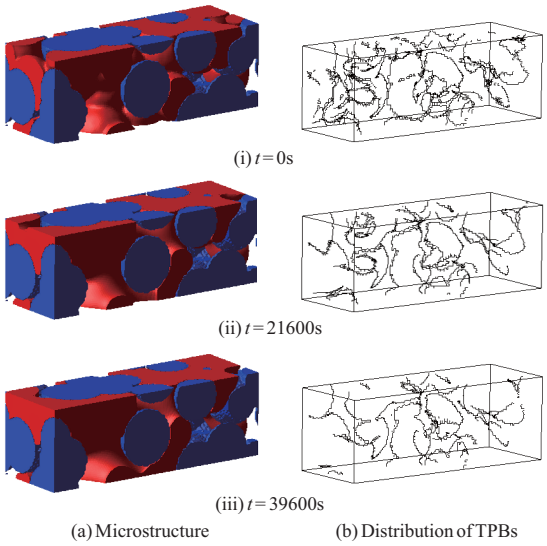


Figure 4: Temporal changes of microstructure

interconnect in this study. Note that the formulation does not limit the materials of cells. To evaluate the macroscopic material properties of the porous anode and cathode electrodes by performing the homogenization analyses, homogenization method are used to obtain the macroscopic material properties. In this study, numerical material testing [2] is employed for homogenization. Especially, the microstructure of anode, which is Ni-YSZ cermet, changes temporally due to Ni-sintering. The macroscopic properties of anode are calculated on the basis of the temporal microstructures.

3.2 Temporal change of microstructure in anode

Figure 2 shows the microscopic analysis models of anode (Ni-YSZ) and cathode (LSCF) microstructures, which are generated by the 3-dimensional porous material simulator POCO [9]. A phase-field simulation with the above-described evolution equations is carried out only for the Ni particles in Fig. 2 (a) that are assumed to be exclusively sintered. The finite difference model is generated with 5050 146 grids ($\Delta x \times \Delta y \times \Delta z$ = 0.062μm × 0.062μm × 0.062μm). The material constants are chosen as $\alpha_1 = \alpha_2 = 14.0\times10^{-2}$ J/cm^3, $A_1 = A_2 = 22.0\times10^{-3}$ J/cm^3, $B_1 = B_2 = 1.0$ J/cm^3, $\beta_{12} = 5.0\times10^{-3}$ J/cm^3, $\beta_{13} = \beta_{23} = 0.0$ J/cm^3, = 100.0×10^{-3}J/cm3, $D_{vol} = 5.0\times10^{-13}$ cm^2/s , $D_{vap} = 5.0\times10^{-15}$ cm^2/s and $D_{surf} = 2.0\times10^{-10}$ cm^2/s. It is noted here that the diffusivity of the interface is set at a value larger than the ones for the solid and gas phases to promote the sintering near the interfaces. Figure 3 is the schematic of the condition of temperature control in a virtual SOFC operation. After the temperature increases to 1073K, this constant value of temperature is kept for 11 hours, and then decreases to the room temperature by natural cooling. Under this temperature control, the sintering simulation is assumed to start when the temperature reaches at 1073K, since Ni particles are sintered mainly in steady operations [1]. Here, the simulation time is 11h with Δt = 1s. The results of the sintering simulation of Ni particles are provided in Fig. 4 that shows the time-variation of the anode microstructure with the TPBs, which are generation sites of reaction currents. The

amount of the TPBs shows that the frequency of TPBs is reduced with time during the sintering of Ni particles.

3.3 Homogenization

Based on the geometrical and material information about anode, which is shown in section 3.1, and cathode, numerical material tests are conducted to evaluate the macroscopic properties such as oxygen ionic and electronic conductivities, elastic constants, creep parameters and coefficient of thermal expansion. The obtained properties are used for the macroscopic analysis.

4 ANALYSIS RESULT

4.1 Electric conduction problem

Figure 5 is the analysis model and boundary conditions for macroscopic electric conduction problem. Figures 6, 7 and 8 show the time variations of the distributions of the potentials of oxygen ion, electron and oxygen, respectively. The contour plot on the left of each of these figures (Figures 6(a), 7(a) and 8(a)) shows their time varying distributions in the x_1-x_2 plane, while the graph on the right of each figure (Figures 6(b), 7(b) and 8(b)) shows their time-varying profiles on line P-Q indicated in Fig. 5. An increase of the uniform distributions of the electrochemical potential of electrons, which can be observed in Figs. (i) and (ii), reflects the supplying condition of fuel in the anode in the numerical analysis. After 3,600s, the potential is kept constant as can be seen from Figs. (ii) and (iii), realizing the steady states.

4.2 Deformation problem

The analytic model and boundary conditions for macroscopic analysis for the macroscopic deformation problem are shown in Fig. 9. As can be seen from Fig. 10, at first, the overall cell structure is bent toward the interconnect side This initial bending until the temperature reaches at 1073K has been caused by the significant development of larger thermal expansion strains in the cathode than those in the interconnect. Subsequently, the deflection is gradually decreased and the head of the cell moves toward the cathode side during 10h-steady operation

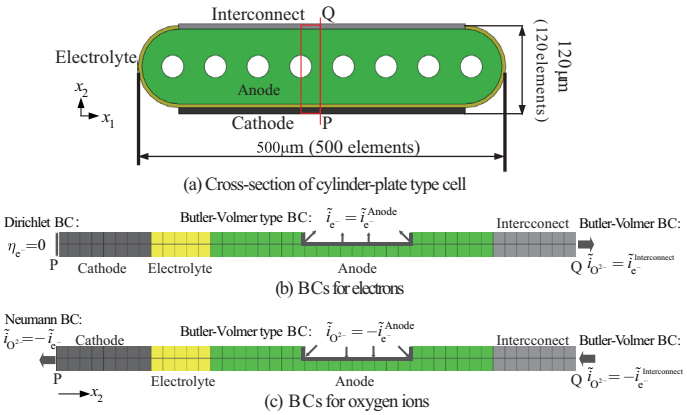
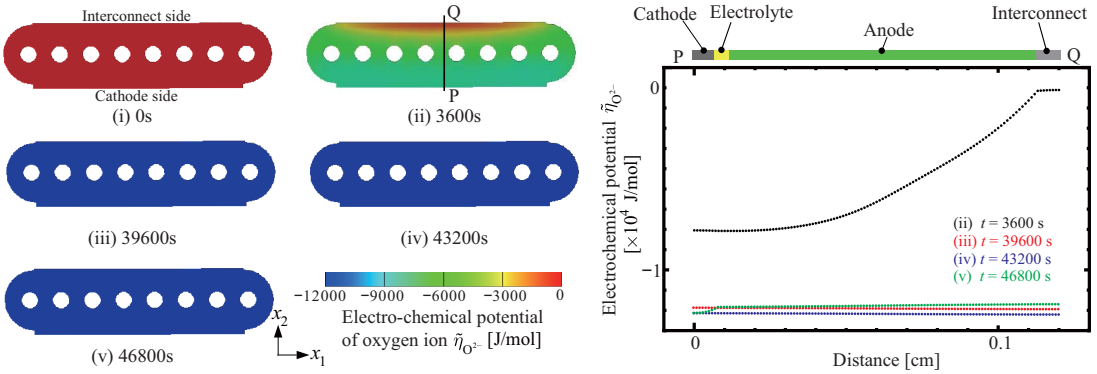


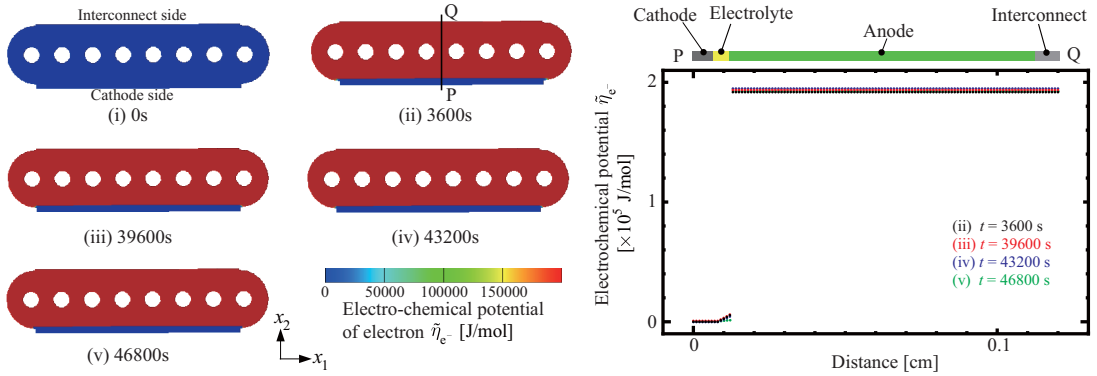
Figure 5: Boundary condition of electric conduction problem



(a) Distributions of electrochemical potential of oxygen ion

(b) Profiles of potential distributions on line P-Q

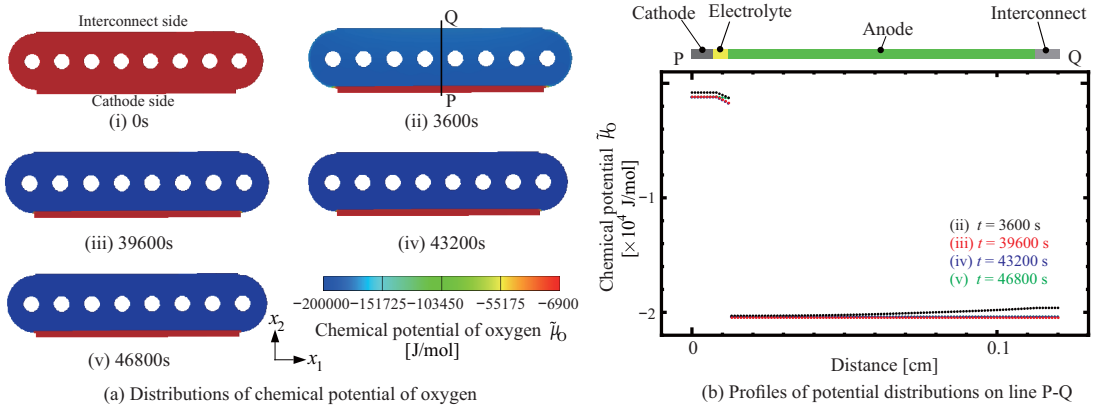
Figure 6: Distribution of electrochemical potential of oxygen ion



(a) Distributions of electrochemical potential of electron

(b) Profiles of potential distributions on line P-Q

Figure 7: Distribution of electrochemical potential of electron



(a) Distributions of chemical potential of oxygen

(b) Profiles of potential distributions on line P-Q

Figure 8: Distribution of chemical potential of oxygen

and cooling, while the overall shape of the cell is warped into the interconnect side (See Figs. 10(ii)-(v)). This is probably caused by the reduction induced expansive strains that are expected to developed in the interconnect side during steady operation and subsequent cooling.

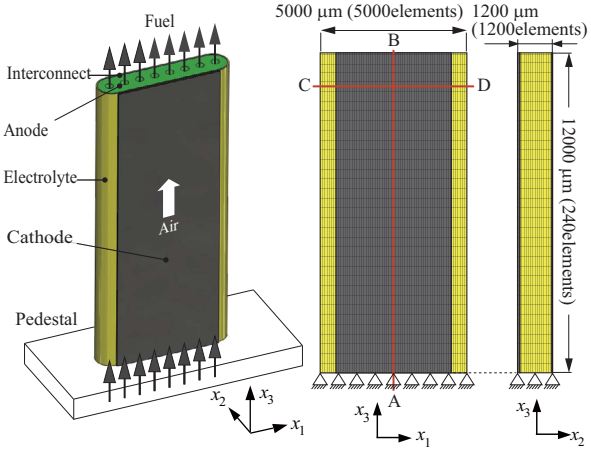


Figure 9: Boundary condition of deformation problem

5 CONCLUSIONS

In this study, to perform stress analyses of SOFCs under operation, mechanical problem and electrochemical problem are coupled taking a time-varying geometries of anode microstructures due to Ni-sintering into account. The conclusions are summarized as follows:

- A phase-field method was presented to simulate the micro-scale morphology change with time, from which the time-variation of the amount of TPBs was directly predicted. In the homogenization analyses, the dependencies of the properties of constituent materials on temperature and/or the oxygen potential, which is supposed to change within an operation period, were also considered.

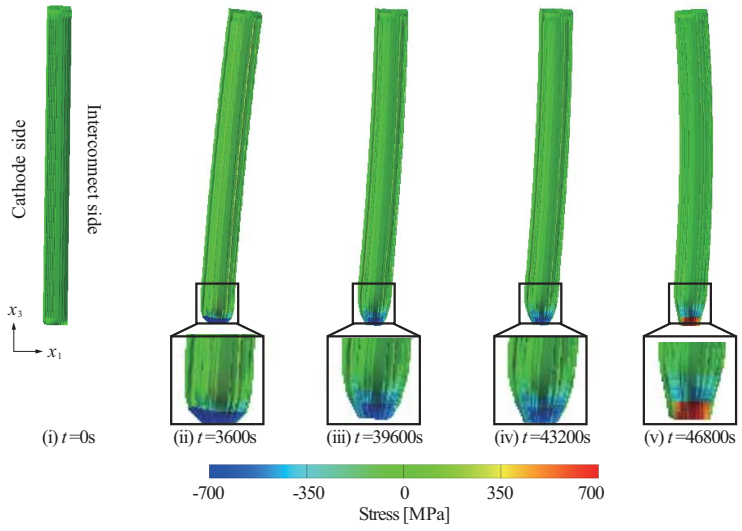


Figure 10: Distribution of stress of the cell

- Under the assumption of an actual start-and-stop operation, the macroscopic electric conduction analysis was conducted to oxygen potential distribution in an overall cell structure under long-period operation with start-and-stop control, which determined reduction-induced expansive/contractive deformation of oxide materials.
- The macroscopic stress analysis was carried out for the overall SOFC structure with the stress-free strains and the homogenized mechanical properties, both of which depend on the operational environment.

ACKNOWLEDGEMENT

This work was carried out as a part of the research project" Technology development for SOFC commercialization promotion, Basic study on rapid evaluation method of SOFC durability", which was supported by the New Energy and Industrial Technology Development Organization (NEDO), Japan.

REFERENCES

Equation Section (Next)

[1] Huang, K. and Goodenough, J.B. *Solid oxide fuel cell technology: principles, performance and operations*. Woodhead Publishing, (2009).
[2] Terada, K., Hirayama, N., Yamamoto, K., Kato, J., Kyoya, T., Matsubara, S., Arakawa, Y., Ueno, Y. and Miyanaga, N. Applicability of micro-macro decoupling scheme to two-scale analysis of fiber-reinforced plastics. *Adv. Compos. Mater.* (2014) **5-6**: 421-450.
[3] Haile, S.M. Fuel cell materials and components. *Acta. Mater.* (2003) **51**:5981-6000.
[4] Chan, S.H., Khor, K.A. and Xia, Z.T. A complete polarization model of a solid oxide fuel cell and its sensitivity to the change of cell component thickness. *J. Power. Sources.* (2001) **93**:130-140.
[5] Terada, K., Kawada, T., Sato, K., Iguchi, F., Yashiro, K., Amezawa, K., Kubo, M., Yugami, H., Hashida, T., Mizusaki, J., Watanabe, H., Sasagawa, T. and Aoyagi, H. Multiscale Simulation of Electro-Chemo-Mechanical Coupling Behavior of PEN Structure under SOFC Operation. *ECS Trans.* (2011) **35**: 923-933.
[6] Jiao, Z. and Shikazono, N. Simulation of solid oxide fuel cell anode microstructure evolution using phase field method. *J. Electrochem. Soc.* (2013) **160**: F709-F715.
[7] Asp, K. and Agren, J. Phase-field simulation of sintering and related phenomena - A vacancy diffusion approach. *Acta. Mater.* (2006) **54**: 1241-1248.
[8] Sasagawa, T., Takahashi, K., Terada, K. and Kawada, T. Estimation of macroscopic material properties using sintering simulation of porous microstructure affected by mechanical effects. *Trans. JSCES* (2012) No. 2012004.
[9] Koyama, M., Ogiya, K., Hattori, T. Fukunaga, H., Suzuki, A., Sahnoun, R., Tsuboi, H., Hatakeyama, N., Endou, A., Takaba, H., Kubo, Del Carpio, C.A. and Miyamoto, A. Development of three-dimensional porous structure simulator POCO2 for simulations of irregular porous materials. *J. Comput. Chem. Jpn.* (2008) **7**: 55-62.

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis (Eds)

CFD SIMULATION OF MICROSCOPIC TWO-PHASE FLUID MOTION ON SOLID BODY WITH EDGES AND HETEROGENEOUSLY-WETTED SURFACE USING PHASE-FIELD MODEL

NAOKI TAKADA*, JUNICHI MATSUMOTO*, SOHEI MATSUMOTO* AND KAZUMA KURIHARA*

* National Institute of Advanced Industrial Science and Technology (AIST)
1-2-1 Namiki, Tsukuba, Ibaraki 305-8564, Japan
e-mail: naoki-takada@aist.go.jp, www.aist.go.jp

Key words: Computational Fluid Dynamics, Multiphase Flow, Wettability, Diffuse-interface Model, Conservative Level-set Method, Lattice Boltzmann Method.

Abstract. Our objective in this study is to examine the applicability of a computational fluid dynamics (CFD) method to simulation of motions of microscopic two-phase fluid on solid surfaces with edges for evaluating fluidic devices and for predicting underground fluid flows through porous media. The method adopts the phase-field model (PFM) for fluid-fluid interfacial dynamics and the lattice-Boltzmann method (LBM) as numerical scheme for solving a set of two-phase fluid-dynamics equations. Based on the free-energy theory, the PFM reproduces an interface as a finite volumetric zone between phases without imposing topological constraints on interface as phase boundary. Wettability of solid surface to fluid is taken into account through minimizing the free energy of the fluid plus the surface energy per unit area. The LBM assumes that a macroscopic fluid consists of fictitious mesoscopic particles repeating collisions with each other and rectilinear translations with an isotropic discrete velocity set. The simply-iterative particle-kinematic operation in a semi-Lagrange form is useful for high-performance computing of complex fluid flow systems. The CFD method therefore has an attractive advantage over the others, efficient simulation of motions of multiple fluid particles on partially-wetted and textured solid surfaces. The major findings in preliminary immiscible liquid-liquid simulations are as follows: (1) the method simulated well departure of single droplet from flat surface due to buoyancy in agreement with the semi-empirical prediction; (2) the droplet had larger static contact angle on hydrophobic grooved surface than that on flat surface in a similar way with available data; (3) the droplet moved more easily under external forcing in the tangential direction to the grooves than that under forcing in the normal direction in qualitative agreement with experimental observations.

1 INTRODUCTION

Microscopic gas-liquid and liquid-liquid flows with a fluid-fluid interface on a heterogeneously or homogeneously wettable solid surface are widely encountered in various science and engineering fields. It is often difficult to experimentally observe such flows and measure the velocity and the pressure simultaneously in three dimensions or to analyse

theoretically them by the classical continuum dynamics approach based on a sharp-interface model. Computational fluid dynamics (CFD) simulations facilitate the understanding and prediction of the two-phase flows for flexible and accurate control of fluid-particle motion and position; as its result, micro-fluidic devices and micro-electro-mechanical-systems (MEMS) device fabrication processes can be optimally designed [1-3].

In last two decades, the phase-field model (PFM) has been attracting much attention from many researchers as one of the mesoscopic models which efficiently simulate behaviours of multi-phase multi-component systems [4]. Based on the free-energy theory [5, 6], the PFM describes an interface as a finite volumetric zone between different phases, across which physical properties vary steeply but continuously. Two-phase coexistence is allowed by a free-energy functional which has a double-well potential of an order parameter (i.e. mass density, molar concentration) and its squared local gradient term, without imposing topological constraints on interface as phase boundary. The contact angle, depending on the energy balance among three types of interfaces, is obtained from a potential of the solid surface through a simple boundary condition of the gradient of the order parameter on the surface [7, 8]. As a result, the PFM-CFD methods do not necessarily require conventional elaborating algorithms for advection and reconstruction of interfaces [4, 7-18]. The method therefore has an advantage over others, efficient simulation of motions of multiple fluid-fluid interfaces attached on solid bodies with edges and partially-wetted textured surfaces. We have recently evaluated two types of diffuse-interface advection equation based on the PFMs [9, 16] to the simulation of two-phase fluid motion on solid surface [17, 18].

Our objective in this study is to examine the applicability of the PFM-CFD method that we have proposed [17, 18] to simulation of microscopic two-phase fluid motion on solid surfaces for evaluating fluidic devices and for predicting underground fluid flows. This paper is organized as follows. In the next section, the basic equations in the PFM-CFD method are outlined for immiscible isothermal liquid-liquid two-phase flow. The third section describes the numerical solution scheme based on the lattice-Boltzmann method (LBM) [8, 9, 19]. In the fourth section, we present preliminary simulation results by use of the method for three-dimensional single droplet attached on flat solid surface in stagnant liquid under gravity, and 3D droplet on textured solid surface. The conclusions are described in the last section.

2 BASIC EQUATIONS IN PHASE-FIELD MODEL (PFM)

The PFM-based CFD methods for simulating incompressible isothermal viscous two-phase fluid flows generally adopt the following set of mass and momentum conservation equations with an interface advection equation [4, 9, 16-18]:

$$\nabla \cdot \mathbf{u} = 0 \tag{1}$$

$$\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u} = \frac{1}{\rho}\left(-\nabla p + \nabla \cdot \boldsymbol{\tau} + \mathbf{F}_I\right) \tag{2}$$

$$\frac{\partial \phi}{\partial t} + \nabla \cdot \left(\phi \mathbf{u} - D(\phi)\nabla \phi\right) = 0 \tag{3}$$

where $\mathbf{u}$ denotes the velocity; t, the time; ρ, the density; p, the pressure; $\boldsymbol{\tau}$, the viscous stress

tensor; $\mathbf{F}_I$, the interfacial-tension force; ϕ, the order parameter which indicates each phase and interfacial regions with its continuous values; and $D(\phi)$, the diffusivity. In the well-known Cahn-Hilliard (CH) form [4-10, 12-15], the diffusion term of Eq. (3) includes the following coefficient:

$$D(\phi) = M_{CH} \frac{\partial \eta(\phi)}{\partial \phi} \tag{4}$$

$$\eta(\phi) = \frac{\partial \psi_0}{\partial \phi} - \kappa_\phi \nabla^2 \phi \tag{5}$$

where M_{CH} denotes the mobility; η, the chemical potential; ψ_0, the double-well potential on ϕ; and κ_ϕ, the capillary coefficient which is related with interfacial width and tension.

In the method we have proposed [17, 18], the conservation-modified Allen-Cahn (AC) form [11, 16] is adopted to Eq. (3). The diffusion coefficient is defined as follows:

$$D(\phi) = M_{AC} \kappa_\phi \left(1 - \frac{\partial \phi(\xi)}{\partial \xi} \left| \nabla \phi \right|^{-1} \right) \tag{6}$$

$$\phi(\xi) = \frac{1}{2} \left\{ 1 + \tanh\left(\frac{\xi}{2\sqrt{\kappa_\phi}} \right) \right\} \tag{7}$$

where M_{AC} denotes the mobility; and ξ, the signed distance in the direction normal to the interface from the central position at $\xi = 0$. The profile of ϕ across an interface in an equilibrium state is theoretically expressed by use of Eq. (7). In the fluid system, two phases correspond to the regions of $\phi = 0$ and 1 respectively, between which the interface is regarded as the finite volumetric region of $0 < \phi < 1$ with width $4\kappa_\phi^{0.5}$.

In each of the CH and AC forms, the diffusion term of Eq. (3) disappears in an equilibrium state. The AC equation, Eq. (3) with Eq. (6), is fully equivalent to a conservative level-set equation [20] in one-step form by which both of interfacial re-initialization and advection calculation are done at the same moment [16]. The same idea for removing the interfacial-curvature dependency on diffusion of ϕ from the original AC equation [11, 16] is applicable to the diffusion coefficient of the CH equation, Eq. (4) [17].

In the same way as the other PFM-CFD methods [7, 8, 10, 15], a wetting boundary condition on solid surface is taken into account in the present method through the following equation [17, 18],

$$\mathbf{n}_S \cdot \kappa_\phi \nabla \phi = -\gamma_S \tag{8}$$

where $\mathbf{n}_S$ is the unit vector normal to the solid boundary and the parameter γ_S is called as wetting potential. The static contact angle θ_W of a fluid phase on the surface depends on the local value of γ_S. The angle can be flexibly set to be heterogeneous on textured solid surfaces by changing the value of γ_S at arbitrary positions on the fluid-solid boundaries.

In addition to Eq. (8), the following constraints are imposed on the stationary solid surface as a non-slip and no-flux boundary [15, 17, 18].

$$\mathbf{n}_S \cdot \nabla p = 0 \tag{9}$$

$$\mathbf{n}_S \cdot \nabla \eta = 0 \tag{10}$$

$$\mathbf{u} = \mathbf{0} \tag{11}$$

3 NUMERICAL METHOD FOR TWO-PHASE FLOW

For solving the above-mentioned set of Eqs. (1), (2) and (3) with Eqs. (6) to (11), we have proposed a numerical scheme [17, 18] based on the lattice Boltzmann method (LBM) [8-10, 19]. The LBM assumes that a macroscopic fluid consists of fictitious mesoscopic particles repeating collisions with each other and linear translations with a set of isotropic discrete velocities. The main variables in the LBM are distribution functions of number density of the particles grouped with their velocities. Time evolutions of the distribution functions, f_a and g_a, at position $\mathbf{x}$ and at time t are expressed by

$$\frac{\partial f_a}{\partial t} + \mathbf{e}_a \cdot \nabla f_a + \frac{\mathbf{F}_I}{\rho} \cdot \frac{\partial f_a}{\partial \mathbf{e}_a} = -\frac{1}{\tau_f}\left[f_a(\mathbf{x},t) - f_a^{eq}(\mathbf{x},t) \right] \tag{12}$$

$$\frac{\partial g_a}{\partial t} + \mathbf{e}_a \cdot \nabla g_a = -\frac{1}{\tau_g}\left[g_a(\mathbf{x},t) - g_a^{eq}(\mathbf{x},t) \right] \tag{13}$$

where $\mathbf{e}_a$ is the particle velocity vector; the subscript a, index of the velocity set; τ_f and τ_g, the relaxation times for f_a and g_a respectively in the BGK approximation collision operator; and the superscript eq denotes a local equilibrium state.

In this study, Equations (12) and (13) are discretized respectively into the semi-Lagrangian forms with second-order accuracies in both space and time as follows [8-10, 19]:

$$f_a(\mathbf{x},t') = f_a(\mathbf{x} - \mathbf{e}_a\Delta t, t) - \frac{\Delta t}{\tau_f}\left[f_a(\mathbf{x} - \mathbf{e}_a\Delta t, t) - f_a^{eq}(\mathbf{x} - \mathbf{e}_a\Delta t, t) \right] \tag{14}$$

$$f_a(\mathbf{x}, t+\Delta t) = f_a(\mathbf{x}, t') + 3 w_a \Delta t \frac{\mathbf{e}_a \cdot \mathbf{F}_I}{c^2} \tag{15}$$

$$g_a(\mathbf{x} + \mathbf{e}_a\Delta t, t+\Delta t) = g_a(\mathbf{x}, t) - \frac{\Delta t}{\tau_g}\left[g_a(\mathbf{x}, t) - g_a^{eq}(\mathbf{x}, t) \right] \tag{16}$$

where Δt is the constant increase in time, $\mathbf{x}+\mathbf{e}_a\Delta t$ denotes the position neighboring $\mathbf{x}$ in the direction of the vector $\mathbf{e}_a$, and w_a is the weight parameter. In three-dimensional Cartesian coordinate system (x, y, z), the space is divided uniformly into unit cubic cells with sides:

$$\Delta x = \Delta y = \Delta z = c\Delta t \tag{17}$$

where c is the lattice constant. The particles are set to be rest or to move according to three-dimensional 15-velocity model [8-10, 19]. The scalar and vector variables of two-phase fluid, ρ, $\mathbf{u}$, p and ϕ, are co-located at each of the spatial cell centers, and defined with f_a and g_a by

$$\rho = \sum_a f_a = \sum_a f_a^{eq} \tag{18}$$

$$\rho\mathbf{u} = \sum_a f_a \mathbf{e}_a = \sum_a f_a^{eq} \mathbf{e}_a \tag{19}$$

$$p = \rho c^2 / 3 \tag{20}$$

$$\phi = \sum_a g_a = \sum_a g_a^{eq} \tag{21}$$

The kinematic viscosity ν and the constant part of the diffusion coefficient Eq. (6), $D_0 = M_{AC}\kappa_\phi$, are defined by the following equations, respectively [8-10, 17-19]:

$$\nu = \frac{c^2}{3}\left(\tau_f - \frac{\Delta t}{2}\right) \tag{22}$$

$$D_0 = \Gamma c^2 \left(\tau_g - \frac{\Delta t}{2}\right) \tag{23}$$

where Γ is the parameter to control the number densities of the moving and rest particles in a stationary equilibrium state of g_a^{eq}.

The boundary conditions for f_a and g_a of the moving particles are taken into account as follows. On non-slip and non-flux solid walls, bounce-back condition is applied to the particles moving with $\mathbf{e}_a$ from fluid side to solid side, which go back to the fluid side in the opposite directions of $-\mathbf{e}_a$. On inflow and outflow boundaries, it is assumed that the particles coming into the computational domain from the outside would be in a local equilibrium state for pressure and fluid velocity given at the outside.

One of main features of the LBM scheme for solving Eqs. (1)-(3) is the simple particle-kinematic operation of iterative local collisions and linear translations in the discrete conservation form on an isotropic spatial lattice. It is useful for high-performance computing on multi-phase fluid dynamics.

4 NUMERICAL RESULTS

4.1 Departure of droplet from solid surface due to buoyancy

For evaluating the PFM-CFD method [17, 18], single droplet in a stagnant continuous liquid phase under gravity was simulated. The computational domain was divided into 64×64×64 cubic cells with sides $\Delta x = \Delta y = \Delta z$ in the coordinate system (x, y, z), which was surrounded with stationary non-slip flat solid wall boundaries. In the initial condition, the droplet took a hemispherical shape with a radius of $16\Delta x$ on the bottom wall with homogeneously-wetted smooth surface at the static contact angle θ_W = 60° or 90° to the continuous phase. The buoyant force acting on the droplet was taken into account inside the dispersed-phase region where $\phi \geq 0.5$ through volumetric forcing at each cell in an upward direction perpendicular to the bottom wall. The two-phase fluid was assumed as an oil-water system under the condition of interfacial tension γ = 51 mN/m, density difference between the phases $\Delta\rho$ of 108.2kg/m^3 and equal viscosities of 1.0 mPa·s at room temperature and at pressure 1atm. The droplet volume V_D was set within the range of 65.0μl to 1.767ml in the actual system.

As shown in Fig. 1, the droplet departs from the wall in cases of (A) at θ_W = 90° and (C) at θ_W = 60° for lager V_D than that in other cases of (B), (D) and (E) at each θ_W. It is also confirmed that the departure takes place more easily at smaller θ_W. The numerical results on

departure and non-departure agreed with the semi-empirical prediction (see Fig.2).

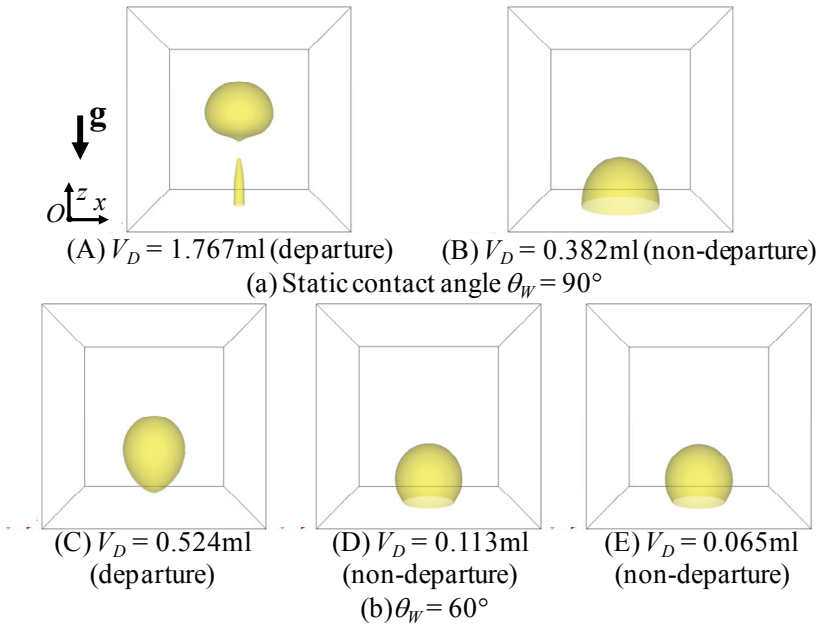


Figure 1: 3D simulation of oil droplet with volume V_D on horizontal flat solid surface with θ_W in stagnant water (continuous phase) under gravity **g**.

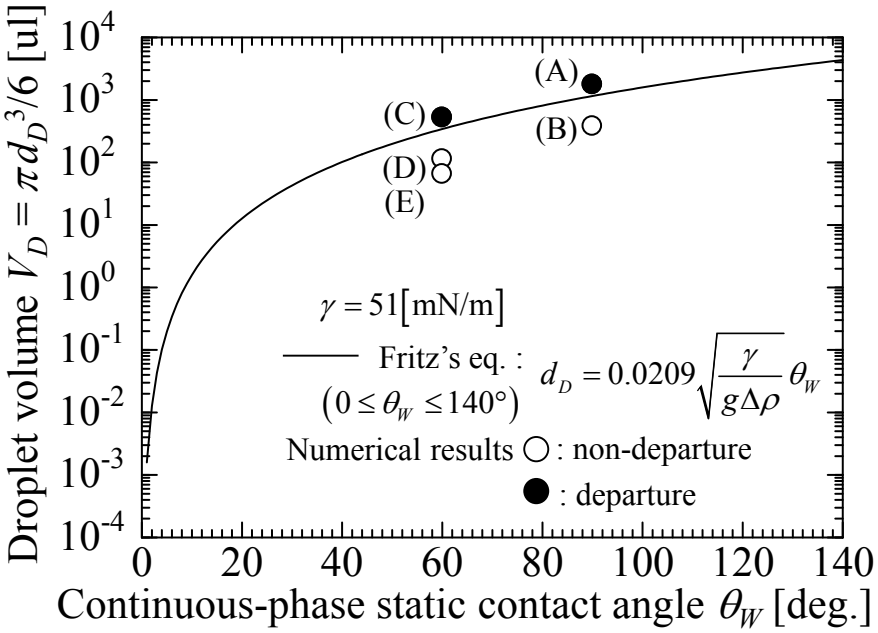
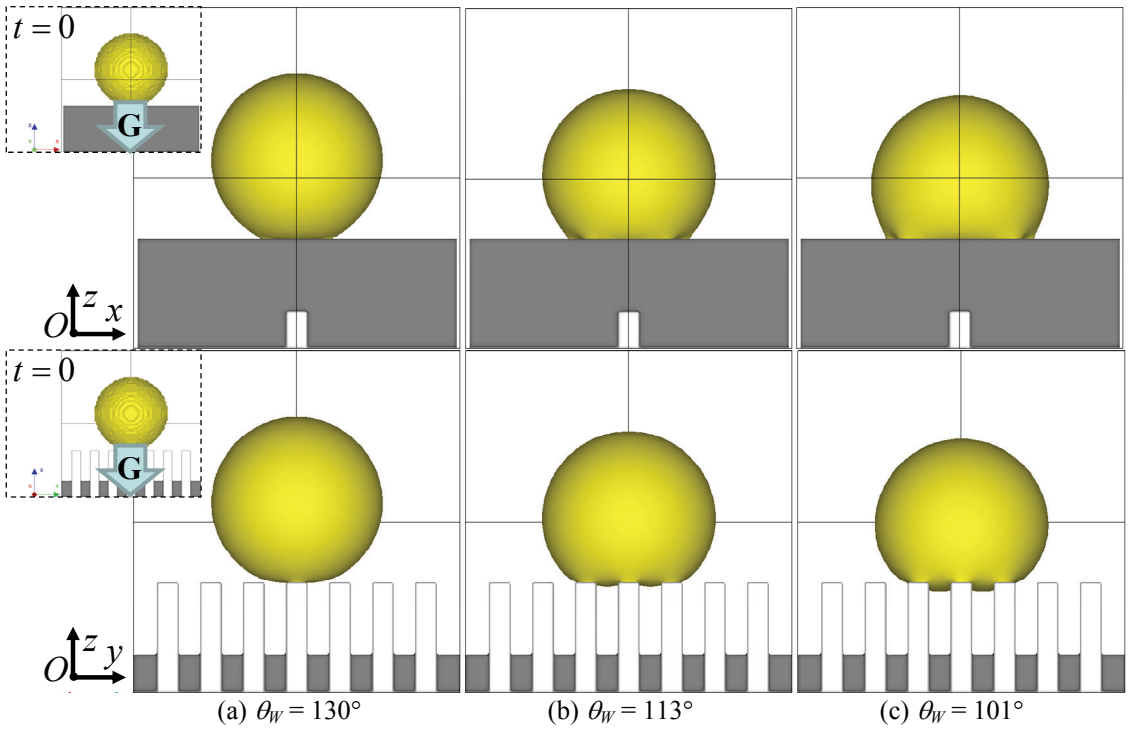


Figure 2: Diagram of departure of droplet from horizontal flat solid surface under gravity g in terms of static contact angle θ_W and volume V_D.

4.2 Droplet on textured solid surface

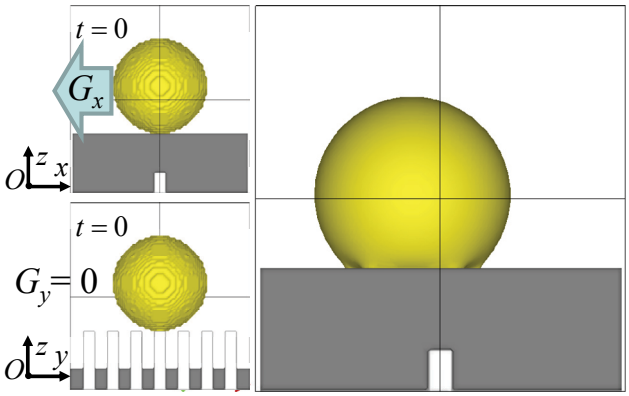
The numerical method was applied to micro-fluidic problem of a droplet attached on a textured solid surface in a stagnant liquid. In the Cartesian coordinate system (x, y, z), the computational domain was uniformly divided into 92×92×96 cubic cells and surrounded with solid wall boundaries on the both sides in the x-axis and z-axis directions and with periodic boundaries in the y-axis direction. Initially a spherical-shaped droplet with diameter d_D = 48Δx was placed at the center of the bottom wall, which was patterned to have rectilinear grooves of width w_F = 6Δx and their intervals of w_S = w_F along the x-axis. The static contact angle θ_W of the droplet on the flat solid surface was set at 130°, 113° or 101° uniformly. After time t = 0, external force **G** had been applied to the droplet. The two-phase fluid was assumed as an immiscible liquid-liquid system with equal density ρ and viscosity μ for dimensionless parameters of $Oh = \mu/(\rho\gamma d_D)^{1/2} = 1.14\times10^{-2}$ and $Eo = |\mathbf{G}|d_D^2/\gamma = 1.76\times10^{-2}$, which were set under the condition of ρ = 998.2kg/m^3, μ = 1mPa·s, γ = 32mN/m and d_D = 240μm.

Figure 3 shows the steady-state shapes of the droplet on the grooved surface under downward **G** in the z direction. At each case of θ_W, the apparent contact angle on the y-z plane became larger than that on the x-z plane in the tangential direction to the grooves. In other cases, as shown in Fig. 4, the droplet moves when **G** has been applied in the x-axis direction, whereas the droplet remained at the initial position on the surface when applying **G** in the y-axis direction. The numerical simulation results in terms of the apparent contact angles and the motion of the droplet agreed qualitatively well with experimental data.

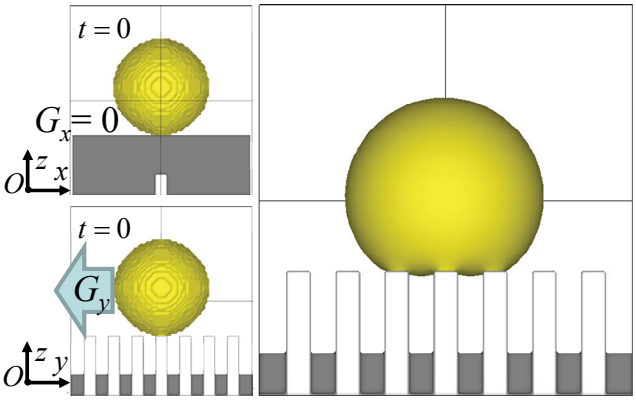


(a) θ_W = 130° (b) θ_W = 113° (c) θ_W = 101°

Figure 3: Steady-state shape of droplet on grooved solid surface under downward external force **G** for static contact angle θ_W and $Oh = \mu/(\rho\gamma d_D)^{1/2} = 1.14\times10^{-2}$ and $Eo = |\mathbf{G}|d_D^2/\gamma = 1.76\times10^{-2}$



(a) Force components $G_x \neq 0$ and $G_y = G_z = 0$



(b) Force components $G_x = 0$, $G_y \neq 0$ and $G_z = 0$

Figure 4: Snapshot of droplet on grooved solid surface under external force $\mathbf{G} = (G_x, G_y, 0)$ for $\theta_W = 113°$, $Oh = 1.14\times10^{-2}$ and $Eo = 1.76\times10^{-2}$

5 CONCLUSIONS

In this study, we examined the applicability of the computational fluid dynamics (CFD) method based on the phase-field model (PFM) [17, 18] to the numerical simulations of motion of microscopic two-phase fluid on heterogeneously-wetted and textured solid surfaces. The PFM-CFD method, that adopted the lattice Boltzmann method (LBM) [8-10, 19] for solving the Navier-Stokes equations with the conservation-modified Allen-Cahn equation [16, 20], was applied to preliminary simulations of an immiscible liquid-liquid two-phase fluid on flat and textured solid surfaces.

The following major findings were obtained:

- The PFM-CFD method simulated well departure of droplet from flat solid surface due to buoyancy in agreement with semi-empirical prediction.
- The droplet on a rectilinearly-grooved hydrophobic solid surface had larger static contact angle than that on a flat surface. The droplet on the textured surface moved more easily in the tangential direction to the grooves than in the transverse direction to them. These numerical results agreed qualitatively with experimental data.

The above-mentioned results prove that the PFM-CFD method can be useful for numerically analyzing the microscopic two-phase fluid motions on solid surfaces for optimal design and evaluation of micro-fluidic devices and micro-fabrication processes, and also for understanding and prediction of underground water flowing with heat and mass transfer.

ACKNOWLEDGEMENTS

This research was granted by the Japan Society for the Promotion of Science (JSPS) through the "Funding Program for World-Leading Innovative R&D on Science and Technology (FIRST Program)," initiated by the Council for Science and Technology Policy (CSTP), and also supported by JSPS KAKENHI Grant Number 26420128.

REFERENCES

[1] Kurihara, K., Suzuki, Y., Sudo, K., Shiba, N., Nakano, T. and Tominaga, J. Wettability control using large-area nanostructured film. *Microelectronic Engineering* (2009) **87**: 1424–1427.

[2] Matsumoto, S., Takada, N. and Matsumoto, J. Microfabrication process of cellular structures in hollow fiber-shaped substrates. *Transactions of the Japan Society of Mechanical Engineers Series C*. (2010) **76**: 1911–1913.

[3] Lu, J., Takagi, H., Nakano, T. and Maeda, R. Size-free MEMS-IC high-efficient integration by using carrier wafer with self-assembled monolayer fine pattern. *Proceedings of ECTC2013*. (2013) CFP13ECT-USB: 1508–1513.

[4] Anderson, D. M., McFadden, G. B. and Wheeler, A. A. Diffuse-interface methods in fluid mechanics. *Annu. Rev. Fluid Mech.* (1998) **30**: 139–165.

[5] van der Waals, J. E. The thermodynamic theory of capillarity under the hypothesis of a continuous density variation. Transl. Rowlinson, J. S., *J. Stat. Phys*. (1979): **20**, 197–244.

[6] Cahn, J. W. and Hilliard, J. E. Free energy of a nonuniform system. I. Interfacial free energy. *J. Chem. Phys*. (1958) **28**: 258–267.

[7] Seppecher, P. Moving contact lines in the Cahn-Hilliard theory. *Int. J. Eng. Sci.* (1996) **34**: 977–992.

[8] Briant, A. J., Papatzacos, P., and Yeomans, J. M. Lattice Boltzmann simulations of contact line motion in a liquid-gas system. *Phil. Trans. R. Soc. Lond. A*. (2002) **360**: 485–495.

[9] Inamuro, T., Ogata, T., Tajima, S. and Konishi, N. A lattice Boltzmann method for incompressible two-phase flows with large density differences. *J. Comput. Phys.* (2004) **198**: 628–644.

[10] Yoshino, M. and Mizutani, Y. Lattice Boltzmann simulation of liquid-gas flows through solid bodies in a square duct. *Math. Comput. Simul.* (2006) **72**: 264–269.

[11] Beaucourt, J., Biben, T., Leyat, A. and Verdier, C. Modeling breakup and relaxation of Newtonian droplets using the advected phase-field approach. *Phys. Rev. E.* (2007) **75**: 021405.

[12] Kim, J. Phase field computations for ternary fluid flows. *Comput. Meth. Appl. Mech. Eng.* (2007) **196**: 4779–4788.

[13] Takada, N. and Tomiyama, A. A numerical method for two-phase flow based on a phase-field model. *JSME Int. J. Ser. B Fluids Therm. Eng.* (2006) **49**: 636–644.

[14] Matsumoto, J. and Takada, N. Two-phase flow analysis based on a phase-field model using orthogonal basis bubble function finite element method. *Int. J. Comput. Fluid Dyn.* (2008) **22**: 555–568.
[15] Takada, N., Matsumoto, J., Matsumoto, S. and Ichikawa, N. Application of a phase-field method to the numerical analysis of motions of a two-phase fluid with high density ratio on a solid surface. *Journal of Computational Science and Technology* (2008) **2**: 318–329.
[16] Chiu, P.-H. and Lin, Y.-T. A conservative phase field method for solving incompressible two-phase flows. *J. Comput. Phys.* (2011) **230**: 185–204.
[17] Takada, N., Matsumoto, J. and Matsumoto, S. Phase-field model-based simulation of motions of a two-phase fluid on solid surface. *Journal of Computational Science and Technology* (2013) **7**: 322–337.
[18] Takada, N., Matsumoto, J. and Matsumoto, S. A diffuse-interface tracking method for the numerical simulation of motions of a two-phase fluid on a solid surface. *The Journal of Computational Multiphase Flows* (2014) **6**: 283–298.
[19] Chen, S. and Doolen, G. D. Lattice Boltzmann method for fluid flows. *Annu. Rev. Fluid Mech.* (1998) **30**: 329–364.
[20] Olsson, E. and Kreiss, G. A conservative level set method for two phase flow. *J. Comput. Phys.* (2005) **210**: 225–246.

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis(Eds)

3D NUMERICAL APPROXIMATION OF RELATIVISTIC PARTICLE BEAMS BY ASYMPTOTIC EXPANSION

FRANCK ASSOUS* AND ANNE LE BLANC*,†

*Department of Comput. Sc. & Maths
Ariel University
40700 Ariel, Israel
e-mail: franckassous@netscape.net

†Department of Mathematics
Tel Aviv University
Tel Aviv, Israel
e-mail: anneleb@post.tau.ac.il

Key words: Vlasov-Maxwell equations, asymptotic methods, paraxial approximation, numerical schemes, relativistic beams

Abstract. We consider the numerical approximation and simulation of a particle beam, usually modeled by the Vlasov-Maxwell system. Our work will deal with a 3D paraxial approximate model, derived from asymptotic expansions. It includes a finite element numerical implementation of the paraxial Maxwell model coupled with a Particle-In-Cell method for the corresponding paraxial Vlasov model. Both methods are implemented with Freefem++ software. Numerical results illustrated the efficiency of the method.

1 INTRODUCTION

Charged particle beams and plasma physics problems are extensively used in Science and Technology. Although we often associate accelerators with the large machines of high-energy physics, charged particle beams have continually expanding applications in many branches of research and technology. Recent active areas include flat-screen cathode-ray tubes, synchrotron light sources, beam lithography for microcircuits, thin-film technology, production of short-lived medical isotopes, radiation processing of food, and free-electron lasers. Clearly, there exists a significant interest in building mathematical models for these beams.

If we consider collisionless plasma or non-collisional beams, one of the most complete mathematical models is the time-dependent Vlasov-Maxwell system of equations. However, the numerical solution of such models requires a large computational effort. There-

fore, whenever possible, we have to take into account the particularities of the physical problem to derive asymptotic approximate models leading to cheaper simulations.

In this article, we consider the case of high energy short beams. A typical example is the transport of a bunch of highly relativistic charged particles in the interior of a perfectly conducting hollow tube. Numerical simulations are mostly performed using the particle-in-cell method.

Following [1–5], we introduce a paraxial model that approximate the coupled time-dependent Vlasov-Maxwell equations. This model is derived by introducing a frame which moves along the optical axis at the speed of light. Then, considering a scaling of the equations which reflects the characteristics of the high energy short beam, a small parameter η is introduced, and asymptotic expansion techniques are used to derive a paraxial model, accurate up to fourth order in η.

This model is then approximated by a finite element method, for the paraxial Maxwell model, coupled with a Particle-In-Cell method for the corresponding paraxial Vlasov model. This implementation is based on the Freefem++ software [7]. First numerical results are proposed, and show the efficiency of the method.

2 THE 3D GOVERNING EQUATIONS

2.1 The Vlasov-Maxwell model

We consider the transport of a population of highly relativistic charged particles, with a mass m and a charge q, in the interior of a perfectly conducting hollow tube, whose axis is constituted by the z-axis. We denote by Ω the transverse section of boundary Γ. Let $\mathbf{x}$=(x,y,z) denote the position of a particle, $\mathbf{p}$=(p_x,p_y,p_z) its momentum and $\mathbf{v}$=(v_x,v_y,v_z) its velocity. We assume that the beam is non collisional so that its distribution function f=f($\mathbf{x}$,$\mathbf{p}$,t) in the phase space ($\mathbf{x}$,$\mathbf{p}$) is a solution to the Vlasov equation

$$\frac{\partial f}{\partial t} + \mathbf{v}.\nabla_{\mathbf{x}} f + \mathbf{F}.\nabla_{\mathbf{p}} f = 0, \qquad \text{where } \mathbf{p} = \gamma m \mathbf{v}, \qquad \gamma = (1 - \frac{|\ \mathbf{v}\ |^2}{c^2})^{-1/2}. \tag{1}$$

Above, $\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B})$ denotes the electromagnetic force acting on the particles. The electric field $\mathbf{E}$=$\mathbf{E}$($\mathbf{x}$,t) and the magnetic field $\mathbf{B}$=$\mathbf{B}$($\mathbf{x}$,t) are solutions to Maxwell's equations

$$\begin{cases} \dfrac{\partial \mathbf{E}}{\partial t} - c^2 \nabla \times \mathbf{B} = -\dfrac{1}{\varepsilon_0}\mathbf{J}, & \dfrac{\partial \mathbf{B}}{\partial t} + \nabla \times \mathbf{E} = 0, \\ \nabla.\mathbf{E} = \dfrac{\rho}{\varepsilon_0}, & \nabla.\mathbf{B} = 0, \end{cases} \tag{2}$$

where the charge and the current density ρ and $\mathbf{J}$ are obtained from the distribution function f with

$$\rho = q \int f d\mathbf{p}, \qquad \mathbf{J} = q \int \mathbf{v} f d\mathbf{p}. \tag{3}$$

Assuming the tube being perfectly conducting is equivalent to assuming that $\mathbf{E}$ (respectively $\mathbf{B}$) have a vanishing tangential (respectively normal) trace on the tube boundary.

2.2 The paraxial model

Assuming that the beam is highly relativistic corresponds to assume that $\gamma >> 1$. Since $v_z \simeq c$ for any particle in the beam, the Vlasov-Maxwell equations (1-2) can be written in a frame which moves along z-axis with the light velocity c. For this purpose, we set $\zeta = ct - z$, $v_\zeta = c - v_z$ and we perform the change of variables $(x, y, z, v_x, v_y, v_z, t) \rightarrow (x, y, \zeta, v_x, v_y, v_\zeta, t)$. It is also convenient to introduce the transverse quantities

$$\mathbf{x}_\perp = (x, y),\ \mathbf{v}_\perp = (v_x, v_y)$$

and to define the transverse operators

$$\mathbf{grad}_\perp \varphi = (\frac{\partial \varphi}{\partial x}, \frac{\partial \varphi}{\partial y}), \quad \mathbf{curl}_\perp \varphi = (\frac{\partial \varphi}{\partial y}, -\frac{\partial \varphi}{\partial x}), \quad \Delta_\perp \varphi = \frac{\partial^2 \varphi}{\partial x^2} + \frac{\partial^2 \varphi}{\partial y^2},$$

for $\varphi = \varphi(x, y)$ a given scalar function. Similarly, for $\mathbf{A}_\perp = (A_x, A_y)$ denoting a transverse vector field, we set

$$div_\perp \mathbf{A}_\perp = \frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y}, \quad curl_\perp \mathbf{A}_\perp = \frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y}.$$

With these notations, the Vlasov-Maxwell equations (1-2) can be expressed in the new variables. The second step to derive the paraxial model consists in introducing characteristic quantities and rewrite the system of equations in dimensionless variables. Introducing a small parameter, the paraxial model was derived by retaining the terms up to the third order in the asymptotic expansion of the distribution function f. This third order expansion of f is entirely determined from the expansion of the transverse electromagnetic force $\mathbf{F}_\perp$ up to order 2, and from the expansion of the longitudinal electromagnetic force F_z up to order 1 only.

As a consequence, it was proved that the asymptotic paraxial Vlasov- Maxwell model requires the knowledge of the principal parts (zero order) of the transverse electric field $\mathbf{E}_\perp$, the first order part of the longitudinal electromagnetic field (E_z, B_z), and the second order part of the transverse so-called "pseudo-fields" $\mathcal{E}_\perp$, where $\mathcal{E}_\perp = (\mathcal{E}_x, \mathcal{E}_y)$ is defined by $\mathcal{E}_x = E_x - cB_y, \mathcal{E}_y = E_y + cB_x$. Details may be found in [2], [3].

Based on these remarks, we have to approximate by numerical methods the paraxial Vlasov equation, and the electromagnetic components $\mathbf{E}_\perp, E_z, B_z, \mathcal{E}_\perp$, solutions to the following paraxial Maxwell equations:

$$\left\{\begin{array}{l} curl\,_\perp \mathbf{E}_\perp = 0 \ \text{ in } \Omega\,, \\ div\,_\perp \mathbf{E}_\perp = \dfrac{1}{\varepsilon_0}\rho \ \text{ in } \Omega\,, \\ \mathbf{E}_\perp \cdot \boldsymbol{\tau} = 0 \ \text{ on } \Gamma\,, \end{array}\right. \qquad \mathbf{B}_\perp = -\frac{1}{c}\mathbf{E}_\perp \times \mathbf{e}_z\,, \qquad \left\{\begin{array}{l} \mathbf{curl}\,_\perp E_z = -\dfrac{\partial \mathbf{B}_\perp}{\partial t} \ \text{ in } \Omega\,, \\ E_z = 0 \ \text{ on } \Gamma\,, \end{array}\right.$$

$$\left\{\begin{array}{l} \mathbf{curl}\,_\perp B_z = \mu_0\,\mathbf{J}_\perp + \dfrac{1}{c^2}\dfrac{\partial \mathbf{E}_\perp}{\partial t} \ \text{ in } \Omega\,, \\ \displaystyle\int_\Omega B_z\, d\Omega = 0 \ \text{ on } \Gamma\,, \end{array}\right. \qquad \left\{\begin{array}{l} curl\,_\perp \mathcal{E}_\perp = -\dfrac{\partial B_z}{\partial t} \ \text{ in } \Omega\,, \\ div\,_\perp \mathcal{E}_\perp = \mu_0\, c\, J_\zeta - \dfrac{1}{c}\dfrac{\partial E_z}{\partial t} \ \text{ in } \Omega\,, \\ \mathcal{E}_\perp \cdot \boldsymbol{\tau} = 0 \ \text{ on } \Gamma\,, \end{array}\right. \tag{4}$$

where J_ζ is defined by $J_\zeta = \rho c - J_z = q \displaystyle\int v_\zeta f d\mathbf{V}$. In this model, the expression of the Lorentz force $\mathbf{F} = (\mathbf{F}_\perp, F_z)$ is given by

$$\left\{\begin{array}{l} \mathbf{F}_\perp = q(\mathcal{E}_\perp + (\mathbf{v}_\perp \times \mathbf{e}_z)B_z + v_\zeta(\mathbf{B}_\perp \times \mathbf{e}_z)\,, \\ F_z = q(E_z + \mathbf{v}_\perp \cdot (\mathbf{B}_\perp \times \mathbf{e}_z))\,. \end{array}\right. \tag{5}$$

3 NUMERICAL METHODS

Our aim is now to build numerical methods to solve the problem (4-5). We have chosen to derive a finite element approximation for the electromagnetic fields computations. The Vlasov equation will be solved by a particle method.

3.1 Numerical schemes for the electromagnetic fields

The first step to get the numerical schemes consists in deriving variational formulations of equations (4). Since the model is written in a frame which moves along the optical axis at the speed of light, the bunch of particles is evolving slowly in that frame. As a consequence, the 3D computational domain is defined as the product $\Omega \times]0, Z[$, $0 \leq \zeta \leq Z$.

Let us now introduce the variational formulations which will be the basis of the method. For the sake of simplicity, we will only consider the components $\mathbf{E}_\perp$ and E_z, the other components solving more or less similar equations. Moreover, as the regularity of the fields are not an issue for our study, we will assume that they are smooth enough, for instance belonging to a standard Sobolev space. For the sake of simplicity, we will denote by $\mathbf{V}$ the space of the fields and of the test functions, regardless of the boundary conditions they satisfy.

Let $\mathbf{v}_\perp$ denote a sufficiently smooth vector test function. We first apply the $\mathbf{curl}\,_\perp$ operator to the first equation of the system. Then, we take the dot product by $\mathbf{v}_\perp$, and integrate

over Ω. Applying then Green's formula for the $\mathbf{curl}_\perp$, we obtain the following variational formulation

$$\int_\Omega curl_\perp \mathbf{E}_\perp curl\, \mathbf{v}_\perp d\Omega - \int_\Gamma (\mathbf{v}_\perp \cdot \boldsymbol{\tau}) curl_\perp \mathbf{E}_\perp d\Gamma = 0$$

The divergence equation is handled through an augmented Lagrangian formulation. Multiplying it by $\mathrm{div}_\perp \mathbf{v}_\perp$, integrating over Ω, and adding it to the above equation, we get

$$\int_\Omega \{curl_\perp \mathbf{E}_\perp curl\, \mathbf{v}_\perp + \mathrm{div}_\perp \mathbf{E}_\perp \mathrm{div}_\perp \mathbf{v}_\perp\} d\Omega - \int_\Gamma (\mathbf{v}_\perp \cdot \boldsymbol{\tau}) curl_\perp \mathbf{E}_\perp d\Gamma = \frac{1}{\varepsilon_0} \int_\Omega \rho\, \mathrm{div}_\perp \mathbf{v}_\perp d\Omega$$

To handle the boundary condition on the $curl - div$ system, we will use a Nitsche method, as proposed in [6]. This is performed on the discretization level. Consider a regular finite element mesh T_h (where $T_h = \cup K$) of the domain, and a finite element approximation space $\mathbf{V}_h = \{\mathbf{v}_\perp \in \mathbf{V} \mid \mathbf{v}_{\perp_{|K}} \in \mathbb{P}_k(K)\}$, where $\mathbb{P}_k(K)$ denotes the set of all vector fields which are polynomials componentwise on K with degree $\leqslant k$. Let us denote by $\mathbf{E}_h^\perp$ the approximate solution of $\mathbf{E}_\perp$ in $\mathbf{V}_h$, C_h being the trace mesh induced by T_h on the boundary of the domain. Essentially, Nitsche's method imposes the boundary condition via three boundary terms. Two of them contain the weak form of the tangential trace of the solution and the test functions. These two terms cause the method to be symmetric and consistent. The third term (with a parameter β) depends on the domain tetrahedrization, and causes the method to be stable. In our case, the Nitsche method is written

$$\begin{aligned}\int_\Omega \{curl_\perp \mathbf{E}_h^\perp curl\, \mathbf{v}^\perp + \mathrm{div}_\perp \mathbf{E}_h^\perp \mathrm{div}_\perp \mathbf{v}_\perp\} d\Omega - \int_\Gamma (\mathbf{v}_\perp \cdot \boldsymbol{\tau}) curl_\perp \mathbf{E}_h^\perp d\Gamma \\ - \int_\Gamma (\mathbf{E}_h^\perp \cdot \boldsymbol{\tau}) curl_\perp \mathbf{v}_\perp d\Gamma + \beta \sum_{E \in C_h} \frac{1}{h} \int_E (\mathbf{E}_h^\perp \cdot \boldsymbol{\tau})(\mathbf{v}_\perp \cdot \boldsymbol{\tau}) d\Gamma = \frac{1}{\varepsilon_0} \int_\Omega \rho\, \mathrm{div}_\perp \mathbf{v}_\perp d\Omega\,,\end{aligned}$$

where β is some positive sufficiently large constant.

Concerning the computation of E_z, we have first to derive a suitable variational formulation. Basically, we apply the $curl_\perp$ operator to the third equation of (4) and we use the identity $curl_\perp \mathbf{curl}_\perp = -\Delta_\perp$ to get

$$\Delta_\perp E_z = curl_\perp \frac{\partial \mathbf{B}_\perp}{\partial t}$$

As previously, we take the dot product by v_z, a sufficiently smooth scalar test function, and integrate over Ω. We then apply the classical Green's formula for the Laplace operator, and use that E_z vanishes on the boundary Γ. The variational formulation, basis of our finite element method, is finally written

$$\int_\Omega \mathbf{grad}_\perp E_z \cdot \mathbf{grad}_\perp v_z d\Omega = \int_\Omega \partial_t \left(\frac{\partial B_y}{\partial x} - \frac{\partial B_x}{\partial y}\right) v_z d\Omega\,. \tag{6}$$

Similar formulations are obtained for the components B_z and $\mathcal{E}_\perp$, following the same principles. As a result, one derives the finite element conforming P_1 approximations by using the FreeFem++ package [7]. The time discretization is performed with a classical finite difference scheme. Remark that time discretization is not an issue here, since it only appears in the right-hand sides of the formulations. Hence, there is no necessity to satisfy any stability condition.

3.2 Particle approximation of paraxial Vlasov equation

The paraxial Vlasov equation (1), written in a paraxial form, that is for an electromagnetic force $\mathbf{F}$ deduced from the paraxial model, is numerically solved by means of a particle method [8]. One approximates the function $f(\mathbf{x}, \mathbf{p}, t)$ (where $\mathbf{x} = (\mathbf{x}_\perp, \zeta)$) by a linear combination of delta distributions in the phase space $(\mathbf{x}, \mathbf{p})$, namely:

$$f(\mathbf{x}, \mathbf{p}, t) = \sum_k w_k \, \delta(\mathbf{x} - \mathbf{x}_k(t)) \delta(\mathbf{p} - \mathbf{p}_k(t)) \,, \tag{7}$$

where w_k denotes the constant weight of the particle k. Its position in the phase space $\mathbf{x}_k = (x, y, \zeta)$ and $\mathbf{p}_k = (p_x, p_y, p_z)$ is solution to the differential system:

$$\left\{ \begin{array}{lll} \dfrac{dx}{dt} = \dfrac{p_x}{\gamma m}, & & \dfrac{dp_x}{dt} = F_x, \\ \dfrac{dy}{dt} = \dfrac{p_y}{\gamma m}, & & \dfrac{dp_y}{dt} = F_y, \\ \dfrac{d\zeta}{dt} = c - \dfrac{p_z}{\gamma m}, & & \dfrac{dp_z}{dt} = F_z, \end{array} \right. \tag{8}$$

together with initial conditions.

The corresponding particle charge and current densities ρ and $\mathbf{J}$ are obtained by introducing the particle approximation (7) in equations (3) that yields:

$$\rho(\mathbf{x}, t) = q \sum_k w_k \delta(\mathbf{x} - \mathbf{x}_k(t)) \,, \tag{9}$$

and

$$\mathbf{J}(\mathbf{x}, t) = q \sum_k w_k \mathbf{v}_k(t) \delta(\mathbf{x} - \mathbf{x}_k(t)) \,. \tag{10}$$

Such expressions, built at the particle positions, cannot be used in this form for solving paraxial Maxwell equations. Indeed, a P_1 finite element approximation requires values of ρ and $\mathbf{J}$ at the vertices of the tetrahedral mesh. Following the classical procedure [8, 9], we introduce the assignment and interpolation procedures.

According to the general approach, time discretization of system (8) is built from a leapfrog scheme, which is a second-order centered finite-difference scheme. The particle positions are defined at time t_n and the particle momenta are computed at time $t_{n+1/2}$. The equations of momentum $\mathbf{p}_k$ are approximated by

$$\begin{cases} \dfrac{1}{\Delta t}(p_x^{n+\frac{1}{2}} - p_x^{n-\frac{1}{2}}) = F_x^n \,, \\ \dfrac{1}{\Delta t}(p_y^{n+\frac{1}{2}} - p_y^{n-\frac{1}{2}}) = F_y^n \,, \\ \dfrac{1}{\Delta t}(p_z^{n+\frac{1}{2}} - p_z^{n-\frac{1}{2}}) = F_z^n \,, \end{cases} \tag{11}$$

where (F_x^n, F_y^n, F_z^n) is a numerical approximation of the Lorentz force (5) at time t_n, the computation of which requiring the knowledge of the paraxial electromagnetic fields. Since they are determined by finite element methods, an interpolation procedure is necessary to recover the values of the fields at the particle locations. For this purpose, we use an interpolation procedure, similar to the one proposed in [3], where the fields are computed by a finite difference method.

The last step consists in computing the particle position solution to (8), that are obtained by solving the following discretized system

$$\begin{cases} \dfrac{x^{n+1} - x^n}{\Delta t} = \dfrac{p_x^{n+\frac{1}{2}}}{\gamma^{n+\frac{1}{2}} m} \,, \\ \dfrac{y^{n+1} - y^n}{\Delta t} = \dfrac{p_y^{n+\frac{1}{2}}}{\gamma^{n+\frac{1}{2}} m} \,, \\ \dfrac{\zeta^{n+1} - \zeta^n}{\Delta t} = c - \dfrac{p_z^{n+\frac{1}{2}}}{\gamma^{n+\frac{1}{2}} m} \,, \end{cases} \tag{12}$$

where $\gamma^{n+\frac{1}{2}}$ is computed with $\gamma^{n+\frac{1}{2}} = \left(1 + \dfrac{|\mathbf{p}^{n+\frac{1}{2}}|^2}{(mc)^2}\right)^{\frac{1}{2}}$. The final complete time advance algorithm has the same structure as the one described in [3], where two dimensional paraxial Maxwell equations were approached by a finite-difference method. We refer the interested reader to this reference.

4 NUMERICAL RESULTS

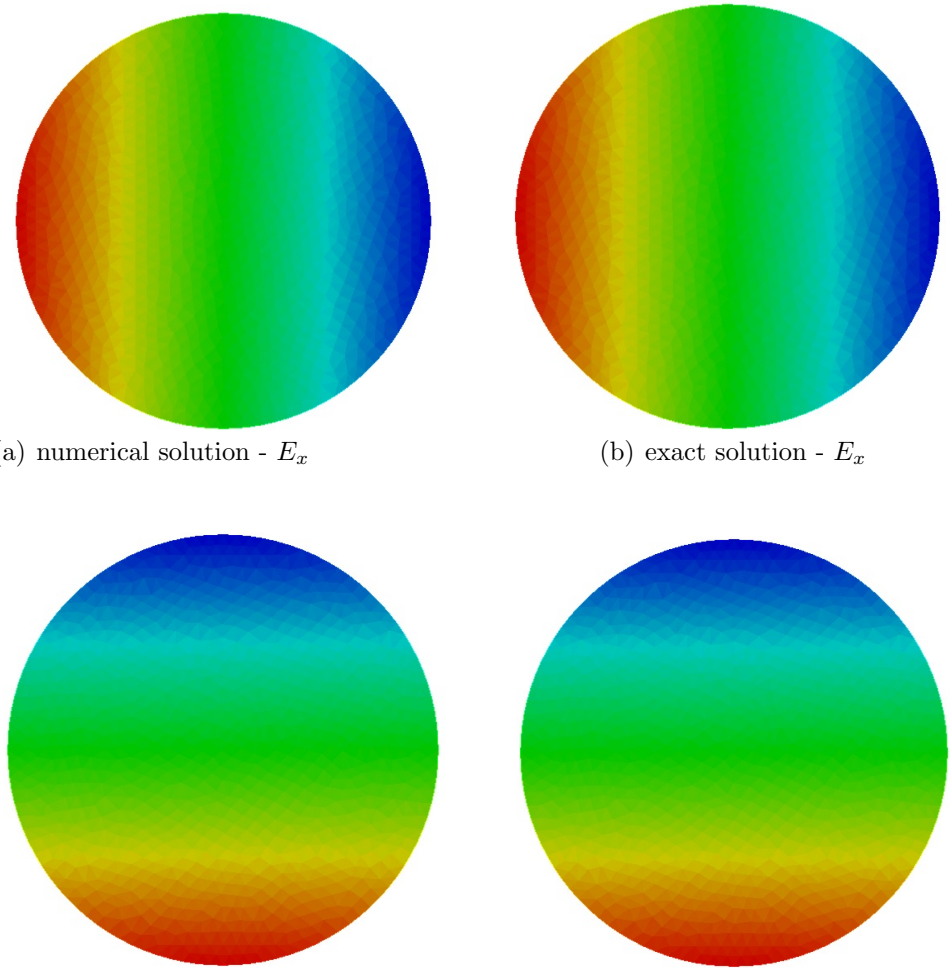
Our aim is to demonstrate the accuracy and the validity of the numerical method, derived from the discrete variational formulations. To this purpose, we consider a 3D

computational domain consisting of a cylinder, the axis of which being the ζ axis. The transverse section Ω is made of a disk of radius $R = 0.1$. We choose a mesh made of 100 edges on the base on the cylinder and 30 layers up, for a cylinder of length 0.3. The total number of degrees of freedom is equal to 86862. We also choose a time step $\Delta t = 10^{-4}$s. All runs were performed on a commercial laptop (MacBook Pro, Processor 2.6 GHz Intel Core i5, Memory 8 GB 1600 MHz DDR3).

We first derive an analytic solution $\mathbf{E}_\perp$ for a given charge density $\rho(\mathbf{x}, t)$. Choosing $\rho(\mathbf{x}, t) = 4\varepsilon_0(\cos(x^2 + y^2) - (x^2 + y^2)\sin(x^2 + y^2))\cos t$, one easily finds that the electric field $\mathbf{E}_\perp = (2x\ \cos(x^2+y^2)\cos t, 2y\ \cos(x^2+y^2)\cos t)$ solves the first equation of (4). With these definitions, we can numerically compute the quantities related to the paraxial model and compare the computed solutions to this exact one. Figure 1 shows respectively the x and y components obtained after 100 time steps of simulation. In order to make the visual comparison between the computed and exact solution more convenient, we have chosen to display the solution in a cut plane of the mesh ($\zeta = 0.15$). As one can see there is an good agreement between the computed solution and the exact one, depicted at the same scale. Indeed, there is no way of distinguishing the difference between the two cut plans of the solutions, even if we used a rather coarse mesh.

5 CONCLUSION

In this paper, we proposed a numerical approximation of a paraxial Vlasov-Maxwell model in three dimensions, adapted to highly relativistic beam. The system of equations we got is simpler and easier to solve than the complete 3D Vlasov-Maxwell equations. We have derived a finite element numerical implementation of the Maxwell part of the model, coupled with a Particle-In-Cell method for the Vlasov part. Both methods were implemented by using the Freefem++ software. This approach seems powerful in its ability to get an accurate, but fast and easy to implement algorithm. Numerical results have been presented to illustrate the feasibility of the method. This solver should give an interesting numerical tool for simulating high energy short beams problems, and could be valuable to the computational accelerator physics community.



(a) numerical solution - E_x (b) exact solution - E_x

(c) numerical solution - E_y (d) exact solution - E_y

Figure 1: $\mathbf{E}_\perp(\mathbf{x}_\perp, \zeta = -0.15)$ after 100 time steps

REFERENCES

[1] P. Degond, P.A. Raviart, On the paraxial approximation of the stationary Vlasov-Maxwell system, *Math. Meth. Appl. Sci.*, **3**, 513–562 (1993).

[2] G. Laval, S. Mas-Gallic, P.-A. Raviart, Paraxial approximation of ultrarelativistic intense beams, *Numer. Math.*, **69(1)**, 33–60 (1994).

[3] F. Assous, F. Tsipis, Numerical paraxial approximation for highly relativistic beams, *Comput. Phys. Comm.*, **180**, 1086–1097 (2009).

[4] F. Assous, J. Chaskalovic, A New Paraxial Asymptotic Model for the Relativistic Vlasov-Maxwell Equations, *C. R. Mecanique*,**340**, 706–714 (2012).

[5] F. Assous, J. Chaskalovic, A Paraxial Asymptotic Model for the Coupled Vlasov-Maxwell Problem in Electromagnetics, *J. Comput. Appl. Maths*,**270**, 369–395 (2014).

[6] F. Assous, M. Michaeli, Solving Maxwell's equations in singular domains with a Nitsche type method, *J. Computat. Phys.*, **230** (2011), 4922–4939.

[7] F. Hecht, New development in FreeFem++, *J. Numer. Math*,**20**, 251–265 (2012).

[8] C.K. Birdsall and A.B. Langdon, *Plasmas Physics via Computer Simulation* (New York: Mac.Graw-Hill, 1985).

[9] R.W, Hockney, J.W. Eastwood, *Computer simulation using particles*, Adam Hilger imprint by IOP Publishing Ltd (1988).

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis(Eds)

ADJOINT HARMONIC BALANCE METHOD FOR FORCED RESPONSE ANALYSIS IN TURBOMACHINERY

ANNA ENGELS-PUTZKA, CHRISTIAN FREY

German Aerospace Center (DLR), Institute of Propulsion Technology
Linder Höhe, 51147 Cologne, Germany
e-mail: Anna.Engels-Putzka@dlr.de, Christian.Frey@dlr.de

Key words: Aeroelasticity, Forced Response, Adjoint Methods, Harmonic Balance

Abstract. This paper describes the derivation and implementation of the discrete adjoint equations based on frequency domain methods (linear harmonics and harmonic balance) within a turbomachinery CFD code. Applications to model problems are presented which demonstrate the potential of the method for multidisciplinary turbomachinery problems, e.g. aeroelastics or aeroacoustics.

1 INTRODUCTION

Computational fluid dynamics (CFD) is increasingly used to improve the performance of turbomachinery components on the basis of numerical simulations. However, typical measures to increase the aerodynamic performance or to reduce the component weight imply a higher susceptibility to blade vibrations. Therefore, as the designs approach the aerodynamic optimum, the design problem becomes more and more multi-disciplinary. The application that the authors have in mind is a design optimisation where an aerodynamic objective (e.g. isentropic efficiency) and aeroelastic constraints (e.g. fatigue strength) are competing goals.

When going from the evaluation of a single design to CFD-based optimisation, it is important to compute also gradients of the objective functions and constraints with respect to design parameters. Since in typical applications the number of design parameters is much larger than the number of objectives, it is advantageous to use the adjoint method [1, 2]. While stationary adjoint methods are nowadays established also in the field of turbomachinery design (see e.g. [3, 4]), the use of instationary adjoint CFD is very limited due to its exceedingly high computational costs, see e.g. [5, 6, 7] and references therein.

The goal of this paper is to demonstrate how adjoint methods can be applied to frequency domain methods. These have been successfully employed for turbomachinery aeroelastic analysis, see [8] for an overview. In particular, the linear harmonic (LH) [9], the nonlinear harmonic (NLH) [10], and the harmonic balance (HB) [11] approaches are

widely used to simulate forced response and flutter. In the context of aeroelastic analysis, adjoint methods have been applied to LH methods [12, 13] as well as HB methods [14, 15, 16], and to the formulation of the HB equations for one harmonic in the time-domain [17]. In this paper, it is shown how to derive the discrete adjoint for the frequency domain methods implemented in the DLR flow solver TRACE [18, 19]. The methodology and solution techniques used for the discrete adjoint steady solver [20] are carried over to the LH and HB solvers [21, 22, 23].

2 THEORY

2.1 Frequency Domain Methods

We present here briefly the harmonic balance method as implemented in TRACE, for details we refer to [22, 23]. For simplicity we restrict the discussion to a single base frequency ω. The time-dependent flow solution $q(x,t)$ is approximated as a Fourier series where a finite number K of higher harmonics is taken into account:

$$q(x,t) = \mathrm{Re}\left[\sum_{k=0}^{K} \hat{q}_k(x) e^{ik\omega t}\right]. \tag{1}$$

Inserting this into the the time-dependent flow equation

$$\tfrac{d}{dt} q(x,t) + R(q) = 0, \tag{2}$$

where R is the discretised RANS residual, one obtains the following system of equations for the Fourier components $\hat{q}_k$:

$$ik\omega\hat{q}_k + \widehat{R(q)}_k = 0, \quad k = 0, \ldots, K. \tag{3}$$

Since R is nonlinear, the k-th harmonic of the residual, $\widehat{R(q)}_k$, may depend on all harmonics of q. Therefore, it is approximated using the Discrete Fourier Transform (DFT) for a set of sampling points $t_1, \ldots, t_n \in [0, 2\pi/\omega]$:

$$\widehat{R(q)}_k \approx \mathscr{F}(R(\mathscr{F}^{-1}(\hat{q})))|_k. \tag{4}$$

$\mathscr{F}^{-1}$ denotes the inverse transform, i.e. the reconstruction $\mathscr{F}^{-1}(\hat{q}) = (q(t_1), \ldots, q(t_n))^T$, where $q(t_j)$ is given by (1) for $t = t_j$. In the case of equidistant sampling points, i.e. $t_j = \frac{2\pi j}{\omega N}$, we obtain

$$q(x,t_j) = \mathrm{Re}\left[\sum_{k=0}^{K} e^{ik\frac{2\pi j}{N}} \hat{q}_k(x)\right], \quad j = 0, \ldots, N-1. \tag{5}$$

The corresponding Fourier coefficients are then given by

$$\hat{q}_0(x) = \frac{1}{N}\sum_{j=0}^{N-1} q(x,t_j) \tag{6}$$

$$\hat{q}_k(x) = \frac{2}{N}\sum_{j=0}^{N-1} e^{-ik\frac{2\pi j}{N}} q(x,t_j), \quad k=1,\dots,K. \tag{7}$$

When the amplitudes of the harmonic perturbations ($|\hat{q}_k|$ for $k>0$) are small, we can approximate the nonlinear Residual $R(q)$ by its linearisation about the time average, i.e.

$$R(q(t)) = R(\hat{q}_0) + \tfrac{\partial R}{\partial q}\big|_{\hat{q}_0}(q(t)-\hat{q}_0). \tag{8}$$

Then the different harmonics decouple and we obtain the steady equation $R(\hat{q}_0)=0$ and a linear equation for each $k>0$ [21, 24]:

$$\left(ik\omega + \tfrac{\partial R}{\partial q}\big|_{\hat{q}_0}\right)\hat{q}_k = 0. \tag{9}$$

2.2 Discrete Adjoint Approach

Similar to the stationary case (see e.g. [20]) we derive the discrete adjoint equations for the frequency domain methods. We assume that the objective functional I depends on a set of parameters α only through the Fourier coefficients of q, i.e.

$$\frac{dI}{d\alpha} = \frac{\partial I}{\partial \hat{q}}\frac{d\hat{q}}{d\alpha}. \tag{10}$$

For the purpose of this paper we assume that α is a parameter which influences only the values prescribed by an inhomogeneous (gust) boundary condition. Starting with the linearised equations , such a boundary condition yields an additional source term, so that (9) takes the form $A_k\hat{q}_k = S_{k,\alpha}$ [24]. From this we obtain immediately an equation for $\frac{d\hat{q}}{d\alpha}$, which can be used to eliminate this term from (10):

$$\frac{dI}{d\alpha} = \frac{\partial I}{\partial \hat{q}}A^{-1}\frac{\partial S_\alpha}{\partial \alpha} = \left((A^{-1})^*\left(\frac{\partial I}{\partial \hat{q}}\right)^*\right)^*\frac{\partial S_\alpha}{\partial \alpha} =: \hat{\psi}^*\frac{\partial S_\alpha}{\partial \alpha}, \tag{11}$$

where we have omitted the index k to simplify the notation. In the last step we have introduced the adjoint variables $\hat{\psi}$, which can be obtained by solving

$$A^*\hat{\psi} = \left(\frac{\partial I}{\partial \hat{q}}\right)^*. \tag{12}$$

For the adjoint harmonic balance method, we have to differentiate (3) with respect to α. Setting $R_{\mathrm{HB},k} := ik\omega\hat{q}_k + \widehat{R(q)}_k$, we get

$$\begin{aligned} 0 &= \frac{dR_{\mathrm{HB},k}}{d\alpha} = \frac{\partial R_{\mathrm{HB},k}}{\partial \hat{q}}\frac{d\hat{q}}{d\alpha} + \frac{\partial R_{\mathrm{HB},k}}{\partial \alpha} \\ &= \sum_j \left(ik\omega\delta_{jk} + \frac{\partial \widehat{R(q)}_k}{\partial \hat{q}_j} \right) \frac{d\hat{q}_j}{d\alpha} + \frac{\partial R_{\mathrm{HB},k}}{\partial \alpha}. \end{aligned} \tag{13}$$

The derivation of the adjoint equations is now analogous to the linear case, with $\frac{\partial S_\alpha}{\partial \alpha}$ replaced by $-\frac{\partial R_{\mathrm{HB},k}}{\partial \alpha}$. The explicit form is not discussed here, since it is only needed for the evaluation of sensitivities, which we do not consider in this article. To compute the system matrix A and its adjoint, we use the approximation (4) for the Fourier coefficients of the residual, and since $\mathscr{F}$ and $\mathscr{F}^{-1}$ are linear operations, we obtain

$$D(\widehat{R(q)}) = \mathscr{F}\left(\mathrm{diag}\left(\frac{\partial R}{\partial q}\Big|_{q(t_j)} \right) \right) \mathscr{F}^{-1} \tag{14}$$

where diag(...) denotes a block diagonal matrix with the corresponding entries on the diagonal. The submatrices are computed by reconstructing the flow solution at the sampling points t_j according to (5) and evaluating the residual Jacobian at each of these flow states.

To determine the adjoint of the matrix $D(\widehat{R(q)})$, we have to find the adjoints of $\mathscr{F}$ and $\mathscr{F}^{-1}$. The Fourier coefficients $\hat{q}$ are complex vectors, but the transformation $\mathscr{F}$ is not linear over the complex numbers. Therefore we consider all vectors as elements in real vector spaces and define the scalar product by

$$\langle \hat{\psi}, \hat{q} \rangle = \mathrm{Re}\langle \hat{\psi}, \hat{q} \rangle_{\mathbb{C}}. \tag{15}$$

The resulting adjoint transformations are given by

$$\psi(t_j) = (\mathscr{F}^* \hat{\psi})(t_j) = \frac{1}{N}\mathrm{Re}(\hat{\psi}_0) + \frac{2}{N}\mathrm{Re}\left[\sum_{k=1}^{K} e^{ik\frac{2\pi j}{N}} \hat{\psi}_k \right] \tag{16}$$

$$\hat{\psi}_k = ((\mathscr{F}^{-1})^* \psi)_k = \sum_{j=1}^{N} e^{-ik\frac{2\pi j}{N}} \psi(t_j). \tag{17}$$

The complete adjoint system matrix is then

$$A^* = \left(\mathscr{F}^{-1}\right)^* \left(\mathrm{diag}\left(\frac{\partial R}{\partial q}\Big|_{q(t_j)} \right) \right)^* \mathscr{F}^* - \mathrm{diag}(ik\omega). \tag{18}$$

3 IMPLEMENTATION

In TRACE, the steady discrete adjoint equations are solved by a preconditioned GMRes (Generalized Minimal Residual) algorithm with restarts. This has now been extended to treat several harmonics at the same time, either uncoupled (adjoint LH) or coupled (adjoint HB). The main difference for the coupled approach is that the multiplication by the system matrix is replaced by several operations according to (18). First, the adjoint solution vector is transformed into the time domain by an adjoint DFT. Then, for each sampling point, the corresponding matrix is applied, and the result is transformed back into the frequency domain. Finally, the frequency term (the original vector multiplied by $-ik\omega$) is added for each harmonic. For the computation of the matrices, the primal flow solution is reconstructed at the same sampling points as are used for the transformation of the solution vectors and the residual Jacobian is evaluated (numerically) at these flow states. For preconditioning, we use the Jacobian computed at the time-mean solution and only modify the diagonal by adding the frequency term corresponding to the current harmonic. The inverses of the modified diagonals are precomputed and stored for all harmonics. The preconditioner used in the following applications is SSOR with a relaxation factor of 0.7.

All boundary conditions are applied in the frequency domain, for each harmonic component separately. For this purpose, the stationary adjoint boundary conditions have been extended to treat complex vectors and the nonreflecting boundary conditions used at entries and exits (see [25]) now take into account the frequency, analogous to the linear solver.

Like the existing adjoint and linear solvers, the modified solver works on structured grids only and employs the constant eddy viscosity assumption. This means that possible dependencies of the eddy viscosity on the parameter α are not taken into account.

For this prototype implementation we consider only one objective functional, namely the entropy at the exit. More precisely, we consider the radial average of the circumferential Fourier coefficients for a given wave number.

4 APPLICATION

4.1 Numerical test case

As a simple numerical test case we use a segment of an annular duct with constant flow conditions. For the forward computation, an entropy wave is prescribed at the entry. The corresponding adjoint computation is done using the entropy functional (evaluated at the exit) with the same circumferential wave number. Since the underlying mean flow is constant, there is in this case no difference between linear harmonic and harmonic balance computations, therefore we present only results for the harmonic balance method. In Fig. 1 the density component of the harmonic balance solution for a plane wave propagating in axial direction, i.e. circumferential wave number zero, and the first component of the corresponding adjoint solution are shown. The adjoint solution represents the sensitivity

of the entropy functional with respect to sources in the flow field. The solutions are reconstructed from the first harmonic at four different points in time. As expected, we observe the same wave length and propagation speed in both cases.

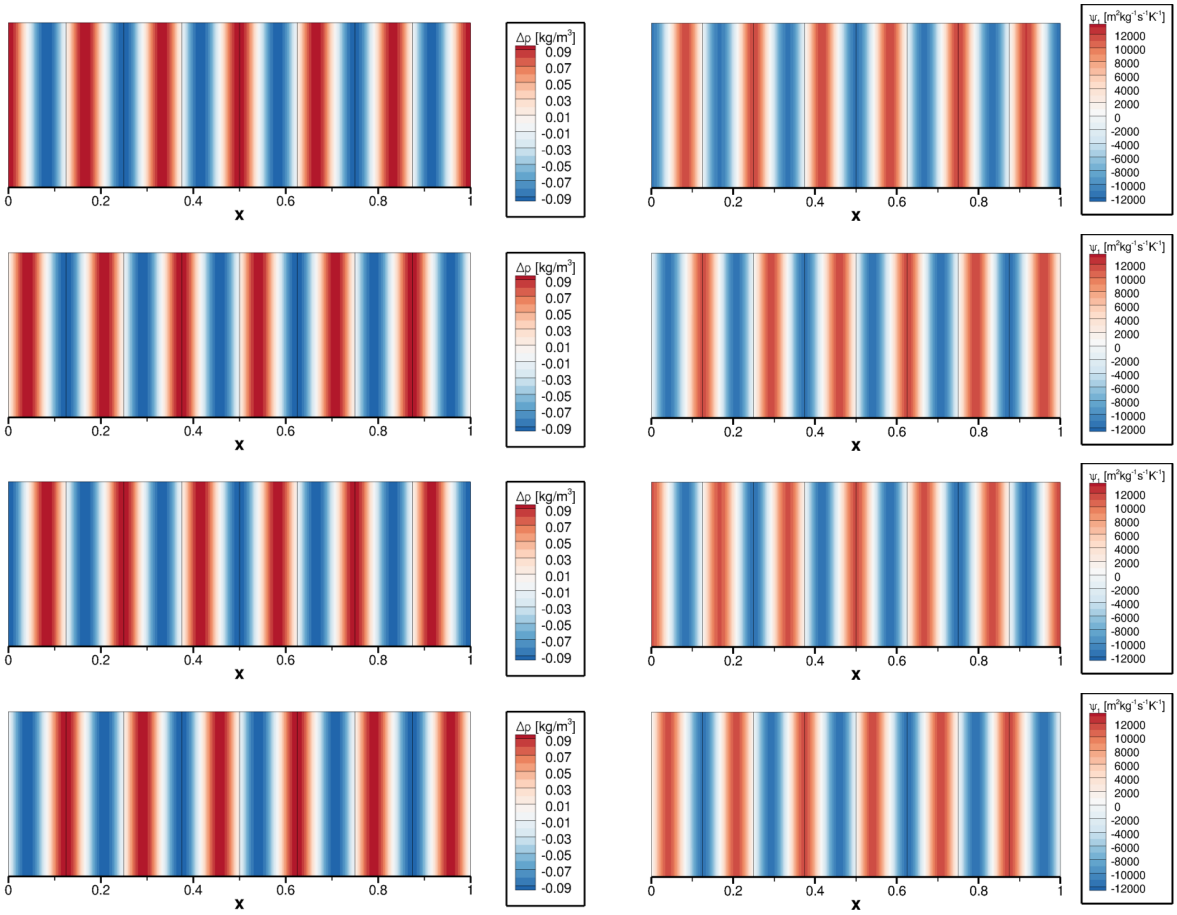


Figure 1: Reconstructed solutions (using the first harmonic) at four different times. Left: HB simulation of the propagation of an entropy wave. Right: Adjoint HB computation for the entropy functional.

In Fig. 2, the same results – but only for one point in time – are shown for a wave with a phase shift corresponding to a circumferential wave number of 16, where a similar relation between forward and adjoint solution can be observed.

4.2 Turbine rotor

As a model problem for turbomachinery applications we consider a configuration consisting of a single blade row, namely the rotor from a high pressure turbine stage. The flow conditions are subsonic with a maximum Mach number of about 0.78. The wake of the stator is extracted from a steady computation and the circumferential component with wave number $m = 70$, which corresponds to a phase shift (inter-blade phase an-

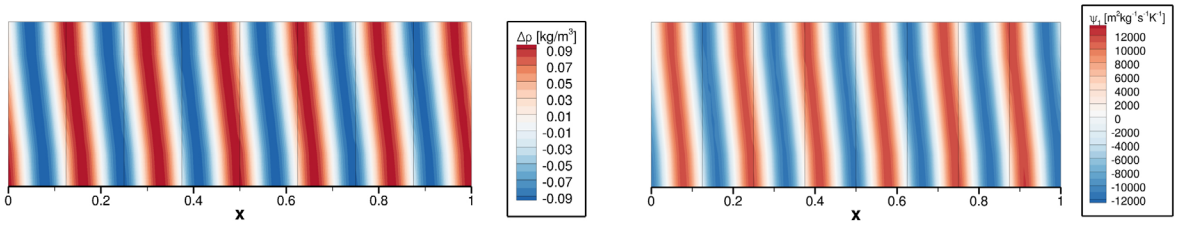


Figure 2: Reconstructed solutions (using the first harmonic) at $t = 0$. Left: HB simulation of the propagation of an entropy wave with circumferential wave number 16. Right: Adjoint HB computation for the entropy functional.

gle) of 60 degrees, is prescribed as gust boundary condition at the entry of the rotor. Computations with one and two higher harmonics are carried out. Figure 3 shows the entropy contours for the time-averaged solution and for the reconstructed solution using two higher harmonics. Although the inhomogeneous boundary condition is only used for the first harmonic, the coupling leads to a nonzero result in the second harmonic, but its magnitude is much smaller than that of the first (see Fig. 4). The effect of the coupling can also be seen in a comparison of a linear harmonic and a harmonic balance computation, each with one higher harmonic (Fig. 5).

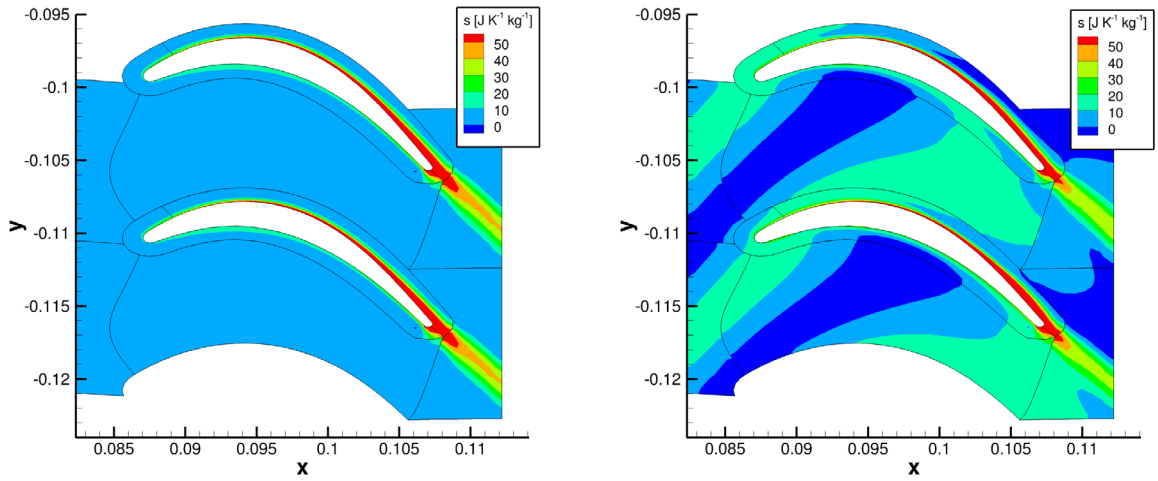


Figure 3: Entropy contours of the time averaged solution (left) and reconstructed instationary solution using two higher harmonics (right).

Similarly, in Fig. 6 we compare the results from the adjoint LH and adjoint HB methods for a computation including the zeroth and first harmonic. Some differences can be observed, although the overall structure of the solution is similar. In addition, we also carried out an adjoint HB computation with the second harmonic added, but only for the

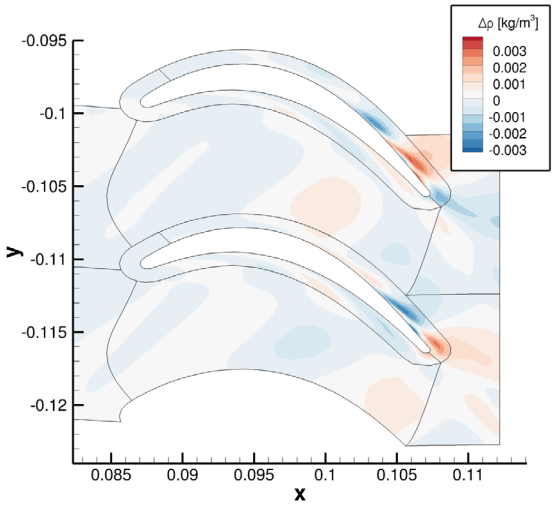


Figure 4: Second harmonic of density, reconstructed at $t = 0$, from a harmonic balance computation using two higher harmonics.

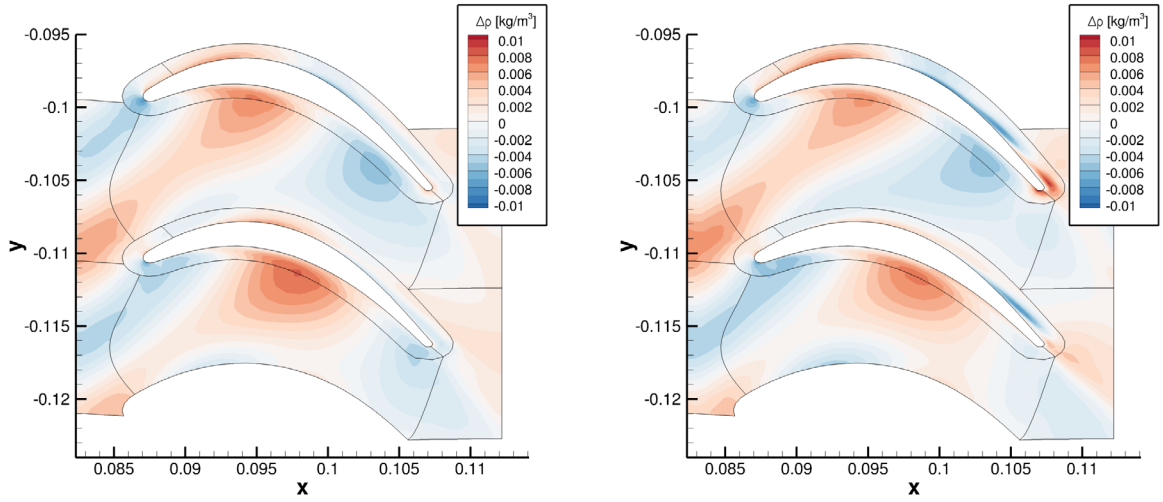


Figure 5: First harmonic of density, reconstructed at $t = 0$, from a linear harmonic (left) and a harmonic balance (right) computation including the zeroth and first harmonics.

first harmonic a non-zero right hand side is prescribed. Figure 7 shows that, as in the forward computation nonlinear effects lead to a non-negligible amplitude in the second harmonic of the adjoint solution, and also the solution for the first harmonic changes due to the coupling.

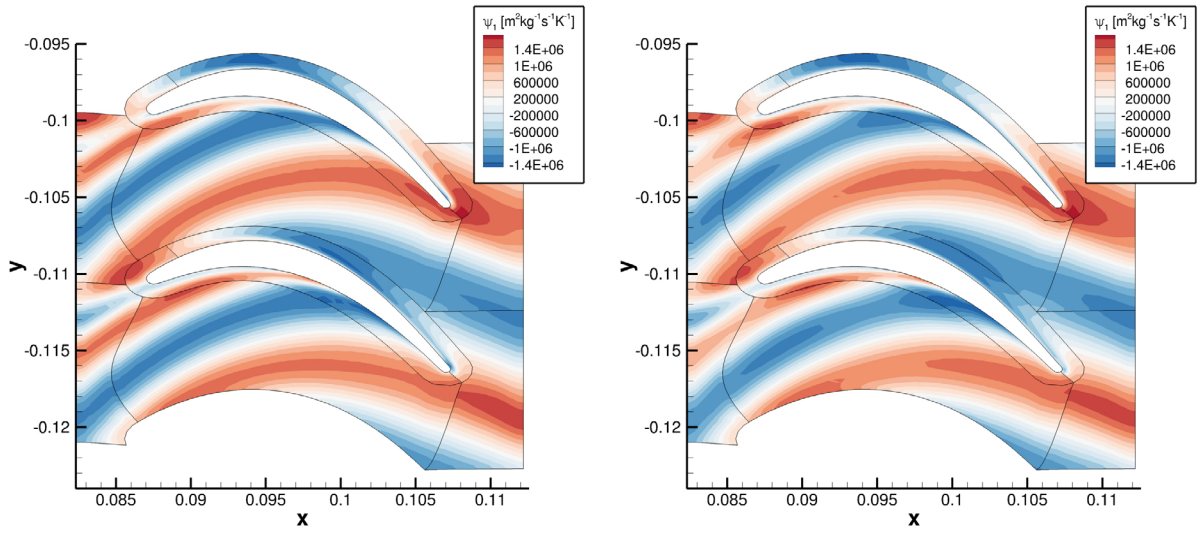


Figure 6: Density component of the adjoint solution for the entropy functional in the time domain (at $t = 0$) reconstructed using the first harmonic for linear harmonic (left) and harmonic balance (right) method.

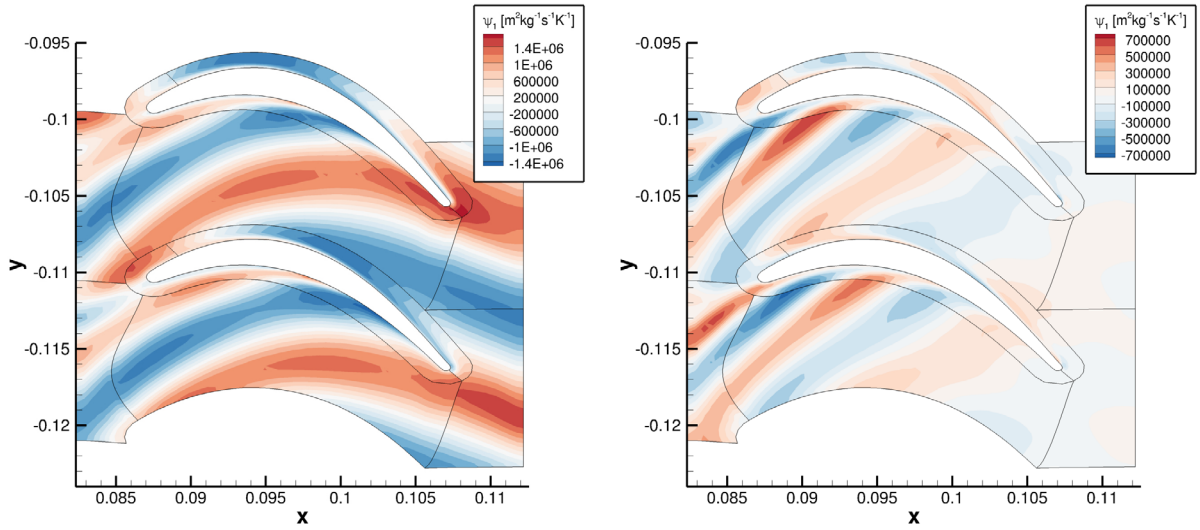


Figure 7: Density component of the first (left) and second (right) harmonic of the adjoint solution in the time domain (at $t = 0$) for the entropy functional from an adjoint HB computation using two higher harmonics.

In Fig. 8, a comparison of the convergence histories for different setups is shown. The convergence of the adjoint linear harmonic computation is similar to the steady adjoint and forward linear harmonic computations, if the linear system is solved over the complex numbers. If it is treated as a real system within the GMRes algorithm, the convergence becomes somewhat slower. The convergence behaviour of the adjoint harmonic balance computations depends strongly on the number of harmonics. If only one harmonic (besides the zeroth) is considered, the convergence is still similar to that of the adjoint linear harmonic computation using GMRes in real mode. If two higher harmonics are included, significantly more iterations are needed. It has to be investigated if the convergence can be improved by different GMRes settings (e.g. restart interval or preconditioner) or if other solution techniques, e.g. pseudo-time marching, are more suitable for this kind of problems.

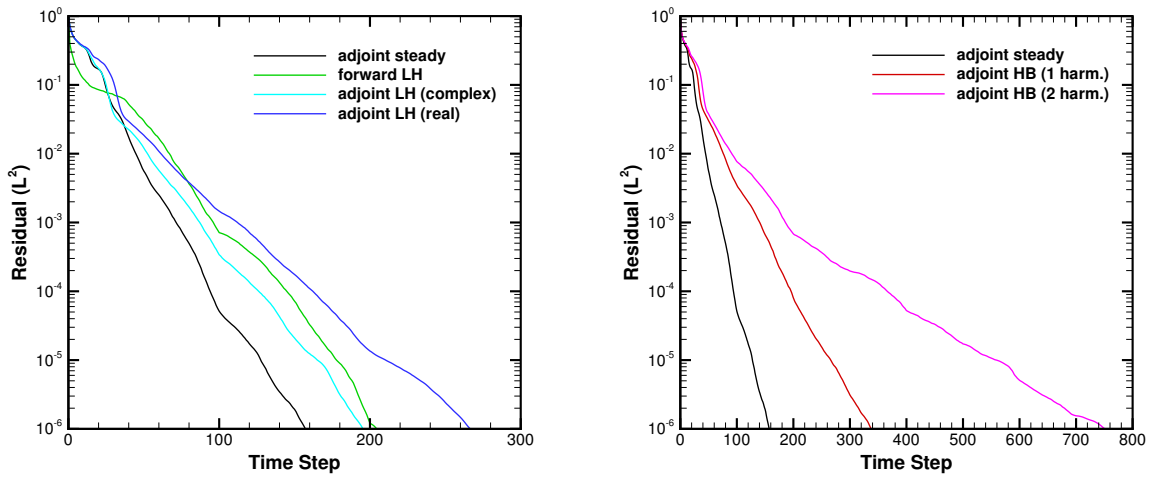


Figure 8: Convergence history for linear harmonic, adjoint linear harmonic and adjoint harmonic balance computations compared with a steady adjoint computation. All computations have been carried out using a restart interval of 100.

5 CONCLUSION

We have presented the derivation and implementation of adjoint frequency domain methods based on existing implementations of the linear harmonic and harmonic balance techniques. The functionality has been demonstrated using two model problems. The application to a turbine rotor shows the potential of the method for turbomachinery applications. Although the adjoint harmonic balance computations need significantly more computational time than the stationary adjoint, there is still a large speedup expected with respect to instationary adjoint methods in the time domain.

However, it has to be investigated if the current approach is also suitable for more complex problems or if different solution strategies are needed. Further topics for future work include the implementation of functionals which are relevant for aeroelastic analysis (e.g. modal work), the evaluation of sensitivities with respect to design parameters, the treatment of different boundary conditions and preconditioners, and the extension to several blade rows.

Acknowledgements Financial support by MTU Aero Engines (co-sponsorship of the first author) is gratefully acknowledged.

REFERENCES

[1] Jameson, A. Aerodynamic design via control theory. *J. Sci. Comput.* (1988) **3**(3):233–260.

[2] Giles, M.B. and Pierce, N.A. An introduction to the adjoint approach to design. *Flow Turbul. Combust.* (2000) **65**:393–415.

[3] Wang, D.X. and He, L. Adjoint aerodynamic design optimization for blades in multistage turbomachines—part I: Methodology and verification. *J. Turbomach.* (2010) **132**(2):021011.

[4] Wang, D.X., He, L. and Li, Y.S., Wells, R.G. Adjoint aerodynamic design optimization for blades in multistage turbomachines—part II: Validation and application. *J. Turbomach.* (2010)**132**(2):021012.

[5] Nadarajah, S.K. and Jameson, A. Optimum shape design for unsteady flows with time-accurate continuous and discrete adjoint methods. *AIAA J.* (2007) **45**(7):1478–1491.

[6] Nielsen, E.J. and Diskin, B. Discrete Adjoint-Based Design for Unsteady Turbulent Flows on Dynamic Overset Unstructured Grids. *AIAA J.* **51**(6), 1355–1373 (2013).

[7] Ntanakas, G. and Meyer, M. Towards unsteady adjoint analysis for turbomachinery applications. Proceedings of the 6th. European Conference on Computational Fluid Dynamics - ECFD VI, pp. 5071–5081 (2014).

[8] He, L. Fourier methods for turbomachinery applications. *Prog. Aerosp. Sci.* (2010) **46**(8):329–341.

[9] Hall, K.C. and Crawley, E.F. Calculation of unsteady flows in turbomachinery using the linearized Euler equations. *AIAA J.* (1989) **27**(6):777–787.

[10] He, L. and Ning, W. Efficient approach for analysis of unsteady viscous flows in turbomachines. *AIAA J.* (1998) **36**(11):2005–2012.

[11] Hall, K.C., Thomas, J.P. and Clark, W.S. Computation of unsteady nonlinear flows in cascades using a harmonic balance technique. *AIAA J.* (2002) **40**(5):879–886.

[12] Duta, M.C., Giles, M.B. and Campobasso, M.S. The harmonic adjoint approach to unsteady turbomachinery design. *Int. J. Numer. Methods Fluids* (2002) **40**(3-4):323–332.

[13] Duta, M., Campobasso, M., Giles, M. and Lapworth, L. Adjoint harmonic sensitivities for forced response minimization. *J. Eng. Gas Turbines Power* (2006) **128**(1):183–189.

[14] Thomas, J., Hall, K. and Dowell, E. Discrete adjoint approach for modeling unsteady aerodynamic design sensitivities. *AIAA J.* (2005) **43**(9):1931–1936.

[15] Nadarajah, S.K., McMullen, M.S. and Jameson, A. Aerodynamic shape optimization for unsteady three-dimensional flows. *Int. J. Comput. Fluid Dyn.* (2006) **20**(8):533–548.

[16] Nadarajah, S.K. and Jameson, A. Optimum shape design for unsteady three-dimensional viscous flows using a nonlinear frequency-domain method. *J. Aircraft* (2007) **44**(5):1513–1527.

[17] He, L. and Wang, D.X. Concurrent Blade Aerodynamic-Aero-elastic Design Optimization Using Adjoint Method. *J. Turbomach.* (2011) **133**(1):011021.

[18] Nürnberger, D., Eulitz, F., Schmitt, S. and Zachcial, A. Recent progress in the numerical simulation of unsteady viscous multistage turbomachinery flow. ISABE 2001-1081 (2001).

[19] Becker, K., Heitkamp, K. and Kügeler, E. Recent progress in a hybrid-grid CFD solver for turbomachinery flows. Proceedings Fifth European Conference on Computational Fluid Dynamics ECCOMAS CFD 2010 (2010).

[20] Frey, C., Nürnberger, D. and Kersken, H.P. The discrete adjoint of a turbomachinery RANS solver. Proceedings of ASME-GT2009 (2009).

[21] Kersken, H.P., Frey, C., Voigt, C. and Ashcroft, G. Time-Linearized and Time-Accurate 3D RANS Methods for Aeroelastic Analysis in Turbomachinery. *J. Turbomach.* (2012) **134**(5):051024.

[22] Frey, C., Ashcroft, G., Kersken, H.P. and Voigt, C. A harmonic balance technique for multistage turbomachinery applications. Proceedings of ASME Turbo Expo 2014 (2014).

[23] Ashcroft, G., Frey, C. and Kersken, H.P. On the development of a harmonic balance method for aeroelastic analysis. 6th European Conference on Computational Fluid Dynamics (ECFD VI) (2014).

[24] Frey, C., Ashcroft, G., Kersken, H.P. and Weckmüller, C. Advanced numerical methods for the prediction of tonal noise in turbomachinery — Part II: Time-linearized methods. *J. Turbomach.* (2013) **136**(2):021002.

[25] Frey, C., Engels-Putzka, A. and Kügeler, E. Adjoint boundary conditions for turbomachinery flows. ECCOMAS 2012 - European Congress on Computational Methods in Applied Sciences and Engineering, e-Book Full Papers (2012).

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis (Eds)

APPROACHES FOR COUPLED NUMERICAL SIMULATION OF HIGH FREQUENCY TUBE WELDING PROCESS

ALEXANDER NIKANOROV[*], EGBERT BAAKE[*], HOLGER BRAUER[†] AND CHRISTOPH WEIL[††]

[*] Institute of Electrotechnology, Leibniz Universität Hannover
Wilhelm-Busch-Str. 4, D-30167 Hannover, Germany
e-mail: nikanorov@etp.uni-hannover.de, baake@ etp.uni-hannover.de

[†] Salzgitter Mannesmann Line Pipe GmbH
Kissinger Weg, 59067 Hamm, Germany
e-mail: Holger.Brauer@smlp.de

[††] Salzgitter Mannesmann Line Pipe GmbH
In der Steinwiese 31, 57074 Siegen, Germany
e-mail: Christoph.Weil@smlp.de

Key words: High Frequency Tube Welding, Induction Heating, Numerical Simulation, Coupled Analysis.

Abstract. Welding processes and installations used nowadays are mainly developed on practical experience and analytical calculations. Nevertheless, high frequency induction tube welding is a very complex three-dimensional dynamic process, where the electromagnetic and thermal characteristics are distributed not only in space but in time as well. A more profound detailed investigation of the induction tube welding process can be only done by numerical modelling. Full and local three-dimensional transient numerical models of induction tube welding process with continuous movement of the welded tube have been developed and tested. Coupled electromagnetic and thermal analyses are carried out at each time step of simulation for correction of temperature dependent material properties. Voltage or current of the induction coil can be individually input into electromagnetic analysis at each time step. This approach allows simulating "quasi" steady-state and transient operation modes.

1 INTRODUCTION

For the production of tubes made of steel (ferritic or austenitic) as well as made of aluminium, brass, copper or zinc various conduction or induction welding technologies are used in industry. The longitudinal seam welding of the tube can be done by the use of high-frequency (HF) resistance heating (conductive) or inductive heating.

The HF induction welding process is executed with the application of a high frequency parallel or series resonant circuit converter for the energy supply. The welding is effected by a ring shaped or profiled inductor, which includes the tube and induces a high frequency current into it. The current passing the strip edges heats the material with an increasing temperature towards the welding spot. The welding itself is now here effected without any additional material by pure pressure from upsetting rolls. The welding beads are trimmed in hot condition behind the upsetting rolls on the outside and on the inside if necessary.

Contrary to other methods of induction heating where the heating is affected directly next to the inductor, the heated area is mainly outside the inductor loop area at longitudinal seam welding. The currents induced under the inductor are passing the back of the tube, then the edges of the strip and meet at the welding spot. There then the required welding temperature of about e.g. 1400°C is achieved.

2 HIGH FREQUENCY APPLICATIONS FOR THE TUBE PRODUCTION

For the production of tubes made of steel (ferritic or austenitic) or steel coated with zinc or aluminium as well as made of aluminium, brass, copper or zinc conduction or induction welding and induction heating technologies are used in industry today as shown in Figure 1.

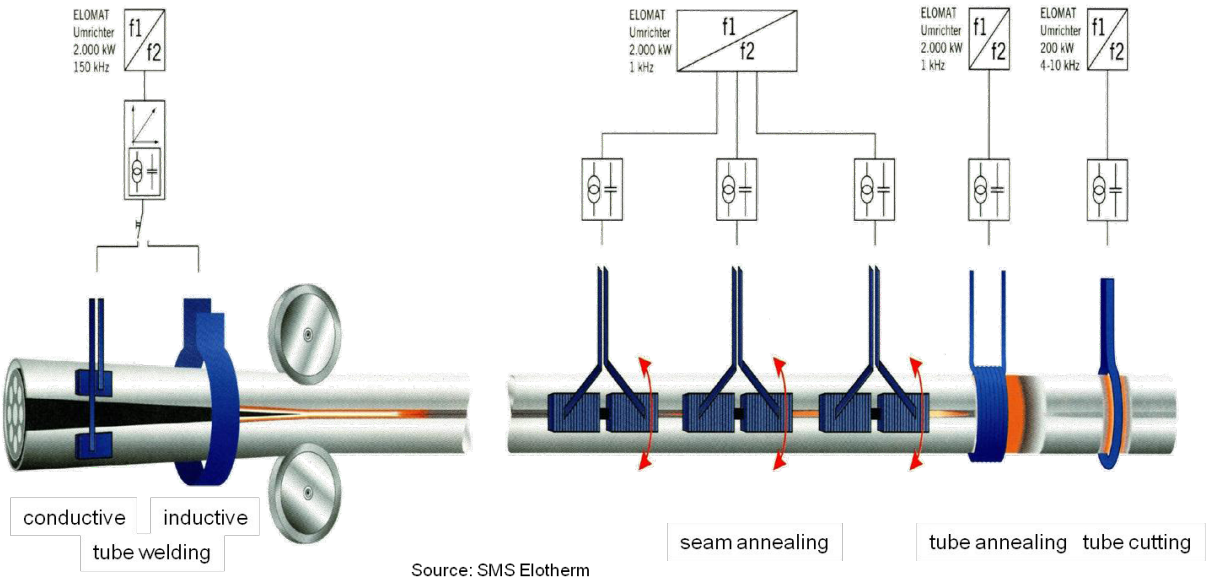


Figure 1: HF applications and induction applications for the production of tubes
(Courtesy: SMS ELOTHERM)

The longitudinal seam welding of the tube can be done by the use of high-frequency (HF) resistance heating (conductive heating) or inductive heating as explained more in detail in following chapter. Due to high cooling rates during the welding process, a different metallurgical structure and different mechanical properties compared to the base material occurs in the seam and in the heat affected zone.

An induction seam annealing installation can be integrated in the welding line to achieve a homogeneous grain structure on the whole circumference. If the annealing of the complete tube is necessary from metallurgical point of view an induction tube heating system can be used. A cutting installation is arranged at the end of the tube production line in order to cut the tubes into conventional lengths. This unit can be operated mechanically, e.g. as a travelling saw, or in certain cases even by induction. In this latter case, the tube is heated by a short one turn inductor and then separated by a tension pulse in axial direction.

The high frequency tube welding involves the application of a high-frequency alternating current in the range 100 - 500 kHz, with the tube forming and energy input operations being performed by separate units. This welding method simultaneously utilizes pressure and heat in order to join the strip edges of the open-seam tube together without the addition of a filler metal. Squeeze and pressure rolls in double- or multi-roll weld stands bring the edges of the open-seam tube gradually together and apply the pressure necessary for welding (see Figures 2 and 3). High-frequency alternating current offers a number of benefits as energy source for generating the heat required for the welding process. Due to the skin and proximity effect the current thus flows along the strip edges of the open-seam tube to the point at which the strip edges abut (welding point), and the ensuing concentration. Below the Curie point, the depth of current penetration only amounts to a few hundredths of a millimetre. Once the steel is heated above this temperature, it becomes non-magnetic and the current penetration depth rises to several tenths of a millimetre at frequencies in the region of 450 kHz.

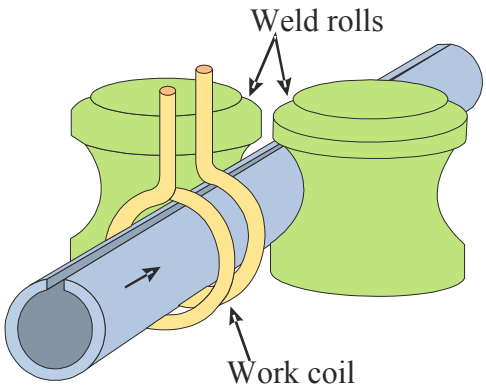


Figure 2: Schematic of the longitudinal high frequency seam welding of tubes using inductive welding

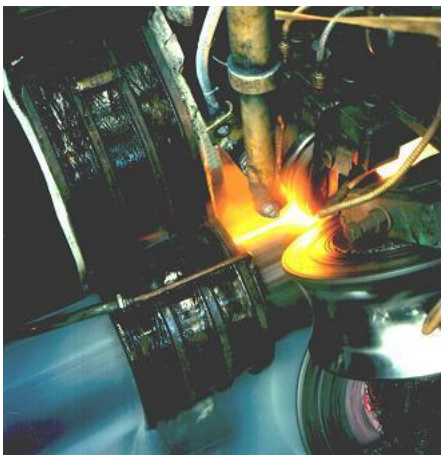


Figure 3: Example of the high frequency induction welding of tubes

The welding current can be introduced into the open-seam tube both by conductive means using sliding contacts and by inductive means using single or multi-wind coils. Consequently, a distinction is made in the nomenclature between high-frequency induction (HFI) welding and high-frequency conduction welding.

High frequency induction tube welding is a very complex multi-physical three-dimensional process. Additionally, the induced current and temperature are distributed not only in space but in time as well. The electromagnetic-thermal process parameters are depending on the geometry of the inductor, impeder and in particular on the tube (e.g. wall thickness, incoming angle etc.), the operation parameter, like total power, inductor current, frequency, welding feed etc. and finally also on the material data like electrical conductivity and other, which are temperature dependent or in case of magnetic permeability temperature and magnetic field strength dependent.

All of these influence factors and physical correlations have to be taken into account for the design and optimization of the HF induction tube welding process. Therefore a more profound detailed investigation of the physical relations of the process can be only done by numerical modelling. In order to simulate the three-dimensional (3D) "quasi" steady state and if necessary transient mode of an induction tube welding system, special numerical models are required. The models must simulate the heating process distributed in space and in time. That is why the models of induction tube welding must be based on special algorithm providing a time loop additionally to coupling between electromagnetic and thermal analysis.

Geometry of the welded edges and their position in the welding angle should be precisely input into the model because of their strong influence to the process. Any simplification in geometrical description of the welded edges made like in [1] distorts the results of simulation.

3 ALGORITHM OF COUPLING BETWEEN ELECTROMAGNETIC AND THERMAL ANALYSES

The developed numerical models of induction tube welding process with continuous movement of the tube are based on one specially created algorithm shown in Figure 4.

For numerical simulation the continuously running physical heating process is replaced by big enough number of time steps. Electromagnetic and thermal analyses are carried out at each time step of simulation. The Joule heat distribution in the tube, calculated in the electromagnetic analysis, is used as an excitation for the thermal one at the running time step. Temperature dependent electro-physical material properties are corrected for electromagnetic analysis at the running time step according to temperature distribution in the tube after the previous time step. Thermal analysis with temperature dependent properties of steel includes also simulation of thermal losses by convection and radiation from all open surfaces of the calculated system. Heat flux by radiation is calculated taking into account view angles. Temperature dependent properties of steel used for simulation are shown in Figure 5.

Input of voltage or current of the induction coil is individually input in electromagnetic analysis at each time step. This approach allows simulating various kinds of transient modes. The "quasi" steady-state operation mode can be reached as well via transient one after long enough time.

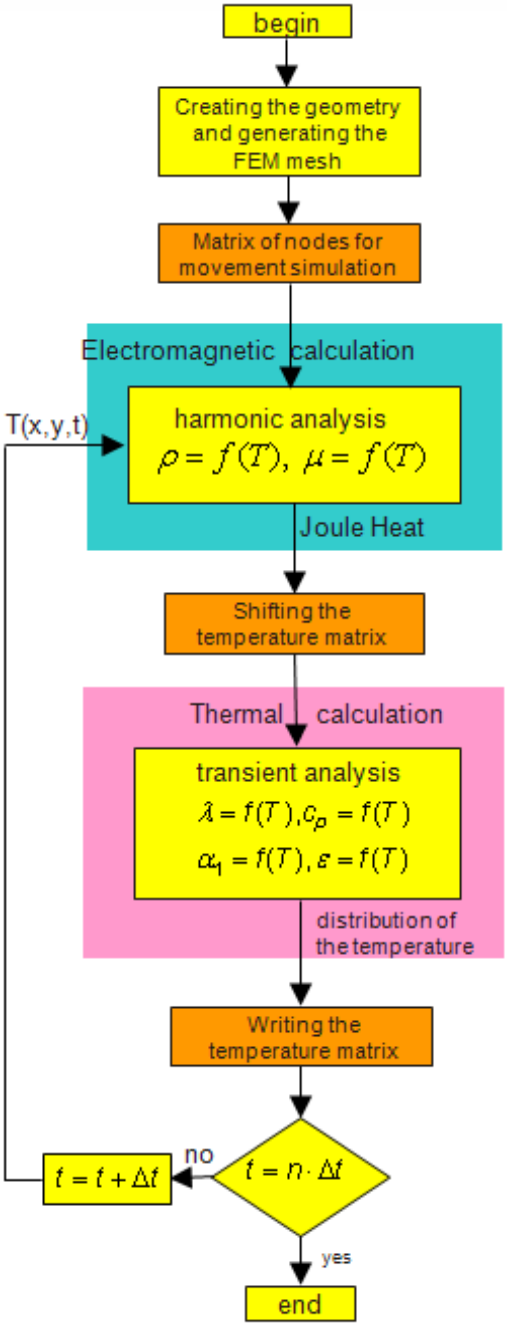


Figure 4: Algorithm of transient coupled electromagnetic-thermal analysis with the tube movement

One robust way to implement movement of the workpiece is based on shifting the temperature field before each time step of thermal analysis [2]. It is very effective for simulation of induction heating systems with continuously moved endless workpiece of constant cross-section. However, uniform numerical mesh in the workpiece in the direction of motion is required for this approach. Speed of the workpiece is taken into account via values of simulation time steps. This approach was modified for the tube welding process to run with non-uniform mesh in the tube which is very important for tube welding simulation. Arbitrary simulation time steps can be applied in this case as well.

Two 3D transient numerical models of the induction tube welding process with continuous movement of the tube have been developed according to the created algorithm. The models are realized using Finite Element Method on the basis of commercial program package ANSYS.

4 NUMERICAL MODELS FOR INDUCTION TUBE WELDING PROCESS

Full model of the welding system includes the welded tube with V-angle, the induction coil, impeder (see Figure 6) and the surrounded air which is necessary for spreading of magnetic field. FEM element mesh in the tube has to be very fine because of small electromagnetic penetration depth at high frequency. Only a proper chosen non-homogeneous mesh allows

reaching an acceptable compromise between the total amount of elements and the computer runtime needed for simulation.

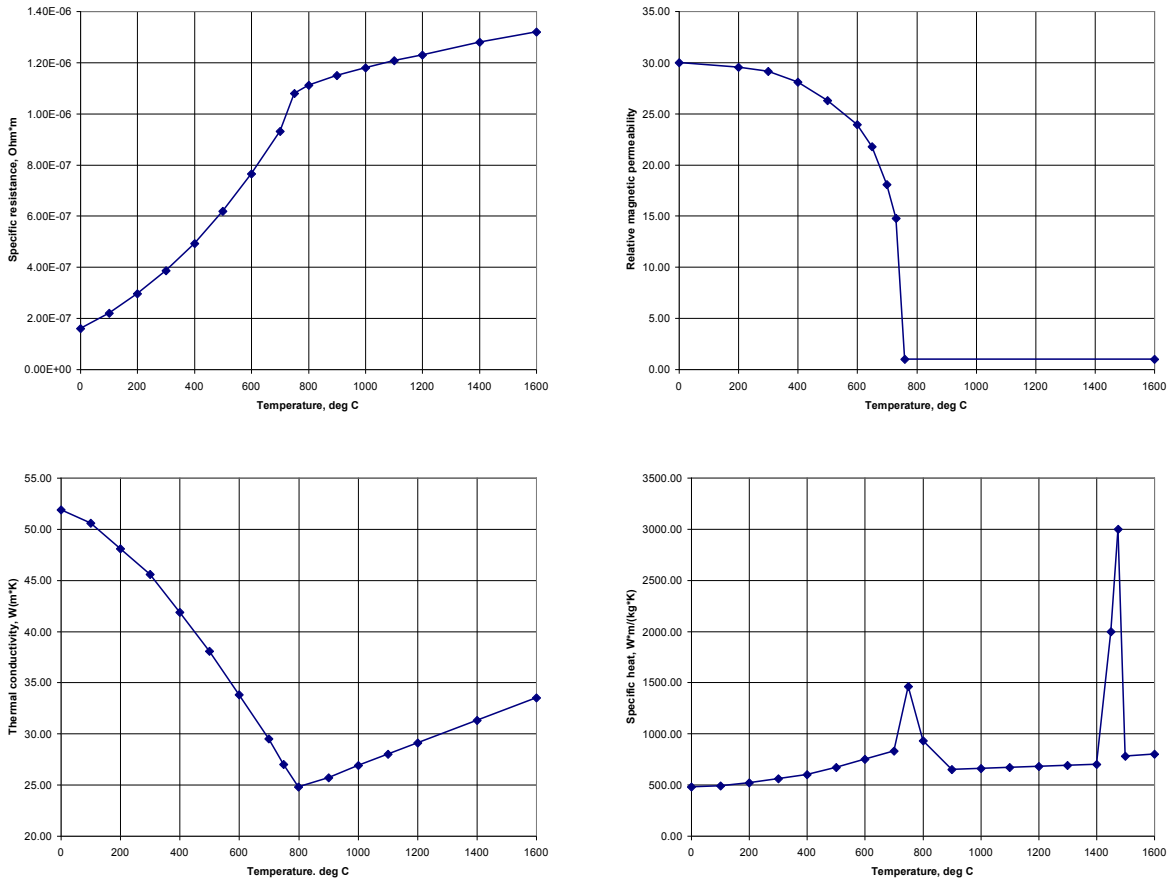


Figure 5: Electro-physical and thermo-physical properties of steel used in the model

In the presented approach the tube edges are of fixed shape even above the melting temperature. That is why the numerical prediction of temperature can exceed the melting point. One example of temperature distribution in the welded tube edge is shown in Figure 7. Temperature distribution in the second welded edge is symmetric to the shown one. Before the welding point temperature grows because of eddy currents induced in the pipe. After the welding point no currents are running, so temperature drops down as a result of thermal equalization.

The full model is absolutely necessary for parametrical study and optimization of the induction system geometry and electrical parameters of the induction coil and the other process improvements.

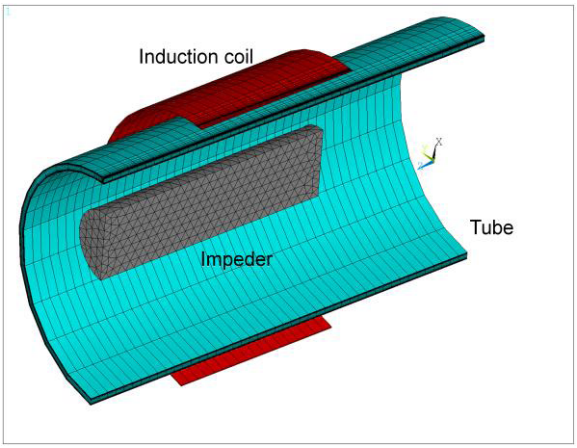


Figure 6: Geometry and FEM mesh in the tube in the full model

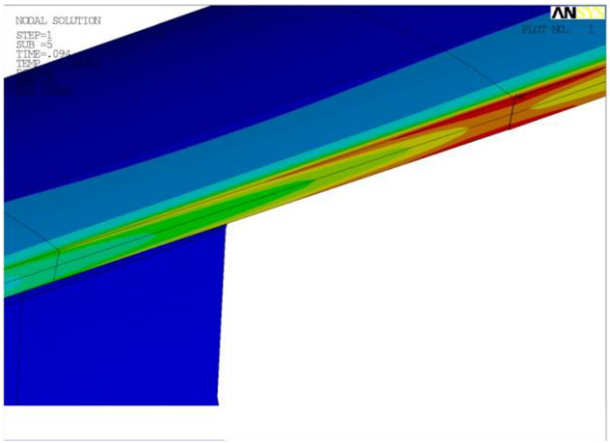


Figure 7: "Quasi" steady-state temperature field in the tube received by the full model

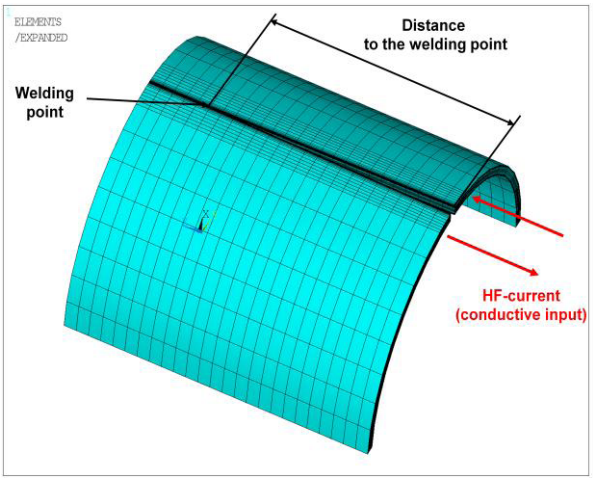


Figure 8: Geometry and FEM mesh in the tube in the local model

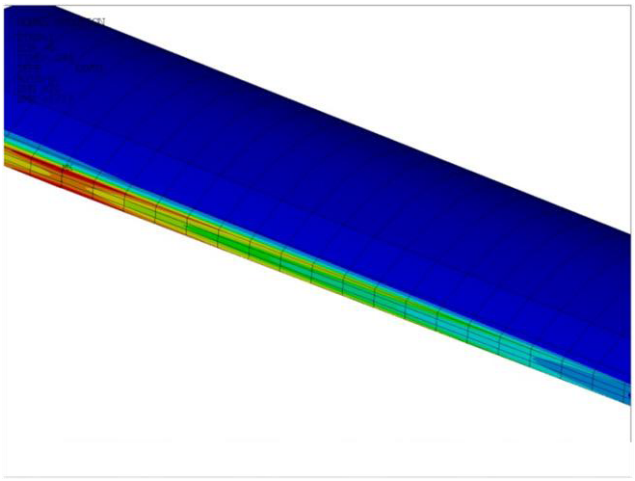


Figure 9: "Quasi" steady-state temperature field in the tube received by the local model

The second model, which is called local, includes the welded tube and the air surrounding only (see Figure 8). In spite of the full model, the current is implemented directly into the welded edges of the tube like it is made in HF conduction welding process. The local model approach allows increasing the element amount in the simulated system without rising of the runtime. It is necessary for deeper analysis of electromagnetic and thermal effects directly in the region of welding point.

One example of temperature distribution in the welded tube edge received by the local model is shown in Figure 9. The local model is very effective for investigation of electromagnetic, thermal and other effects around the welding point.

Temperature distribution is very similar to the results from the full model if the distance from the current implementation to the welding point is chosen in a proper way. In front of the welding point temperature grows because of eddy currents induced in the pipe. Like in the previous case, no currents are running after the welding point, so temperature drops down as a result of thermal equalization.

To compare the both developed full and local models, relative temperature profiles over the tube wall thickness in the welding point are shown in Figure 10.

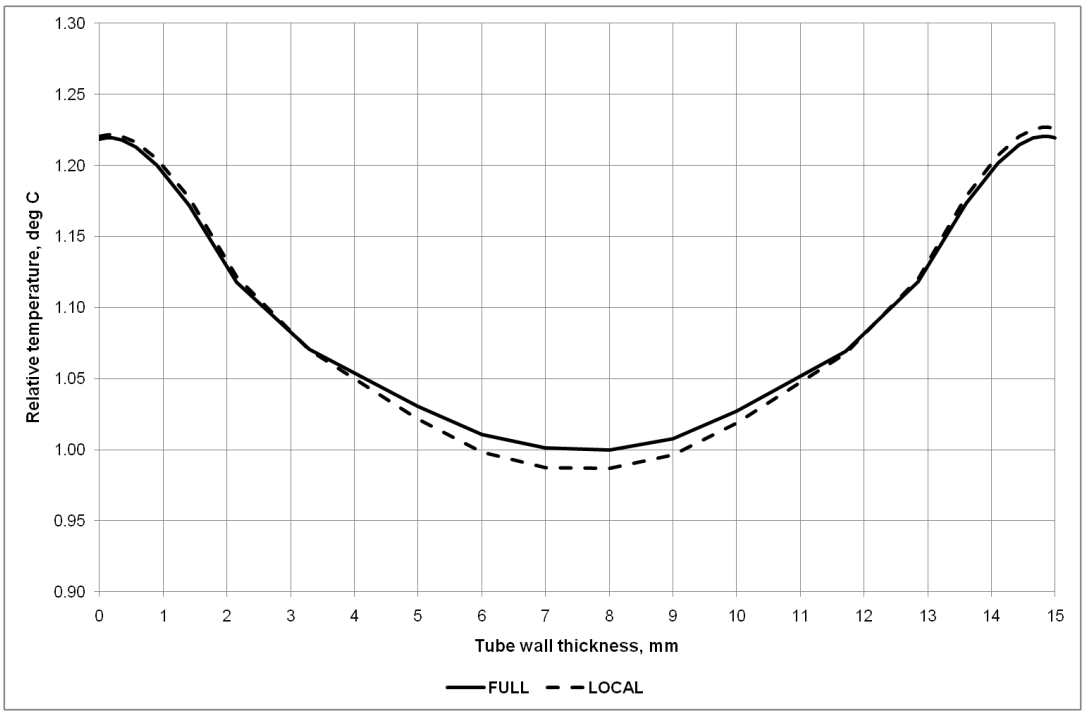


Figure 10: Relative temperature profiles over the tube wall thickness in the welding point received by the full and the local models

The temperature level at the welding point depends on power which is defined by the given induction coil current or voltage. This current or voltage should be tuned to reach the reference temperature given for the certain point of the tube.

Difference between the temperature curves does not exceed 20 K in the middle part of the tube edge. Because of the welded edges in the V-angle are not parallel to each other, the temperature profile is of small asymmetry.

5 CONCLUSIONS

- High frequency tube welding is a very complex process with various physical phenomena. Detailed investigation and optimization of the induction tube welding process can be only done by numerical modelling.
- Full and local three-dimensional transient numerical models of induction tube welding process with continuous movement of the welded tube have been developed. They are based on coupled electromagnetic and thermal analyses with temperature dependent electro- and thermo-physical material properties.
- The full model is absolutely necessary for parametrical study and optimization of the induction system geometry and electrical parameters of the induction coil. The local model is very effective for investigation of electromagnetic, thermal and other effects around the welding point.
- Temperature profiles over the tube wall thickness in the welding point received by full and local models are very similar if the distance from the current implementation to the welding point is chosen in a proper way.

REFERENCES

[1] Dughiero, F.; Forzan, M.; Pozza, C.; Sieni, E.: *A translational coupled electromagnetic and thermal innovative model for induction welding of tubes.* IEEE Transactions on Magnetics, Vol. 48, No. 2, February 2012, pp.483-486

[2] Galunin, S.; Zlobina, M.; Blinov, K.; Nikanorov, A.; Zedler, T,; Nacke, B.: *Numerical analysis of coupled physics for induction heating of movable workpieces.* Proceedings of International Scientific Colloquium Modelling for Electromagnetic Processing (MEP2008), Hannover, Germany, October 27 – 29, 2008, pp.59-64

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis(Eds)

BALLISTIC CHARGE TRANSPORT IN A TRIPLE-GATE SILICON NANOWIRE TRANSISTOR

O. MUSCATO, T. CASTIGLIONE AND C. CAVALLARO

Dipartimento di Matematica e Informatica
Università degli Studi di Catania
Viale A. Doria 6, 95125 Catania, Italy
e-mail: muscato@dmi.unict.it, castiglione@dmi.unict.it

Key words: Silicon nanowires, Schrödinger-Poisson-Boltzmann system, Variational Method, ballistic transport.

Abstract. In this paper we investigate the electrostatics and charge transport in a triple-gate Silicon Nanowire transistor. The quantum confinement in the transversal dimension of the wire have been tackled using the Schrödinger equation in the Effective Mass Approximation coupled to the Poisson equation. This system have been solved efficiently using a Variational Method. The charge transport along the longitudinal dimension of the wire has been considered using the semiclassical approximation, in the ballistic regime.

1 INTRODUCTION

Silicon nanowires (SiNW) are considered an interesting alternative architecture to the conventional planar technology for electronic devices, because different electronic structures and transport properties in one dimension can be utilized to fabricate high performance and highly packed integrated circuits. By shrinking the cross-section of SiNW electronic devices, effects of quantum confinement are observed and the wave nature of the electrons must be taken into account. In addition, it is well known that the behavior of field-effect transistors is dominated by electrostatics, which has therefore to be accurately simulated in order to reproduce the device electrical behavior. In order to accurately model transport charge in SiNW one have to take into account its transversal and longitudinal dimensions. For nanowires with transversal dimensions greater than 5 nm, the 2D Schrödinger equation in the Effective Mass Approximation (EMA) represents a good approximation to describe the quantum confinement [1], but below, atomistic electronic structure models must be employed. For longitudinal lengths (called channel) greater than 20nm, the charge transport can be described with a semiclassical scheme, in which the charges are treated as newtonian particles. Below this channel dimension, quantum mechanical simulations based on the nonequilibrium Greens function (NEGF) formalism

are mandatory. The goal of this paper will be the investigation of the charge transport in a triple-gate SiNW transistor within the effective mass approximation, in the semiclassical regime. The EMA equation will be solved with a Variational Method, with the main advantage to eliminate discretization errors as well as to cut down computational time with respect to a purely-numerical approach. The charge transport along the channel will be tackled in a ballistic regime.

2 Transport equations

In the following we shall consider a SiNW with rectangular cross section. For a quantum wire with longitudinal expansion in z-direction, and confined in the plane x-y, the normed electron wave function $\psi(x,y,z)$ can be written in the form

$$\psi(x,y,z) = \chi_\alpha(x,y)\frac{e^{ik_z z}}{\sqrt{L_z}} \tag{1}$$

where $\chi_\alpha(x,y)$ is the wave function of the α-th subband and the term $e^{ik_z z}/\sqrt{L_z}$ describes an independent plane wave in z-direction confined to the normalization length, where $z \in [0, L_z]$ and k_z is the wave vector number. In general the electron is subject to external confining potential U, such as by a discontinuity in the band gap at an interface between two materials, and also to the effect of the other electrons in the system. The simplest approximation, called *Hartree approximation*, is to assume that the electrons as whole produce an average electrostatic energy potential, and that a given electron feels the resulting total potential V_{tot}

$$V_{tot}(x,y,z) = U(x,y) - e\Phi(x,y,z) \quad . \tag{2}$$

The normed wave function satisfies the Schrödinger equation in the Effective Mass Approximation, i.e.

$$\left[-\frac{\hbar^2}{2m^*}\Delta + V_{tot}(x,y,z)\right]\psi = E\,\psi \tag{3}$$

where E is the total energy, and m^* denotes the effective mass of the electron in the conduction band. By inserting eq.(1) into eq.(3), in each z-th cross section of the device, one obtains the following equation for the envelope function $\chi_{\alpha z}(x,y)$

$$\begin{cases} H\chi_{\alpha z} = \varepsilon_{\alpha z}\chi_{\alpha z} \\ H = \left[-\frac{\hbar^2}{2m^*}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right) + V_{tot}(x,y,z)\right] \\ E_{\alpha z} = \varepsilon_{\alpha z} + \frac{\hbar^2 k_z^2}{2m^*} \end{cases} \tag{4}$$

where $\varepsilon_{\alpha z}$ is the kinetic energy associated to the confinement in the x-y plane, and we have assumed the parabolic band approximation. The term Φ satisfies the Poisson equation

$$\nabla\cdot[\epsilon\nabla\Phi(x,y,z)] = e(n - N_D + N_A) \tag{5}$$

where N_D, N_A are the doping profile (due to donors and acceptors) and $n(x,y,z,t)$ is the electron density, which depends on $\chi_{\alpha z}$

$$n(x,y,z,t) = \sum_{\alpha} \rho^{\alpha}(z,t)|\chi_{\alpha z}(x,y,t)|^2 \tag{6}$$

where ρ^{α} is the subband linear density in the z-direction

$$\rho^{\alpha}(z,t) = \frac{2}{2\pi}\int f_{\alpha}(z,k_z,t)dk_z \tag{7}$$

f_{α} being the electron distribution function in the α-subband. For an assigned confining potential, one has to solve a coupled problem formed by the eqs.(4), (5) and (6) to find $\varepsilon_{\alpha z}, \chi_{\alpha z}$ in each cross-section. The subband electron distribution function can be obtained by solving the 1-D Multiband Boltzmann Transport Equation (MBTE), which forms an integro-differential system in two dimensions in the phase-space and one in time, with a complicate collisional operator. The full solution of the MBTE can be obtained or by using cpu-demanding methodologies such as the Monte Carlo one [2]-[10] or deterministic numerical solvers [11],[12],[13], or by introducing hydrodynamic models [14]-[18].

3 Solution of the EMA equation

For a fixed total potential $V_{tot}(x,y,z)$ in each cross-section (with z= const.) one has to solve the eigenvalues problem (4). This can be achieved or by using a finite difference numerical scheme, or by using a Variational Method. With a finite difference scheme, a 2D spatial mesh with N_1 grid points in the x-dimension and N_2 in y-dimension is introduced. If we approximate the differential operator $\frac{\partial^2}{\partial x^2}$ with a central difference formula, then the eigenvalues problem (4) reduces to a system of linear equations having a $N_1 \times N_2$ matrix. The discretization error, in this case, is quadratic in the mesh size [19].

With the Variational Method [20], the eigenvalue problem of the Hamilton operator H is approximated by the eigenvalue problem of the matrix $\mathcal{H}$

$$\mathcal{H}\mathbf{v} = \lambda\mathbf{v} \quad , \quad \mathcal{H}_{\{i\}\{j\}} = \left\langle \psi_{\{i\}}|H|\psi_{\{j\}}\right\rangle \quad , \quad \mathbf{v} = (v^1, v^2, ...) \tag{8}$$

where the matrix elements $\mathcal{H}_{\{i\}\{j\}}$ are expectation values of the Hamilton operator H with respect to a set of orthonormal functions $\{\left|\psi_{\{1\}}\right\rangle, \left|\psi_{\{2\}}\right\rangle, ...\}$ which span the Hilbert space of H. Then the α-th approximated eigenvalue of the Hamilton operator H is the α-th eigenvalue of the matrix $\mathcal{H}$ i.e. $\varepsilon_{\alpha} \simeq \lambda_{\alpha}$, and the α-th eigenfunction χ_{α} of the Hamilton operator H is approximates as

$$\chi_{\alpha} \simeq \sum_{\{i\}} v_{\alpha}^{\{i\}} \left|\psi_{\{i\}}\right\rangle \quad . \tag{9}$$

We observe that the number of the orthonormal functions that span the Hilbert space of H is infinite. Since the number of used orthonormal functions is finite, the solutions of

$\mathcal{H}$ span in general a subspace of H. Hence the variational method is only an advantage if the set of orthonormal functions is chosen appropriately so that the matrix elements $\mathcal{H}_{\{i\}\{j\}}$ are easily calculable and the accuracy of the approximation can be ensured for a small number of orthonormal functions. A good choice for a set of orthonormal functions are the eigenstates of the two-dimensional anisotropic harmonic oscillator which are given by

$$|\psi_{j_x,j_y}\rangle = \frac{(ab)^{\frac{1}{4}}}{\sqrt{\pi 2^{j_x+j_y} j_x! j_y!}} \exp\left(-\frac{ax^2}{2} - \frac{by^2}{2}\right) H_{j_x}(\sqrt{a}x) H_{j_y}(\sqrt{a}y) \tag{10}$$

$$a = \frac{m^*\omega_x}{\hbar} \quad , \quad b = \frac{m^*\omega_y}{\hbar} \quad , \quad \{j\} = (j_x, j_y) \in \mathbb{N} \times \mathbb{N} \tag{11}$$

where j_x, j_y are the quantum numbers, $H_{j_x}(\sqrt{a}x)$ is the Hermite polynomial of order j_x and ω the angular velocity. Then, the elements of the matrix $\mathcal{H}_{\{i\}\{j\}}$ are formed by two contributions:

$$\mathcal{H}_{\{i\}\{j\}} = \mathcal{H}^{(1)}_{\{i\}\{j\}} + \mathcal{H}^{(2)}_{\{i\}\{j\}} = \left\langle \psi_{\{i\}} \left| -\frac{\hbar^2}{2m^*}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right) \right| \psi_{\{j\}} \right\rangle + \left\langle \psi_{\{i\}} | V_{tot} | \psi_{\{j\}} \right\rangle \quad . \tag{12}$$

The first term can be evaluated as:

$$\mathcal{H}^{(1)}_{\{i\}\{j\}} = \frac{\hbar\omega_x}{4}\delta_{i_y,j_y}\left[(2j_x+1)\delta_{i_x,j_x} - \sqrt{j_x^2 - j_x}\,\delta_{i_x,j_x-2} - \sqrt{i_x^2 - i_x}\,\delta_{i_x-2,j_x}\right] + \frac{\hbar\omega_y}{4}\delta_{i_x,j_x}\left[(2j_y+1)\delta_{i_y,j_y} - \sqrt{j_y^2 - j_y}\,\delta_{i_y,j_y-2} - \sqrt{i_y^2 - i_y}\,\delta_{i_y-2,j_y}\right] \quad . \tag{13}$$

The second term depends on $V_{tot}(x, y, z)$. Let us suppose that this term reduces to a two-dimensional finite rectangular potential well $U(x, y)$, i.e.

$$U(x,y) = \begin{cases} 0 & \forall (x,y) \in [-L_x/2, L_x/2] \times [-L_y/2, L_y/2] \\ U_0 & \text{otherwise} \end{cases} \tag{14}$$

where the depth of the potential well is usually taken U_0 = -4.05 eV (the work function of Si) and L_x and L_y are the lengths of the well in x and y direction respectively. In this case, after some calculations, we have

$$\mathcal{H}^{(2)}_{\{i\}\{j\}} = \frac{U_0}{\pi}\sqrt{\frac{1}{2^{i_x+i_y+j_x+j_y} i_x! i_y! j_x! j_y!}} \exp\left[-\frac{1}{4}\left(aL_x^2 + bL_y^2\right)\right] \times \left\{\sum_{k=0}^{\min(i_x,j_x)} 2^k k! \binom{i_x}{k}\binom{j_x}{k}\left[H_{i_x+j_x-2k-1}(-\sqrt{a}L_x/2) - H_{i_x+j_x-2k-1}(\sqrt{a}L_x/2)\right]\right\} \times \left\{\sum_{k=0}^{\min(i_y,j_y)} 2^k k! \binom{i_y}{k}\binom{j_y}{k}\left[H_{i_y+j_y-2k-1}(-\sqrt{b}L_y/2) - H_{i_y+j_y-2k-1}(\sqrt{b}L_y/2)\right]\right\} \tag{15}$$

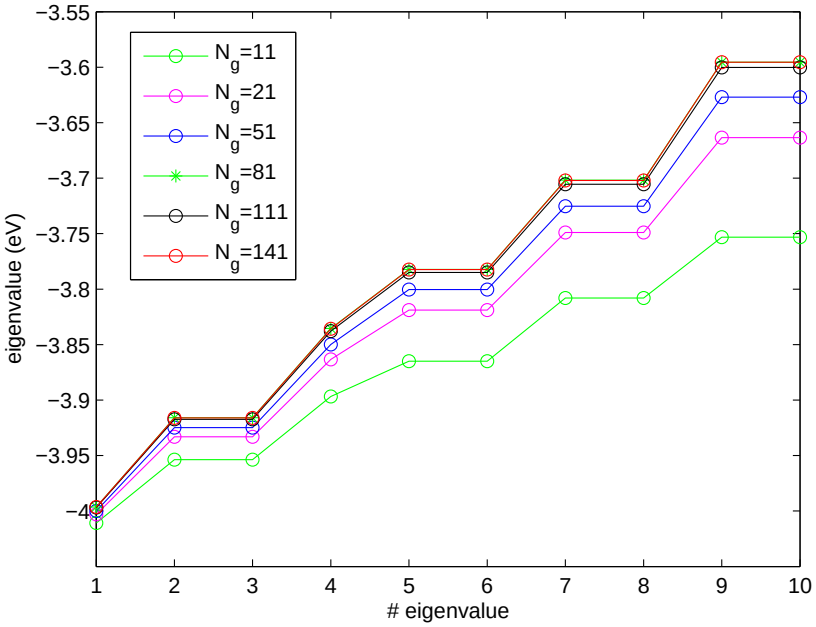


Figure 1: The subband energies ε_α (4) obtained with a finite difference solver, with N_g grid points in each axes.

If $j_x = 0, 1, ..., N_x$ and $j_y = 0, 1, ..., N_y$ then the matrix $\mathcal{H}$ has $(N_x + 1)(N_y + 1)$ rows and columns. The main advantage of the Variational Method is that no discretization error is introduced, as in the finite difference case, and the convergence of the results depends only on how many quantum numbers are taken. Just to have a comparison, we have taken a SiNW with cross-section $L_x = L_y = 10$ nm, and we have evaluated the eigenvalues with the two methods. In the figure 1 we plot the first ten eigenvalues ε_α obtained with the finite difference scheme, by increasing the grid points $N_g = N_1 = N_2$. From this figure we can clearly see that the convergence is reached for N_g= 141. In such a case, the matrix dimension is 19881 × 19881 and the CPU time was $\simeq$ 19000 sec (using MATLAB and an AMD Phenom II X6 1090T processor). If we want to obtain a similar result with the Variational Method, we have to take into account $N_x = N_y = 16$ quantum numbers. In this case the matrix dimension is 289 × 289 and the CPU time was $\simeq$ 163 sec. The percentage error between the two methods is shown in the figure 2. Similar errors are obtained for the envelope functions.

4 Ballistic transport

Once the quantum confinement in the transversal cross-section of the wire has been dealt with by efficiently solving the EMA equation, another issue is the description of the charge transport in the longitudinal direction of the nanowire. If the channel size of the device is far larger than the scattering length, carriers undergo a large number of

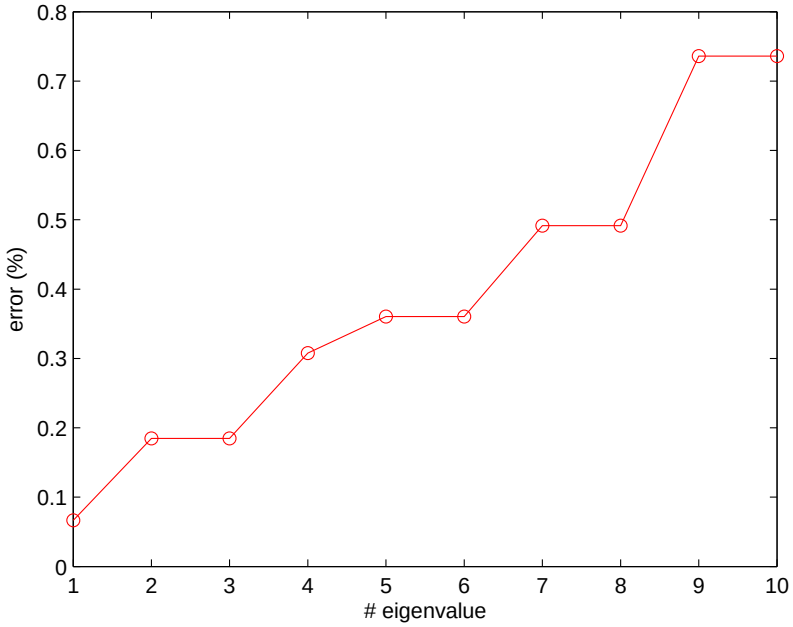


Figure 2: The percentage error between the subband energies ε_α (4) obtained with a finite difference solver and the Variational Method.

scattering processes which result in the diffusive carrier motion. If the channel device size is further decreased to less than the scattering length, ballistic carrier motion dominates the transport. The upper limit of the channel size to have a ballistic regime is an open question. Experimental results obtained in deeply scaled silicon MOSFETs [21] have shown that the carrier transport is near ballistic for channel sizes lesser than 50 nm.

In this paper we have considered the optimistic case of fully ballistic transport. The electrons are in equilibrium with their injecting reservoir (source or drain) which adsorbs and feed carriers into the channel without reflections. The charges are injected into the channel with unity probability, if their energy in the longitudinal direction is larger than the maximum subband energy $\varepsilon_{\alpha z}$. Then the carriers are transferred from the source to the drain without being scattered, and by neglecting tunneling effects. The source-to-drain current I can be evaluated with use of the method proposed in Landauer's formula [22], which is expressed as a sum of many one-dimensional subband components. Each subband current component flowing in one direction is given by the product of the unit charge, the number of carriers flowing into the subband per unit time, the transmission coefficient of the subband, all integrated over the carrier energy. The number of carriers flowing into the subband is further expressed by the product of the input carrier group velocity, the density of states, and the probability that the state is occupied by the carrier. In the not-degenerate case, the probability of carrier occupancy is given by the maxwellian distribution function with the source Fermi level on the source side, and that with the

drain Fermi level on the drain side of the subband. Both current directions, the one from the source to the drain and that in the opposite direction, should be considered. Then we have

$$I = q \sum_{\alpha} \int_0^{+\infty} vD(E_\alpha) \left\{ f_\alpha^{(eq)}(E_\alpha, E_{FS}) - f_\alpha^{(eq)}(E_\alpha, E_{FD}) \right\} T(E_\alpha)\, dE_\alpha \tag{16}$$

where $v = \hbar k_z / m^*$ is the group velocity, and $D_\alpha(E_\alpha)$ is the one-dimensional density of state

$$D(E_\alpha) = \frac{\sqrt{2m^*}}{\hbar\pi} \frac{\Theta(E_\alpha - \varepsilon_{\alpha z})}{\sqrt{E_\alpha - \varepsilon_{\alpha z}}} \quad . \tag{17}$$

$T(E_\alpha)$ is the subband transmission coefficient which, in the ballistic transport, implies

$$T(E_\alpha) = \begin{cases} 1 \quad \text{if} \quad E_\alpha \geq \varepsilon_\alpha^{Max} = \max_z \varepsilon_{\alpha z} \\ 0 \quad \text{otherwise} \end{cases} \tag{18}$$

In the not-degenerate case the equilibrium distribution function is the maxwellian, i.e.

$$f_\alpha^{(eq)}(E_\alpha, E_F) \propto \exp\left(-\frac{E_\alpha - E_F}{k_B T} \right) \tag{19}$$

where E_F is the Fermi energy and T the lattice temperature. E_{FS}, E_{FD} are the source and the drain Fermi energies respectively. If V_D and V_S are the applied voltage biases at the drain and source respectively, $V_{DS} = V_D - V_S$ then $E_{FD} = E_{FS} - qV_{DS}$.

The Fermi energy is obtained by imposing charge neutrality along each cross-section of the device, i.e.

$$\int\int n^{(eq)}(x, y, z)\, dx\, dy = \int\int (N_D - N_A)\, dx\, dy \tag{20}$$

and, if we suppose that the doping depends only on the longitudinal dimension z, we obtain

$$E_F(z) = -k_B T \log \left\{ \frac{\sqrt{2m^* k_B T}}{\pi\hbar \left[N_D(z) - N_A(z) \right] L_x L_y} \sum_{\alpha} \exp\left(-\frac{\varepsilon_{\alpha z}}{k_B T} \right) \right\} \quad . \tag{21}$$

The equilibrium linear electron density (7) is evaluated using eqs.17),(19),(21)

$$\rho_\alpha^{(eq)}(z) = \int D_\alpha(E_{\alpha z}) f_\alpha^{(eq)}\, dE_{\alpha z} = (N_D - N_A) L_x L_y \frac{\exp\left(-\frac{\varepsilon_{\alpha z}}{k_B T} \right)}{\sum_\alpha \exp\left(-\frac{\varepsilon_{\alpha z}}{k_B T} \right)} \quad . \tag{22}$$

Then by easy calculations eq.(16) reduces to

$$I = \frac{2q}{\hbar\pi} k_B T \sum_{\alpha} \exp\left(-\frac{\varepsilon_\alpha^{Max}}{k_B T} \right) \left\{ \exp\left(\frac{E_{FS}}{k_B T} \right) - \exp\left(\frac{E_{FD}}{k_B T} \right) \right\} \quad . \tag{23}$$

5 Simulation of a triple-gate SiNW transistor

We have considered a SiNW transistor with a parallelepiped shape, having square cross-section with dimension $L_x = L_y = 10$ nm and longitudinal dimension (i.e. the free transport direction) $L_z = 120$ nm. The device consists in an internal parallelepiped filled by Si (having dimension 8 x 8 nm^2) surrounded by SiO_2 of 1 nm thickness, producing the confining potential (14). The silicon, in the internal parallelepiped, is doped in the n^+ region with $N_D^+ = 10^{18}\,\mathrm{cm}^{-3}$ and in the n region with $N_D = 10^{16}\,\mathrm{cm}^{-3}$, with a regularization at the two junctions given by a hyperbolic tangent profile, i.e.

$$N_D(z) = N_D^+ - \frac{N_D^+ - N_D^-}{2}\left[\tanh\left(\frac{z-z_1}{s}\right) - \tanh\left(\frac{z-z_2}{s}\right)\right] \tag{24}$$

where $s = 5$ nm and $z_1 = 10$ nm, $z_1 = 80$ nm. The bottom of the parallelepiped (for z=0) is the source contact, and the top (for $z = 120$ nm) the drain contact. The triple-gate contact surrounds three faces of the parallelepiped with an extension $20 \le z \le 100$ nm. The applied biases are V_S= 0, V_D= 0.5 V and V_G is considered variable. The simulation procedure is the following:

- the EMA equation (4) is solved with $V_{tot} = U$ in each cross-section (z = const.), obtaining the subband energies $\varepsilon_{\alpha z}$ and envelope functions $\chi_{\alpha z}$.
- These functions are used to evaluate the charge density (6) in equilibrium condition, i.e. with ρ_α given by eq.(22).
- The Poisson equation (5) is solved leading to a new $V_{tot} = U - e\Phi$.
- If the new V_{tot} does not coincide sufficiently well with the old one, we start a further loop by solving the EMA equation with the new confining potential until the maximum difference between these two quantities is smaller than a predetermined tolerance.
- For each α-th subband, we evaluate $\varepsilon_\alpha^{Max} = \max_z \varepsilon_{\alpha z}$, and finally the current (23) is obtained.

We have taken into account only the first four subbands ($\alpha = 1, 2, 3, 4$) since, numerical experiments, shown that the other ones are very scarcely populated. The simulation results are shown in the figure 3. This characteristic curve, obtained in the ballistic approximation, gives us an upper limit to the the performance of this device, and the chosen approximations ensure a very good trade off between accuracy and computational resources required. Further extensions, such as including a more sophisticated transport model per subband according to the guidelines in [23]-[29], is under current investigations and will be presented in the next future.

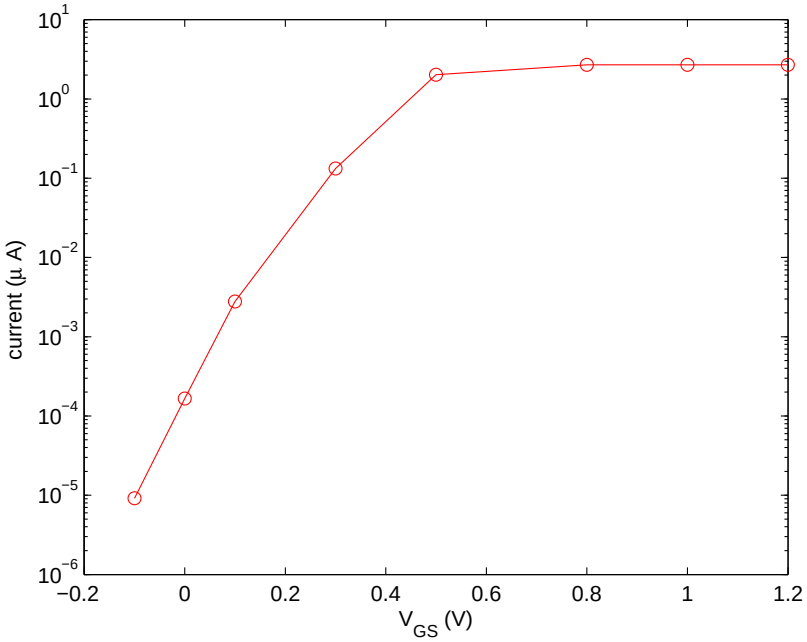


Figure 3: The current (23) versus the gate voltage V_{GS}.

Acknowledgment

We acknowledge the support of the Università degli Studi di Catania, FIR 2014 "Charge Transport in Graphene and Low dimensional Structures: modeling and simulation".

REFERENCES

[1] Ferry, D.K. Goodnick, S.M. and Bird, J. *Transport in nanostructures*. Cambridge University Press, (2009).

[2] Ramayya, E.B. Vasileska, D. Goodnick, S.M. and Knezevic, I. Electron mobility in Silicon Nanowires. *IEEE Trans. Nanotech.* (2007) **6**(1):113-117

[3] Ramayya, E.B. Vasileska, D. Goodnick, S.M. and Knezevic, I. Electron transport in silicon nanowires: The role of acoustic phonon confinement and surface roughness scattering. *J. Appl. Physics* (2008) **104**:063711.

[4] Ramayya, E.B. and Knezevic, I. Self-consistent Poisson-Schrödinger-Monte Carlo solver: electron mobility in silicon nanowires. *J. Comput. Electron.* (2010) **9**:206-210.

[5] Aksamija, Z. and Knezevic, I. Thermoelectric properties of silicon nanostructures, *J. Comput. Electron.* (2010) **9**: 173-179.

[6] Muscato, O. Monte Carlo evaluation of the transport coefficients in a n+ - n - n+ silicon diode. *COMPEL* (2000) **19**(3): 812-828.

[7] Muscato, O. and Wagner, W. Time step truncation in direct simulation Monte Carlo for semiconductors. *COMPEL* (2005) **24**(4): 1351-1366.

[8] Muscato, O. Wagner, W. and Di Stefano, V. Numerical study of the systematic error in Monte Carlo schemes for semiconductors. *ESAIM: M2AN* (2010) **44**(5): 1049-1068.

[9] Muscato, O. Wagner, W. and Di Stefano, V. Properties of the steady state distribution of electrons in semiconductors. *Kinetic and Related Models* (2011) **4**(3): 809-829.

[10] Muscato, O. Di Stefano, V. and Wagner, W. A variance-reduced electrothermal Monte Carlo method for semiconductor device simulation. *Computers & Mathematics with Applications* (2013) **65**(3):520-527.

[11] Lenzi, M.L. Palestri, P. Gnani, E. Gnudi A., Esseni, D. Selmi, L. and Baccarani, G. Investigation of the transport properties of silicon nanowires using deterministic and Monte Carlo approaches to the solution of the Boltzmann transport equation. *IEEE Trans. Electr. Dev.* (2008) **55**(8):2086-2096

[12] Ossig, G. and Schuerrer, F. Simulation of non-equilibrium electron transport in silicon quantum wires. *J. Comput. Electron.* (2008) **7**: 367-370.

[13] Majorana, A. Muscato, O. and Milazzo, C. Charge transport in 1D silicon devices via Monte Carlo simulation and Boltzmann-Poisson solver. *COMPEL* (2004) **23**(2): 410-425.

[14] Muscato O. and Di Stefano V., Hydrodynamic modeling of silicon quantum wires. *J. Comp. Electr.* (2012) **11**:45-55

[15] Muscato O. and Di Stefano V., Coupled quantum-classical transport in silicon nanowires. In: S. Idelsohn, M. Papadrakakis and B. Schrefler. Computational Methods for Coupled Problems in Science and Engineering V. p. 1109-1119, BARCELONA:CIMNE (Int. Center for Num. Meth. in Engineering), ISBN: 978-84-941407-6-1, Ibiza, 17-19 June 2013

[16] Muscato O. and Di Stefano V., Electro-thermal behaviour of a sub-micron silicon diode, *Semicond. Sci. Technol.* (2013) **28**: 025021

[17] Muscato O. and Di Stefano V., Hydrodynamic simulation of a n+ - n - n+ silicon nanowire, *Contin. Mech. Thermodyn.* (2014) **26**:197-205

[18] Muscato, O. Pidatella, R.M. and Fischetti, M.V. Monte Carlo and hydrodynamic simulation of a one dimensional $n^+ - n - n^+$ silicon diode. *VLSI Design* (1998) **6**(1-4): 247-250.

[19] Potter D., *Computational Physics*. Wiley, (1972).

[20] Suzuki Y. and Varga K., *Stochastic Variational Approach to Quantum-Mechanical Few Body Problems*. Springer, (1998).

[21] Lochtefeld A. and Antoniadis D.A., On Experimental Determination of Carrier Velocity in Deeply Scaled NMOS: How Close to the Thermal Limit?, *IEEE ELEC. DEV. LETTERS* (2001) **22**(2):95-97

[22] Natori K., Ballistic metal-oxide-semiconductor field effect transistor, *J. Appl. Phys.* (1994) **76**(8):4879-4890

[23] Muscato, O. and Di Stefano, V. Modeling heat generation in a sub-micrometric $n^+ - n - n^+$ silicon diode. *J. Appl. Phys.* (2008) **104**(12): 124501.

[24] Muscato, O. and Di Stefano, V. Hydrodynamic modeling of the electro-thermal transport in silicon semiconductors. *J. Phys. A: Math. Theor.* (2011) **44**(10): 105501.

[25] Muscato, O. and Di Stefano, V. An Energy Transport Model Describing Heat Generation and Conduction in Silicon Semiconductors. *J. Stat. Phys.* (2011) **144**(1): 171-197.

[26] Muscato, O. and Di Stefano, V. Local equilibrium and off-equilibrium thermoelectric effects in silicon semiconductors. *J. Appl. Phys.* (2011) **110**(9): 093706.

[27] Muscato, O. and Di Stefano, V. Heat generation and transport in nanoscale semiconductor devices via Monte Carlo and hydrodynamic simulations. *COMPEL* (2011) **30**(2): 519-537.

[28] Di Stefano, V. and Muscato, O. Seebeck Effect in Silicon Semiconductors. *Acta Appl. Math.* (2012) **122**(1): 225-238.

[29] Davoody, A.H. Ramayya, E.B. Maurer, L.N. and Knezevic, I. Ultrathin GaN nanowires: Electronic, thermal, and thermoelectric properties. *Phys. Rev. B* (2014) **89**(11):115313

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis (Eds)

COUPLED PROBLEMS OF LOCALIZATION OF DUST AND GAS

O. A. AVERKOVA*, K. I. LOGACHEV*, I. N. LOGACHEV* AND A. E. CANAR*

* Belgorod State Technological University named after V.G. Shukhov
(BSTU named after V.G. Shukhov), 308012 Belgorod, Russia
e-mail: kilogachev@mail.ru, web page: http://www.bstu.ru

Key words: dedusting, granular materials, exhaust ventilation

Abstract. The paper outlines the developed mathematical models, computational algorithms forcalculation of dust and gas streams in the spectrum of action of local ventilation device of the closed type (aspiration shelter) from nodes overload of granular materials.

1 INTRODUCTION

The most reliable way of localization and capturing of dust and gas emissions in the production premises of industrial enterprises is the use of dedusting ventilation systems (aspiration), main element of which is local ventilation exhaust (LVE).Performance of dedusting ventilation system should reduce particulate air pollution to the level of maximum permissible concentration at a minimum flow of air entering in LVE.

The aim of this work is to develop methods for the calculation of the velocity field in the spectra of actions of LVE, which allowstaking into account: influence of the rotating elements of the technological equipment on the velocity distribution of airflow; vortex structures, spreading in the closed LVE - aspiration shelters; separated flows, generated at the entrance to the leaks of aspiration shelters and exhaust openings; distribution of dust aerosols in the obtained velocity field of airflows.

2 METHOD OF REDUCING VOLUME OF EJECTED AIR

A mathematical model of ejection of air in a circular perforated pipe was developed for determining the effect of the flow of granular material on the air environment and the effectiveness of exhaust cover, equipped with a bypass chamber.

It was considered axisymmetric particle flow in a circular pipe with cross-sectional area $\tilde{s}_t$, m^2 (Fig.1-2). A cylindrical bypass-chamber is provided around the tube (with cross-sectional area $\tilde{s}_b$, m^2), aerodynamically connected with a tube with perforated wall.

Because of overflowing of air from pipe to the bypass chamber, speed of ejected air in the pipe ($\tilde{u}$, m/s) and speed of the upward flow in the bypass chamber ($\tilde{\omega}$, m/s) are changing along the length of the channel.

To determine these velocities we used the equation of conservation of momentum of the air in a fixed volume $\tilde{V}\,(\mathrm{m}^3)$, bounded by the surface $\tilde{S}(\mathrm{m}^2)$: $\int_{\tilde{S}} \tilde{\rho}\vec{\tilde{u}}\,\tilde{u}_n d\tilde{S} = \int_{\tilde{V}} \vec{\tilde{M}} d\tilde{V} + \int_{\tilde{S}} \vec{\tilde{p}}_n d\tilde{S}$, where $\tilde{u}_n$ -

projection of the vector the air velocity on the external $\vec{n}$ surface normal $\tilde{S}$, м/с; $\vec{\tilde{u}}$ - velocity vector of air, м/с; $\tilde{\rho}$ - density of air kg/m^3; $\vec{\tilde{M}}$ - vector of mass forces, N/m^3; $\vec{\tilde{p}}_n$ - vector of surface forces, applied to an elementary area $d\tilde{S}$ with an external normal $\vec{n}$, Pa.

Whence, neglecting small values of higher order, we have come to the basic equation of one-dimensional problems of dynamics of ejected air in the pipe in a dimensionless form:

$$dp + 4udu = \mathrm{Le}(v-u)|v-u|/v \cdot dx, \tag{1}$$

where $\mathrm{Le} = \psi\beta_k \tilde{l}\tilde{F}_м / \tilde{V}_ч$; $\beta_k = \tilde{G}_ч / (\tilde{\rho}_ч \tilde{v}_k \tilde{s}_t)$; $p = 2\tilde{p}/(\tilde{\rho}\tilde{v}_k^2)$; $u = \tilde{u}/\tilde{v}_k$; $v = \tilde{v}/\tilde{v}_k$; $x = \tilde{x}/\tilde{l}$; $\tilde{p}$ - static pressure in pipe, Pa (hereinafter we are talking about excessive static pressure); $\tilde{v}_k$ - velocity of particles at the end of pipe, m/s; $\tilde{l}$ - total length of pipe, m; dimensionless number Le («ejection parameter») is the ratio of maximum ejection pressure forces (at $\tilde{v}-\tilde{u}=\tilde{v}_k$) and dynamic pressure of the ejected air.

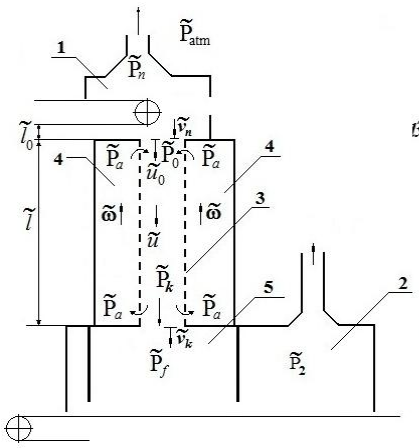


Figure 1: The scheme of aspiration of handling node, equipped by upper (1) and lower (2) aspirated shelters, a perforated chute (3) with the bypass chamber (4) and form camera (5) in the lower shelter

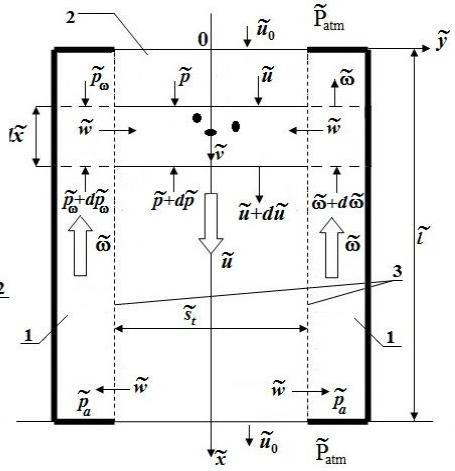


Figure 2: An exemplary scheme of perforated chute with the bypass chamber: 1 – bypass chamber; 2 – a pipe;3 – perforated walls of chute

Similarly, we obtained an equation of dynamic of rising air in the bypass channel with cross-sectional area $\tilde{S}_b$ (in the absence of aerodynamic resistance forces of channel walls):

$$dp_\omega + 4\omega d\omega = 0; \quad p_\omega = 2\tilde{p}_\omega/(\tilde{\rho}\tilde{v}_k^2); \quad \omega = \tilde{\omega}/\tilde{v}_k, \tag{2}$$

where p_ω - dimensionless static pressure in transverse cross sections of the bypass channel.

After a number of transformations, we have a system of differential equations:

$$du/dx = E \cdot \gamma\sqrt{|\triangle p|}; \qquad dp/dx = -4uE\gamma\sqrt{|\triangle p|} + \mathrm{Le}(v-u)|v-u|/v,$$

where parameters are calculated by formulas

$\triangle p = p_\omega - p = p_a - 2\left((u-u_0)/r\right)^2 - p$; $E = \hat{S}_t/\sqrt{\zeta_0} = \tilde{I}\tilde{l}\varepsilon/(\tilde{s}_t\sqrt{\zeta_0})$; $v = \sqrt{(1-n^2)x+n^2}$; $n = \tilde{v}_0/\tilde{v}_k$; γ – mark of pressure difference $(p_\omega - p)$;

$\tilde{v}_0$ - flow rate of particles at the inlet of the pipe; $\hat{S}_t$ - the ratio of the total area of perforation holes in the walls of tube to its cross-sectional area; $\tilde{I}\tilde{I}$ - perimeter of pipe; ε - degree of perforation of the walls of the pipe ($\varepsilon = 0$ in the absence of perforation, $\varepsilon = 1$ in the absence of the pipe walls); ζ_0 - coefficient of local resistance (c.l.r.) of perforation hole.

We use the following boundary conditions:

$$u(0) = u_0; \quad p(0) = -\zeta_n u_0^2; \qquad (3) \qquad u(1) = u_0; \quad p(1) = \zeta_k u_0^2, \qquad (4)$$

where ζ_n, ζ_k - coefficients of local resistance to movement of air at the inlet of pipe and outlet of pipe. Some difficulties arise in connection with the fact that the desired data are exactly the velocity of the air at the entrance to the pipe (at the outlet of pipe) - u_0 and overpressure at the beginning (at the end) of the bypass channel - p_a.

Therefore, the solution of the boundary value problem is solved by the method of random search:specifying of u_0, is solved the Augustin-Louis Cauchy problem with initial conditions (3), wherein a predetermined value of p_a from the condition $p(1) > p_a > p(0)$, $p_a \approx (\zeta_k - \zeta_n)u_0^2 / 2$ and then verifies the condition (4). To facilitate the search of values u_0 and p_a you can use the bisection method (method of half division).

The increase in the number of ejection Le, as in the case of flow in a pipe with impermeable walls, promotes growth of u_0. The asymptotic of the growth (fig.3) and decline in ejection are notable in compare with the flux of particles in the non-perforated pipe ($u_0 < u_2$ - dash-dotted curve) in the investigated range of numbers Le.

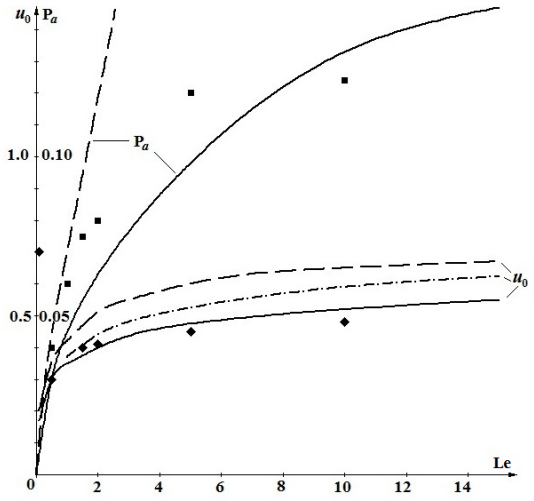


Figure 3: Dependence of ejected air velocity u_0 and pressure in the bypass chamber p_a from the number of Le (at $E = 1$; $r = 1$; $n = 0.4319$; $\zeta_n = 0.5$; $\zeta_k = 1$; $\zeta_0 = 1.5$) ;solid curves are constructed according to the formulas of the linearized problem; dashed line - at maximum forces $k = 1$, Le$(v-u)$; dash-dotted line is for the case of pipes with impermeable walls ($E = 0$); rhombuses ◆ - for u_0 and squares ▪ - for p_a according to the results of the numerical solution of «exact» equation

An even greater effect of minimizing u_0 is observed when increasing the degree of perforation and there is asymptotic - when $E \geq 2.5$ the decrease of u_0 practically stops.

The decrease volumes (speed u_0) of ejected air due to its recirculation in the bypass channel is higher, the greater the number of Le and less than n.

The decrease in ejection is even more noticeable with increasing aerodynamic resistance at the exit and entrance of air into the perforated tube. This fact confirms the significant role on the volume of recirculated air in the conventional (non-perforated) bypass channel pressurization of the top shelter and the presence of the buffer capacity (form-chamber) in the lower cover, which create a greater vacuum at the top of the pipe and an afflux - at the bottom. This creates conditions for a more intense overflow of air through the perforation holes. Volumes of ejected air can be reduced by 1.3-1.5 times by increasing ζ_k to 8-32 and 1.45-1.55 times with increasing of ζ_n to 4-16 (if $\mathrm{Le} \geq 10;\quad r = 2;\quad E = 2.5$).

3 REDUCTION OF THE VOLUME OF AIR ENTERING THROUGH LEAKAGES

Initially, we considered a potential separation of airflow at the entrance to flat slot, in front of which there are two screens, perpendicular to its axis and investigated separated flow in the slit-like intake channel extending beyond the flat wall, in the spectrum of action of which there is an impervious screen. Then, using the theory of functions of a complex variable was able to solve the more general problem of separated flow in a horizontal channel, the inlet section of which protrudes from the vertical wall at a distance S (Fig. 4).

The flow is limited by impenetrable screen, remote at a distance G from the inlet section of flat channel with height 2B, and divided at the distance M by vertical shield (screen with a central hole in height of 2R) into two regions. An impenetrable screen in this case allows simplifying the task of determining the constants of the integral of the Christoffel-Schwarz, in particular to estimate the magnitude of increments at the infinite point A.

The function of Zhukovsky in this case has the form: $\omega = \frac{1}{2}\ln\frac{\sqrt{t}+\sqrt{b}}{\sqrt{t}-\sqrt{b}} + \frac{1}{2}\ln\frac{\sqrt{t}+\sqrt{p}}{\sqrt{t}-\sqrt{p}}$, and for complex potential, we obtained an expression: $w = \frac{h-m}{1-m}\cdot\frac{q}{\pi}\cdot\ln\left(\frac{t-m}{m}\right) + \frac{1-h}{1-m}\cdot\frac{q}{\pi}\cdot\ln(t-1)$.

A connection of points of auxiliary plane t and points of the physical region defined by the formula:

$$z = i + (1-k)\frac{\delta_\infty}{\pi}\cdot A + k\frac{\delta_\infty}{\pi}\cdot B \text{, where } A = \int_0^t e^{\omega}\frac{dt}{t-m};\ B = \int_0^t e^{\omega}\frac{dt}{t-1}.$$

Shield with a central aperture divides intake torch into two parts, wherein, despite low consumption of closest to cut-off part of hole, its velocity is higher than the other parts, which increases the inertial preload of jet to axis of the channel and, as a consequence, reduces the thickness of the jet δ_∞.

On the value of δ_∞ is influenced as location of shield M, as magnitude of the hole R. Influence of value δ_∞ on the coefficient resistance of input environment ζ in aspirating hole is determined by the formula of Idelchik I.E .: $\zeta = \left(1/(\delta_\infty)^m - 1\right)^2$, where $m = 1$ for slit intake canal and $m = 2$ for round pipe.

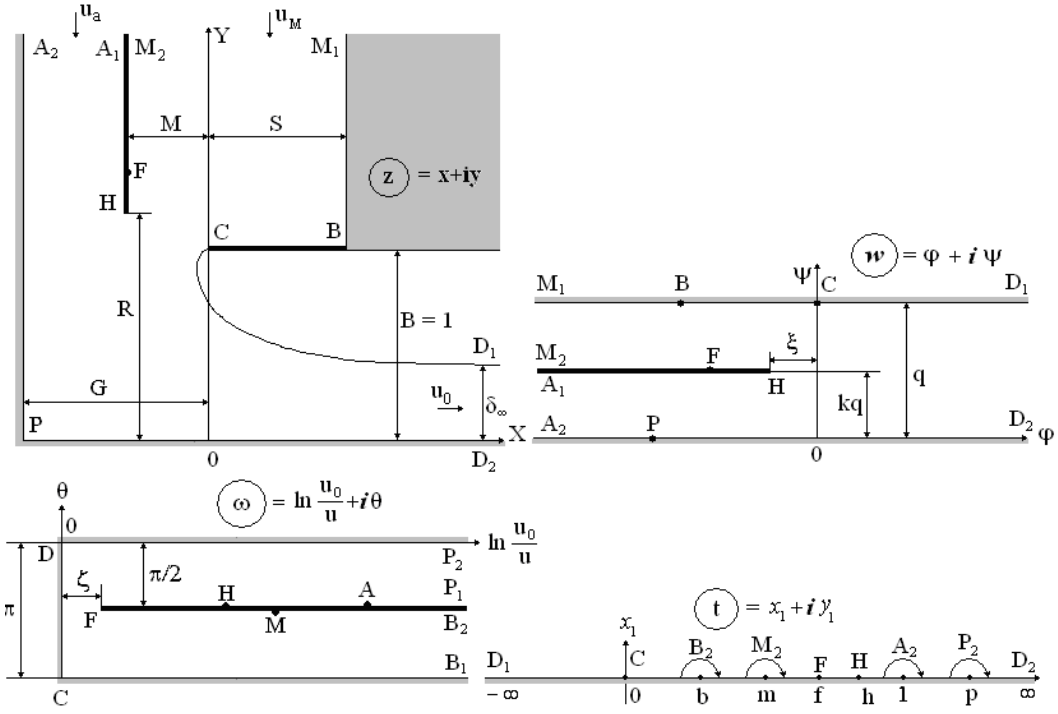


Figure 4: To the definition of an orthogonal grid and velocity field of suction channel of protruding torch with screen

As research has shown, the change in relative thickness of the jet $\bar{\delta} = \delta_\infty(1)/\delta_\infty(\infty)$ and the relative coefficient of resistance $\bar{\zeta} = \zeta(1)/\zeta(\infty)$ (here $\zeta(1) = \zeta|_{R=1}$; $\zeta(\infty) = \zeta|_{R\to\infty}$) have clearly defined extrema while removing the shield at $M \approx 0.75$ not only at a considerable distance of blank wall, but also with a noticeable approximation of this wall to this aspiration hole. Thus, value $\bar{\delta}$ has a minimum, and $\bar{\zeta}$ - the maximum.

The use methods of theory of functions of a complex variable and Zhukovsky`s method allows accurately determine parameters of flow separation, but does not give the opportunity to explore the flow in plane multiply connected regions containing screens of finite length as well as the flows near the round suction channels.

Therefore, there was a need to develop a method in which these disadvantages have been eliminated.

We considered multiply connected flow region (fig.5 a) at the entrance into a flat (or round) intake channel, in the spectrum of action of which there is thin screen (a circular disc with a central opening), at its circulation flow around. Flow separation occurs and forms a free current line from a sharp edge *C*.

A numerical procedure was developed to determine its position, velocity of flow at any given point and the coefficient of local resistance (c.l.r.) at the entrance to the inlet hole

Discrete mathematical model for a plane problem is constructed as follows (fig.5 b). Let us denote: N is number of attached vortices, the same number will be the control points. A vortex lying on the sharp edge of visor C is considered to be free. Attached vortices were

located at the points of fracture boundaries. Between the attached vortices were located control points.

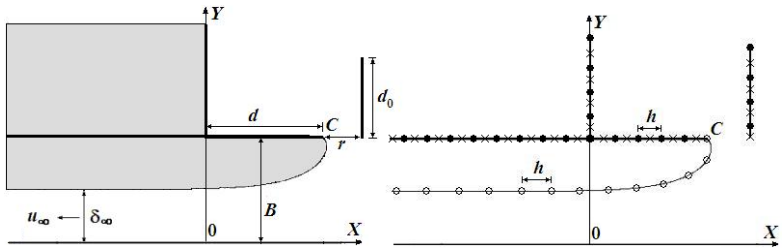


Figure 5:The statement of the problem::a) physical region of flow; b) discretization of the boundary of area (● −attached vortices, ○ −free vortices, × −control points)

Point of $\xi^k(\xi_1,\xi_2)$ − is point location of k-th attached vortex; $x^p(x_1,x_2)$ − *p*-th control point. To ensure impermeability of axis OX we reflect symmetrically with respect to it all the vortices; circulations of symmetric vortices should be opposed.Compliance with this condition is automatically leads to a condition of non-circulatory flow. It was assumed that desired free line of vortices circulation is constant and equal to γ. The total impact of all jets on a control point x^p in the direction of the outward normal is expressed by the equation:

$$v_n(x)=\sum_{q=1}^{N}\left(G\left(x,\xi^q\right)-G\left(x,\xi^{q+N}\right)\right)\Gamma\left(\xi^q\right)+\gamma\sum_{k=1}^{N_S}\left(G\left(x,\zeta^k\right)-G\left(x,\zeta^{k+N_S}\right)\right), \tag{5}$$

where ζ^k –point location of the free vortex. Function $G(x,\xi)=\left((x_1-\xi_1)n_2-(x_2-\xi_2)n_1\right)/\left(2\pi\left((x_1-\xi_1)^2+(x_2-\xi_2)^2\right)\right)$ expresses by itself the velocity caused at the point $x(x_1,x_2)$ along a predetermined direction $\vec{n}=\{n_1,n_2\}$ of a single vortex, located at the point $\xi(\xi_1,\xi_2)$. Since $v_n\left(x^p\right)=0$ in all control points,i. e. condition of impermeability is performed, at change of $p=1,2,\ldots,N$ the expression (5) is transformed into a system of linear algebraic equations for determining the unknown circulations $\Gamma\left(\xi^q\right)$ of attached vortices.

The second approximation for the free streamline is constructed using the Runge-Kutta method for the numerical solution of systems of ordinary differential equations $dx/dt=v_x;\ dy/dt=v_y$.

Line current begins to build with sharp edges C. As soon as the distance between the point (x,y) and sharp edge becomes *h*, then a free vortex comes in this point, i.e. it will be the second approximation for this point free line current. Then again, the current line is built, while again the distance between the point (x,y) and the previous position of the free vortex will be h.A free vortex is placed at this point and so on.

After determining of the second approximation for the free streamline is required to re-solve the system of equations (5) and to determine the circulation of attached vortices. Then, the third approximation of free streamline is being built and etc. This iterative process continues until the distance between the subsequent position of N_S-th free vortex and the

previous one will be no more than a given accuracy ε. Discrete mathematical model is constructed in a similar manner for the axisymmetric problem.

The infinitely thin vortex rings were used as discrete features, without self-induction. The system of equations for determining the unknown intensities of attached vortex rings will take the form:

$$\sum_{q=1}^{N} G\left(x^p, \xi^q\right) \Gamma\left(\xi^q\right) = -\gamma \sum_{k=1}^{N_S} G\left(x^p, \zeta^k\right), \tag{6}$$

and speed is determined by the formula:

$$v_n(x) = \sum_{q=1}^{N} G\left(x\ , \xi^q\right) \Gamma\left(\xi^q\right) + \gamma \sum_{k=1}^{N_S} G\left(x^p, \zeta^k\right), \tag{7}$$

where $G(x,\xi) = \dfrac{(A_1 b + A_2 a)}{b} \cdot \dfrac{4}{(a-b)\sqrt{a+b}} E(t) - \dfrac{A_2}{b} \cdot \dfrac{4}{\sqrt{a+b}} F(t)$ при $b \neq 0$, $A_1 = \dfrac{{\xi_2}^2 n_1}{4\pi}$, $2x_2\xi_2 = b > 0$,

$t = 2b/(a+b)$, $G(x,\xi) = \dfrac{\xi_2^2 n_1}{2a\sqrt{a}}$ при $b = 0$, $a = (x_1 - \xi_1)^2 + {\xi_2}^2 + {x_2}^2 > 0$, $E(t) = \int_0^{\pi/2} \sqrt{1 - t^2 \sin^2\theta}\, d\theta$;

$F(t) = \int_0^{\pi/2} \dfrac{d\theta}{\sqrt{1 - t^2 \sin^2\theta}}$, $F(t) = \sum_{i=0}^{4} c_i (1-t)^i + \sum_{i=0}^{4} d_i (1-t)^i \ln\dfrac{1}{1-t}$,

$E(t) = 1 + \sum_{i=1}^{4} c_i (1-t)^i + \sum_{i=1}^{4} d_i (1-t)^i \ln\dfrac{1}{1-t}$, c_i, d_i formula are taken from the tables of special functions.

We considered various modes of flow around of vertical screen that is contained in the spectrum of the inlet channel. The circulation around the screen was the closest to the experimental data with the condition of finiteness of the speed on the lower edge. (Fig.6).

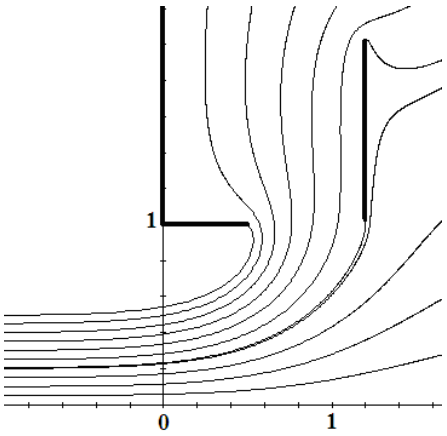


Figure 6: Line current at circulation flow around screen (δ_∞ = 0.4957; r = 0.7)

For a fixed value of length d_0 of the vertical screen and the change of distance r there is a minimum value of δ_∞. In particular, the value of δ_∞ for the plane problem for different values of d_0 has a minimum in the range $0{,}55 < r < 0{,}75$, and for axisymmetric one – $0{,}3 < r < 0{,}3$. With a significant increase of d_0 for a plane problem, extremum occurs when $r = 0{,}75$, that

corresponds with calculations according to the method of N. E. Zhukovsky at $d_0 \to \infty$.

Numerical experiments have shown that increasing the length of the profile of more than one caliber does not give significant practical effect.

4 APPLICATION OF SWIRLED AIRFLOWS

In this paragraph is being developed a method of modeling dusty flows in the spectrum of action of LVE from rotating cylindrical parts and offered a new direction to reduce the dust discharge from aspirated shelters due to the use of the properties of swirling flows. Initially we developed a method of computer simulation based on the method of boundary integral equations with plane jets simulating rotating cylinders.

There were simulated dusty flows near open LVE, which shows the necessity of taking into account the rotation of the cylinders, which have a significant effect on the velocity field of the airflows and the value of the maximum diameter $d_{\max}$ of dust particles captured by LVE.

It decreases by 3-4 times. Thus, the accuracy of the model significantly influences on correct choice of an efficient dust collecting apparatus. It is shown that significant energy-saving effect and increase the effectiveness of LVE can be achieved by installing a mechanical screen.

In local exhausts of closed type - aspiration shelters can also be used rotating cylinders to reduce their energy consumption.

Using the developed computer program, we determined the value of $d_{\max}$ finely dispersed aerosols of different physical properties, that being carried into aspiration network with the aim to develop constructive proposals for designing of aspiration shelters with the functions of dust precipitation chamber.

Computational experiments have shown that reduction of 140-200 μm of the maximum diameter of dust particles can be achieved due to equipping of aspiration shelters with screening visors, rotating cylinders and exhaust-cylinder, conducing due to the airflow induced by their rotation of precipitation of dust particles on the bottom of cover. To study the processes of vortex is developed a mathematical model of vortex flows inside the shelter, in which may be located rotating cylinders, its computer implementation.

We considered the region with the boundary of S, in which is set the normal component of the velocity.

In the area may be located the rotors and exhaust-cylinders with known linear speeds v_k, $k = 1,2,\ldots,M$. To simulate the boundary S is used a simple layer: continuously distributed fictitious sources with intensities $q(\xi)$. To account the influence of the airflow, initiated by rotation of a cylinder of radius of R_k, locate in their centers c_k plane vortices with known quantities of circulations $\Gamma_k = 2\pi R_k v_\kappa$.

At the time of $t = m \cdot \Delta t$ the system for determining the unknown circulation of attached vortices and intensities of sources (runoffs) will have the form:

$$\begin{cases} -0,5q(x^p) + \sum\limits_{\substack{k=N+1,\\ k\neq p}}^{N+W} q(\xi^k) \int\limits_{\Delta S^k} F_2(x^p,\xi^k)dS(\xi^k) + \sum\limits_{k=1}^{N} G(x^p,\xi^k)\Gamma(\xi^k) + \Lambda = \\ = v_n(x^p) - 2\pi\sum\limits_{k=1}^{M} R_k v_{ki} G(x^p,c_k) - \sum\limits_{\tau=1}^{m}\sum\limits_{l=1}^{L} G(x^p,\zeta^{l\tau})\gamma^{l\tau}, \quad \sum\limits_{k=1}^{N}\Gamma(\xi^k) + \sum\limits_{\tau=1}^{m}\sum\limits_{l=1}^{L}\gamma^{l\tau} = 0, \end{cases}$$

where function of $F_2(x,\xi) = \frac{1}{2\pi}\frac{(x_1-\xi_1)n_1 + (x_2-\xi_2)n_2}{(x_1-\xi_1)^2 + (x_2-\xi_2)^2}$ expresses by itself an impact on the point $x(x_1,x_2)$ a single source located at the point $\xi(\xi_1,\xi_2)$ along the unit vector $\vec{n} = \{n_1, n_2\}$; $v_n(x^p)$ - speed in the direction of the external normal to the boundary of the region in the p-th control point at $p = 1,2,\ldots,N$ orthe middle of the p-th segment at $p = N+1, N+2,\ldots,N+W$; Λ - regularizing variable of I. K.Lifanova; $\gamma^{l\tau}$ - circulation of free vortex descended from the l-th sharp edge at time $t = \tau\cdot\Delta t$; W - number of boundary elements with a simple layer; N - number of attached vortices.

The speed is determined from the expression:

$$v_n(x) = \sum_{k=N+1}^{N+W} q(\xi^k) \int_{\Delta S^k} F_2(x,\xi^k)dS(\xi^k) + \sum_{k=1}^{N} G(x,\xi^k)\Gamma(\xi^k) + 2\pi\sum_{k=1}^{M} R_k v_k G(x,c_k) + \sum_{\tau=1}^{m}\sum_{l=1}^{L} G(x^p,\zeta^{l\tau})\gamma^{l\tau}$$

The trajectory of the dust particle is constructed based on the integration of the equation of its movement:

$$\frac{1}{6}\rho_p\pi d_e^3 v_p'(t) = -\psi\left|\vec{v}_p - \vec{v}_a\right|\left(\vec{v}_p - \vec{v}_a\right)\rho_a\chi S_m / 2 + \frac{1}{6}\rho_p\pi d_e^3\vec{g},$$

where $\vec{v}_a$ – air speed; ρ_a – air density; $\vec{v}_p$ – particle velocity; ρ_p – density of the particle; d_e– equivalent diameter of the particles; $\vec{g}$ – acceleration of gravity; $S_m = \pi d_e^2/4$ – the square of middle section of the particle; χ– coefficient of dynamic form; ψ–the coefficient of air resistance, calculated by the formulas of Stokes, Klyachko, Adam.

With the help of developed set of numerical algorithms, computer program was identified patterns of behavior of dust aerosols, arising during overloads of granular materials in aspiration shelter of standard design, depending on the availability of non-stationary vortex structures in the flow domain.

A set of particles of different sizes came in area of shelter from supply opening. We can enter dispersed composition and concentration of dust and performed simulation of its movement until their complete precipitation or capturing. Total mass of particles and concentration in the aspirated air have been determined, that were caught in the extraction, equal to the ratio of this mass to the volume of air in which they were held. The particle disperse structure in the suction pipe was determined by counting particles of different fractions, captured by suction.

As an example, it was implemented modeling of movement of dust particles (Fig. 7) in an aspiration shelter of node overload clinker on conveyor. It was considered the movement of 30,000 dust particles with a density of 3050 kg/м^3 with dynamic coefficient shape of 1.8 (sharp-grained particles).

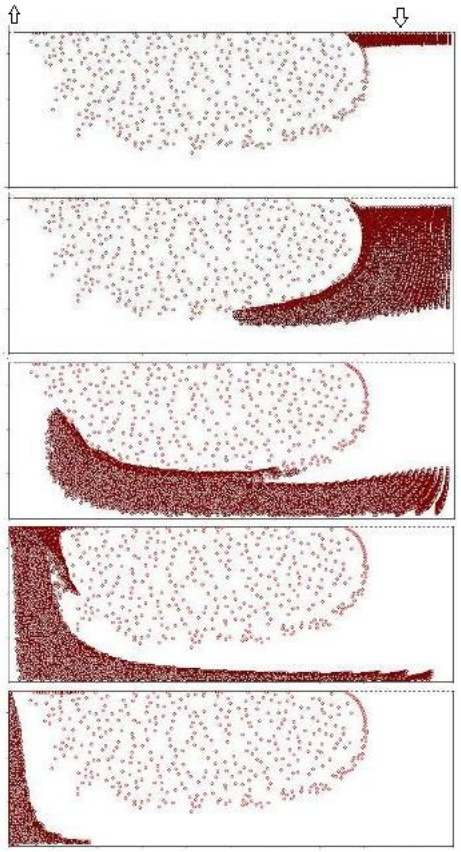


Figure7: Movement of dust cloud of 30,000 particles of various fractions in aspiration shelter node overload of clinker conveyor

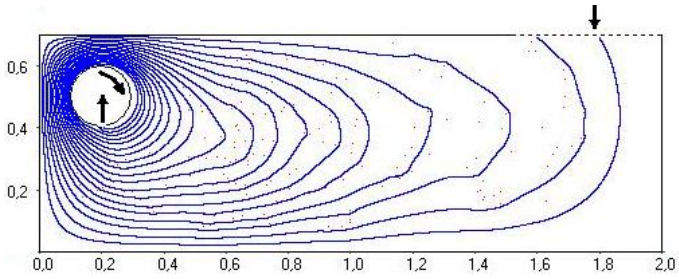


Figure 8: Swirledairfield

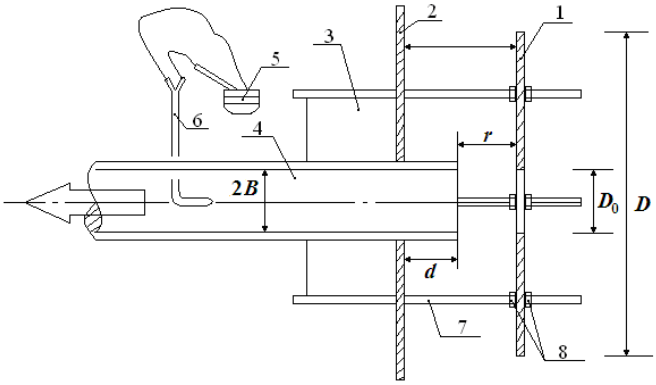


Figure 9: Scheme of experimental setup for determining resistance at the entrance of air into the screened round hole: 1 –screen with a central hole; 2 –shield; 3 –directing triangular prism; 4 –pipe; 5 –micromanometer with inclined tube; 6 –pneumometric tube of Pitot-Prandtl; 7 – steel rods-studs; 8 – screw-nuts for fixing screen

Replacement of exhaust hole on rotating cylinder-suction significantly changes the aerodynamics inside the shelter (fig.8).

Numerical experiments have shown that concentration of dust in intake air can be reduced from maximum permissible up to zero. It depends on speed of rotation of exhaust-cylinder. It is obvious that it should be the most removed from loading chute and conveyor belt. The direction of rotation should contribute to deposition of dust.

5 COMPARISON WITH THE EXPERIMENT

We have developed an experimental device, which is illustrated in fig.9, for study the flow separation at the entrance to the round aspiration channels. Comparison of calculated and obtained experimental values of coefficient of local resistance ζ (fig.10) demonstrates their satisfactory matching. The overestimation of calculated values of ζ, not more than 15%, and it is observed for small values of *r*.

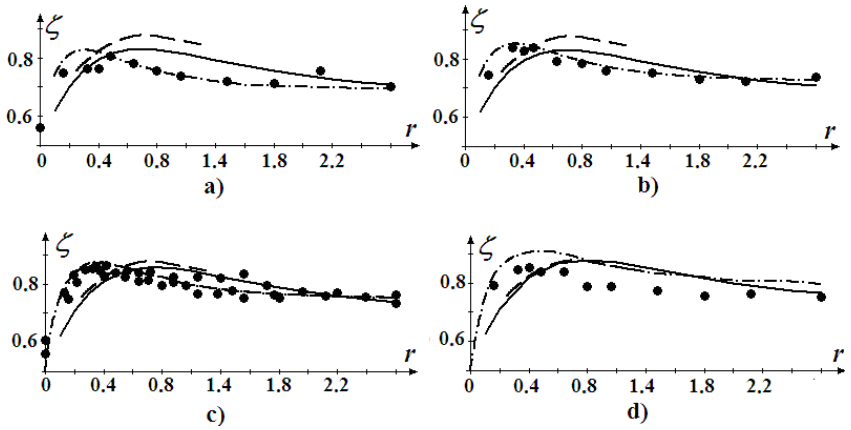


Figure 10: Comparison of calculated values of coefficient of local resistanceζ from a distance$r$ when $d_0 = 1.55$: a) $d = 0.24$; b) $d = 0.48$; c) $d = 0.56$; d) $d = 1.2$ (solid curve –calculations for a plane problem,dash-dotted curve –for axisymmetric one, dashed curve –method of N. E. Zhukovsky, circles – an experiment)

Value of ζ is the closest to experimental data, that has been found for the problem in axisymmetric formulation, which is naturally, because full-scale experiment was placed in the same formulation. Calculated and experimental extrema of coefficient of local resistance (fig. 10) are the same, which enables to conclude about the reliability of developed method of mathematical modeling and research results as for the plane problem (i.e. of separated flow at the inlet of rectangular aspiration aperture with an aspect ratio of not less than 1 to 10), as axisymmetric one.

6 CONCLUSIONS

1. For improving the efficiency of dedusting ventilation systems is necessary a package of measures for reducing:volume of ejected air, air consumption, entering through leaks and technological openings of aspirating shelters, dust discharge in aspiration network by using of properties of vortex, separate and recirculating flows.

2. A method of reducing the volume of ejected air during overloads of granular materials was developed. The effectiveness of creating a load device is analytically proved in the form of a vertical perforated pipe with a bypass chamber when loading granular materials. The process of recirculation of air in the system "boot perforated pipe - bypass camera" was investigated, determined the influence of their geometric and aerodynamic characteristics on volume of ejected air and performance of the aspiration system.

By calculation, it was shown that the performance of dedusting ventilation system might be reduced by at least 40% due to closed circulation of dusty flows without additional boosters of traction.

3. A non-contact method of reducing volume of air was developed, entering through leaks and technological openings of aspiration shelters for localization dust emissions, based on the effect of separation of airflow from sharp edges of mechanical screens.

We proposed and patented a method of controlling flow separation at the inlet to aspiration channels. We have developed mathematical models of separated flows at the entrance of round and slotted aspiration channels in the range of which can be arranged end-thin profiles.

We investigated different regimes of flow profiles, determined their geometrical dimensions and conditions, contributing to high aerodynamic resistance of aspiration channels.

We developed recommendations to reduce the volume of air entering through leaks of aspiration shelters. Calculated data experimentally are shown that airflow is possible to reduce by at least 20%, thereby increasing the efficiency of the aspiration system by reducing energy consumption.

4. A method for reducing dust discharge from aspiration shelters was developed, based on the use of swirling air currents, initiated by rotating exhaust-cylinders.

We proposed program-algorithmic support for research of processes dust discharge in aspiration network from localizing dust emission devices and investigated the dynamics of dust aerosols in aspirated flows, containing mechanical screens and rotating cylinders.

Based on computational experiments, we have investigated processes: influence of rotation cylindrical detail on airflow in the spectrum of absorption LVE open type and dynamics of dust particles in this flow; capturing of dust particles of different shape and density at changing ratio velocity of adsorption to speed of rotation of cylindrical part.

We identified regularities of the influence of thin impermeable screens, rotating exhausts-cylinders and cylinders on entrainment of dust in the aspiration network. It has been shown, that a significant reduction in dust discharge in aspiration network up to complete precipitation of dust on the conveyor belt can be achieved by using a rotating exhaust-cylinder.

5. Performed experimental research on installation of aspiration shelter with slotted leakages and stands for study of flow separation in the inlet to the intake of circular and slit-like shape channels, allowed to confirm the data obtained by analytical and computational experiment. We established qualitative and quantitative concurrence of the velocity field in the spectra of absorption and regularities of changes of the coefficient of local resistance to the entrance of the intake channels.

The reported study was partly supported by RFBR, research project No. 14-41-08005r_ofi_m and the President of the Russian Federation, project MK-103.2014.1.

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis(Eds)

DROP COALESCENCE SIMULATIONS USING LEVEL SETS COUPLED WITH BOUNDARY INTEGRAL METHODS

M. Garzon*, L. J. Gray† AND J. A. Sethian††

* Dept. of Applied Mathematics
Universidad de Oviedo
Calvo Sotelo s/n, 33007 Oviedo, Spain.
e-mail: maria.garzon.martin@gmail.com

† Bergen Software Services Int.,
Post Box 2921, 5852 Bergen, Norway.
e-mail:len@bssi-tt.com

†† Dept. of Math., U.C. Berkeley and Math. Dept., Lawrence Berkeley National Laboratory
1 Cyclotron Road, Berkeley, CA, USA
e-mail: sethian@math.berkekey.edu

Key words: Level Set Method, Boundary Integral method, Potential Flow, Drop Coalescence

Abstract. The study of singularities in free surface flows remains a subject of considerably interest. Regarding droplet dynamics the pinch-off events and merging of droplets have been extensively studied due its enormous interest in industrial applications. The level set techniques allows to embed the partial differential equations posed on a free boundary into one higher dimension equations posed on a fixed domain, in such a way that the classical potential flow model can be re-formulated in a complete Eulerian frame work, with the advantage that free boundary topological changes are automatically included. The Laplace equation for the velocity potential is solved via its integral formulation and a boundary element approximation, whereas the evolution of the level set function and extended velocity potential function is approximated using first order finite differences schemes. Merging and splitting events are therefore computationally possible. In the case of two equal drops coalescing, initial instants are very difficult to compute and also to see experimentally. After initial contact a liquid bridge connecting the two drops grows on time and a capillary wave, generated at the point of contact, propagates towards the drop end points. Numerical results regarding two droplet coalescence are presented and a detail discussion of the main flow characteristics is addressed. Comparison with previous computations and laboratory experiments will be also included.

1 INTRODUCTION

The phenomenon of coalescence entails a singularity in the free surface flow when the two liquid droplets initially touch and merge to form a single body. Understanding this complex flow is very important from both the theoretical view point as well as with regard to various industrial applications such as emulsion stability, cloud formation and nanofluidic technologies.

The general mathematical model to simulate this kind of flows is the Navier-Stokes equations but its complexity and computational cost have motivated the use of various simplifying assumptions. For low viscous fluids like water, the motion can be considered inviscid and irrotational up to length scales of few nanometers. One method for potential flow computation in moving and breaking domains, uses the level set embedding techniques to establish an Eulerian formulation of the classical Lagrangian equations [1, 2]. The advantage of this approach is that it seamlessly allows topologically breakup of droplets. It has proven to be a robust method in simulating various physical situations, such as wave overturning and breaking, see [1]; the Taylor-Rayleigh instability of a fluid jet, [2]; droplet and bubble evolution in a two fluid system, [3, 4]; and more recently electrical droplet deformation, [5].

In this paper we use the same model and numerical algorithm to simulate drop coalescence of two equal size droplets. In particular we analyze how our numerical method handles the onset of the singularity and the subsequent propagation of the surface capillary waves generated at the contact point. The evolution of the bridge radios is carefully studied to establish possible self similar solutions and scaling laws.

2 THE MODEL EQUATIONS

To introduce the model equations we follow the presentation in [2] and consider two time-dependent regions occupied by different fluids and separated by a moving boundary. We use the subscript F for the liquid and E for the exterior fluid. Let $\Omega_F(t)$, $\Omega_E(t)$ be three-dimensional moving fluid domains, $\Gamma(t)$ a parameterization of the free surface between both domains at time $t \in [0, T]$ and $\mathbf{R}(\mathbf{s}, t)$, $\mathbf{s} = (s_1, s_2)$ the position vector of a fluid particle on the moving front. See Fig. 1 for a 3D sketch of the physical domain.

We assume that the fluid in $\Omega_E(t)$ is at rest and p_a is the constant reference pressure. The fluid occupying $\Omega_F(t)$ is considered incompressible, irrotational and inviscid, and thus the conservation laws of mass and momentum in $\Omega_F(t)$ lead to the classical potential Eulerian-Lagrangian formulation:

$$\mathbf{u} = -\nabla\phi \quad \text{in } \Omega_F(t) \tag{1}$$

$$\Delta\phi = 0 \quad \text{in } \Omega_F(t) \tag{2}$$

$$D_t\mathbf{R} = \mathbf{u} \quad \text{on } \Gamma(t) \tag{3}$$

$$D_t\phi = f \quad \text{on } \Gamma(t), \tag{4}$$

where $\mathbf{u}$ is the velocity field, ϕ the velocity potential, p the pressure field and D_t stands

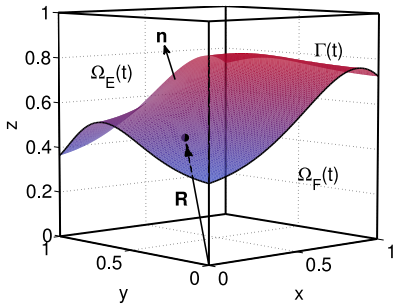


Figure 1: Schematic figure of the physical setting. $\Omega_F(t)$ is the fluid domain and $\Omega_E(t)$ is the exterior domain.

for the convective derivative, $D_t\phi = \frac{\partial \phi}{\partial t} + \mathbf{u} \cdot \nabla \phi$. The right hand side of Eqn. (4) is the function

$$f = -U + \frac{1}{2}|\mathbf{u}|^2 - \frac{p - p_a}{\rho}, \tag{5}$$

where the different terms represent the competing forces acting on the free boundary. Here ρ is the fluid density and U the potential function for the remainder of the body forces per unit mass in $\Omega_F(t)$. Depending upon the flow assumptions, some of these terms may be negligible compared to the others. For example, in the case of gravity wave propagation, see [1], the gravitational field $\mathbf{g} = -\nabla U$ has to be considered and surface tension forces can be neglected. Alternatively, the Rayleigh-Taylor instability and breakup of an infinite jet and the coalescence of liquid droplets are processes driven by surface tension and gravity terms can be neglected, see [2].

The pressure jump across the free surface $\Gamma(t)$ is therefore given by

$$p - p_a = \gamma \kappa \tag{6}$$

where κ is twice the mean curvature of the surface and γ the surface tension coefficient. Then equation 5 reads

$$f = \frac{1}{2}|\mathbf{u}|^2 - \gamma \kappa \tag{7}$$

Note that gravity forces have been neglected as inertia, surface tension are the dominant forces in the physical applications of drop coalescence presented here.

To make the equations dimensionless we introduce the usual characteristic scales for length r_0, capillary time t_0. The non-dimensional model equations remain the same except for f that becomes

$$f = \frac{1}{2}|\mathbf{u}|^2 - \kappa, \tag{8}$$

In what follows all the variables are dimensionless. We have thus avoided a separate notation for the dimensionless variables.

Next, we formulate the above Eulerian-Lagrangian equations in a complete Eulerian framework.

2.1 The Eulerian formulation of the model equations

Let Ω_D be a fixed three dimensional domain that will contain the free boundary for all times. In Fig. 1, $\Omega_D = \Omega_F \cup \Omega_E$. Eqns. (3) and (4), which are posed on a moving surface, can be reformulated in this fixed domain using the level set/extended potential technique described in [1, 2]. In this approach, the moving front $\Gamma(t)$ and velocity potential $\phi|_{\Gamma(t)}$ are embedded into functions Ψ and G of one higher dimension, respectively. The level set function Ψ and the extended velocity potential function G are defined on the fixed computational domain Ω_D that should contain the free boundary for $t \in [0, T]$ and such that

$$\Psi(\mathbf{R}(\mathbf{s}, t), t) = 0 \tag{9}$$

$$G(\mathbf{R}(\mathbf{s}, t), t) = \phi|_{\Gamma(t)} \tag{10}$$

for $t \in [0, T]$. Following the derivation in [2], Eqns. (3) and (4) transform into:

$$\Psi_t + \mathbf{u}_{\text{ext}} \cdot \nabla \Psi = 0 \quad \text{in } \Omega_D \tag{11}$$

$$G_t + \mathbf{u}_{\text{ext}} \cdot \nabla G = f_{\text{ext}} \quad \text{in } \Omega_D, \tag{12}$$

respectively. The subscript "ext" in Eqns. (11) and (12) denotes the extension of f and $\mathbf{u}$ onto Ω_D, the details of how to perform these extensions will be explained later. Note that we here have set $\Omega_D = \Omega_F \cup \Omega_E$, but it could be chosen in any other way. The only requirement is that the free boundary should always be included in Ω_D.

The Eulerian model equations for the coalescence problem are thus:

$$\mathbf{u} = -\nabla \phi \quad \text{in } \Omega_F(t) \tag{13}$$

$$\Delta \phi = 0 \quad \text{in } \Omega_F(t) \tag{14}$$

$$\Psi_t + \mathbf{u}_{\text{ext}} \cdot \nabla \Psi = 0 \quad \text{in } \Omega_D \tag{15}$$

$$G_t + \mathbf{u}_{\text{ext}} \cdot \nabla G = f_{\text{ext}} \quad \text{in } \Omega_D \tag{16}$$

$$\tag{17}$$

The rest of the boundary conditions and domain geometries have to be specified for each particular application.

The main advantage of the level set/extended potential formulation relies on the fact that any topological change of the free surface and evolving magnitudes within this boundary are directly taken into account by Eqns. (15) and (16). Moreover, the embedding provides a convenient regularization of possible singular geometries and prevents the blow up of diverging flow variables. This situation often occurs at fluid break-up and merging.

3 NUMERICAL APPROXIMATION

The time discretization is done using a explicit forward Euler scheme with time step Δt. At each time step t_n, given $\Psi^n = \Psi(t_n)$ and $G^n = G(t_n)$ the following semi-discretized system has to be solved:

$$\mathbf{u}^n = -\nabla\phi^n \quad \text{in } \Omega_F(t_n) \tag{18}$$

$$\Delta\phi^n(x,y,z) = 0 \quad \text{in } \Omega_F(t_n) \tag{19}$$

$$\phi^n|_{\Gamma_{t_n}} = G^n|_{\Gamma_{t_n}} \tag{20}$$

$$\frac{\Psi^{n+1} - \Psi^n}{\Delta t} = -\mathbf{u}^n_{\text{ext}} \cdot \nabla\Psi^n \quad \text{in } \Omega_D \tag{21}$$

$$\frac{G^{n+1} - G^n}{\Delta t} = -\mathbf{u}^n_{\text{ext}} \cdot \nabla G^n + f^n_{\text{ext}} \quad \text{in } \Omega_D. \tag{22}$$

Eqns. (18) to (20), which will be approximated using the Boundary element method, are strongly coupled to Eqns. (21) and (22), which are approximated by applying suitable finite difference level set schemes. First, the Dirichlet boundary condition (20) is obtained from the spatial mesh values of G^n in Ω_D and secondly $\mathbf{u}$ has to be provided from the boundary element calculation to yield $\mathbf{u}^n_{\text{ext}}$ and f^n_{ext} in the mesh points of Ω_D.

Regarding the spatial discretization, linear elements are used to discretize the free boundary and finite difference upwind schemes are used to approximate the level set equations (21), (22).

The basic algorithm can be summarized as follows:
For $n = 1, \ldots,$ number of time steps

1. Calculate $\mathbf{u}^n$ using the Boundary element method to solve Eqns. (19),(18).
2. Extend $\mathbf{u}^n$ on to Ω_D using the level set extension techniques.
3. Update Ψ at t_{n+1} using Eqn. (21) and the finite difference level set schemes.
4. Calculate the free boundary curvature κ^{n+1}
5. Calculate f^n as $f^n = \frac{1}{2}|\mathbf{u}^n|^2 - \kappa^{n+1}$.
6. Extend f^n on to Ω_D using the level set extension techniques.
7. Calculate G^{n+1} using Eqn. (22) and the finite difference level set schemes.
8. Interpolate G^{n+1} from the finite difference points to the boundary.

End For

4 DROP COALESCENCE SIMULATIONS

During coalescence of two drops of the same size they first touch and then merge as the liquid bridge connecting them grows in time. A capillary wave develops at the contact point and propagates away from the singularity. At early stages of the merging process the drop end points barely move, the dynamics occurs mainly in the bridge zone. Occasionally the capillary wave reaches the drop ends causing visible oscillations until the equilibrium state is finally achieved. For the numerical simulations presented here we take as initial geometry two spherical droplets of $r_0 = 1$ centered at $z = 0$ and $z = 2$, respectively, such that the initial contact point lies at $z = 1$. The fixed domain for the level set computations is set to $\Omega_D = [-1, 5, 3.5] \times [-1.5, 1.5]$ and the time span considered is $t \in [0, 1]$. A good practice to validate numerical results is to check for convergence with respect the discretization parameters. Simulations are run with (a) coarse and (b) fine grids with sizes:
(a) $\Delta r = \Delta z = 0.005$
(b) $\Delta r = \Delta z = 0.0025$.
To discretize the free boundary a high resolution is needed near the initial contact point, whereas at both ends of the droplet the size of the surface grid can be larger. We start with $\Delta s = 0.02$ near $z = 1$ and increase this distance gradually away from the center point using the regridding technique established in [2]. Therefore the number of points N_p to represent the free surface varies with time between $N_p = 141$ and $N_p = 135$. The time step is chosen adaptatively according with the criteria

$$\Delta t \leq \min(\frac{\Delta r}{|u_{\max}|}, 0.2\Delta s^{3/2}),$$

which leads to time steps as low as $\Delta t = 1. \times 10^{-5}$ to accurately resolve for the initial perturbation and can be increased to $\Delta t = 1. \times 10^{-3}$ once a steady neck growth is achieved.

A very challenging issue that remains open regarding full numerical computations is how to establish the initial conditions in the drop coalescence problem to avoid the singularity at $t = 0$. The general approach is to assume that immediately after the two free surfaces touch, a bridge of small but yet finite size already exists [6, 7, 8]. Another recent approach is to use the so called interface/disappearance model described in [9], which suppresses the initial singularity but entangles a much more complex mathematical model and the tuning of various model parameters. Within the present model and numerical framework, and due to the robustness of the level set method to handle topological changes, the initial condition is just two spheres touching tangentially at a single point and no artificial smoothing (besides the inherent to the discretization procedure) is needed. We do not claim the physical correctness of our initial contact evolution (molecular forces would probably play an important role) but just its computational simplicity, provided enough spatial-temporal resolution is given to resolve the small scales involved. We restrict ourselves to report the numerical results obtained which can be of interest to the computational community. In the present computations the early stages occur

Figure 2: Initial stage evolution: Reconnection event at $t = 0.00155$

in the time interval $t \in [0, 0.002]$, where the length of the contact line is stable or even diminishes with time. It is worthy to report that at $t = 0.00111$ and $t = 0.00155$ two free boundaries re-connections take place, hypothetically entrapping a thoroidal bubble of radius 5×10^{-5} non dimensional units, see a zoom of the area at $t = 0.00155$ depicted in Fig. 2. These events are easily handled by the level set technique and the computation proceeds smoothly.

Once the onset of the liquid bridge between the drops has taken place, how its minimum neck radios $r_{\min}$ evolves with time is a subject of interest, as theoretical works and experiments indicate the existence of certain scaling laws. Duchemin *et al* [6] established that, for inviscid fluids, capillary pressure should balance dynamical pressure at early stages of the process, that is $\rho(\frac{dr_{\min}}{dt})^2 \approx \Delta p$, from which the following scaling law is obtained:

$$r_{\min} \approx (\frac{\gamma R}{\rho})^{1/4} t^{1/2}$$

More recently, Paulsen *et al* [10] distinguished two different regimes regarding inviscid drop coalescence: an initial regime (never identified before) which they named the inertial-limited-viscous regime and a pure inertial regime. The scaling laws proposed by these authors are $r_{\min} \approx t$ and the above $r_{\min} \approx t^{1/2}$ respectively. They conclude that the initial regime should apply for drops of any viscosity. Nevertheless it is difficult to establish in which time range (or $r_{\min}$ range) these two regimes should apply due to the difficulty of obtaining reliable experimental or numerical data at such early stages of bridge formation.

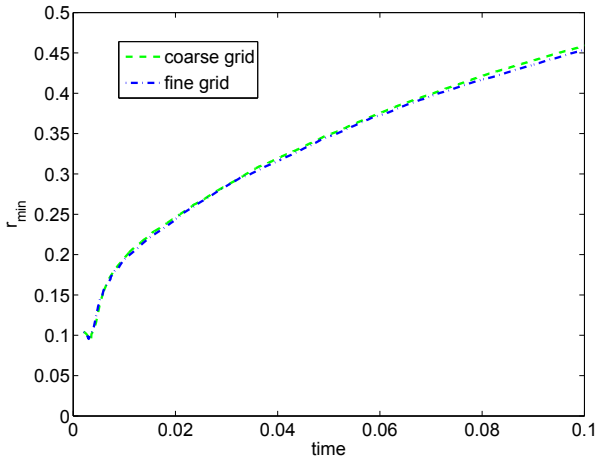


Figure 3: Minimum neck radios evolution

Therefore, one unavoidable characterization of the merging process is the time history of the minimum neck radios. In Fig. 3 we plot $r_{\min}$ versus t for both mesh sizes, from which we can conclude the independence of the results with respect the discretization parameters at the present refinement level. In what follows the numerical results correspond to the coarse grid. To check for the scaling laws proposed previously we plot in Fig. 4 $\log(r_{\min})$ versus $\log(t)$. Looking at this graph we can distinguish two very different slopes. A linear fit performed with Matlab gives an exponent of $\alpha = 1$ for very early times in $t \in [0.004, 0.007]$, and a very stable exponent of $\alpha = 0.4$ for times in the range $[0.018, 0.15]$ (and even beyond). The transition law between these two scaling laws would correspond to the theoretical $\alpha = 0.5$. The more persisting exponent of $\alpha = 0.4$ found in this work agrees very well with previously reported experimental results, see [8], [11], [7].

In Fig. 5 the evolution of the left end drop axial coordinate is depicted. As it can be observed from this graph, droplet deformation from $t \in [0, 0.4]$ is very much localized on the bridge region as the end points barely move. The capillary wave reaches these end points at around $t = 0.4$. In Fig. 6 we show the amplitude of the capillary wave with respect the superimposed initial condition at time $t = 0.1$. We have focused the capillary wave amplitude over the radial direction $y = x$. The amplitude of the wave with respect the sphere of $R = 1$ is approximately $4 - 5 \times 10^3$.

Drop profiles at selected times are depicted in Fig. 7. To compare our simulated results with the laboratory experiments in [11] we have to transform our non-dimensional times to dimensional ones multiplying by the characteristic time scale $t_o = \sqrt{\frac{\rho r_0^3}{\gamma}} = 1720 \mu s$. Here we have taken the usual water properties and $r_0 = 0.6$mm, which is the radios of the droplets in the experiments by Thoroddsen (see figure 12 of [11]). We observe that not only the drop profiles are in very good agreement with the experimental ones, but also the time occurrences.

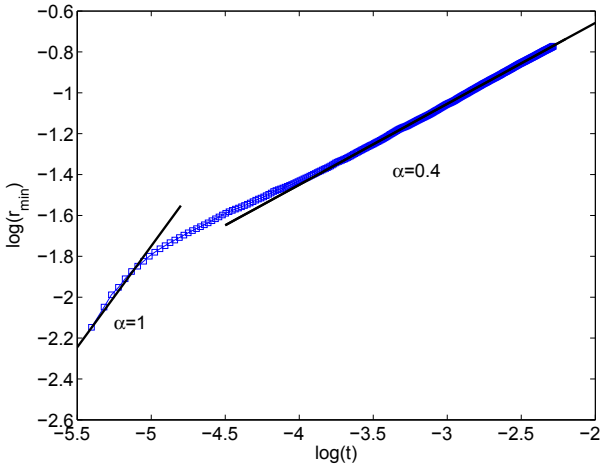


Figure 4: Log-log plot of minimum neck radios evolution

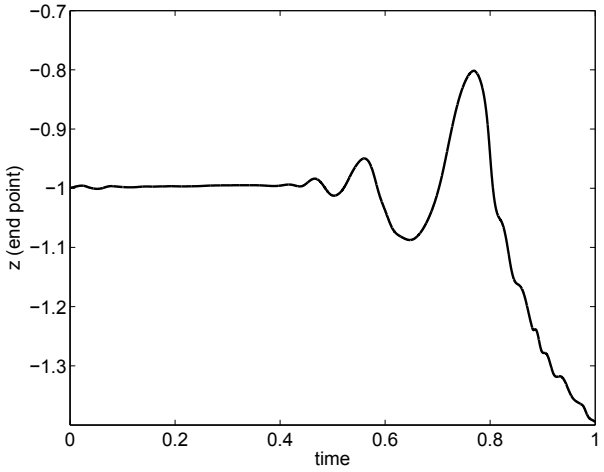


Figure 5: Left end drop axial coordinate evolution

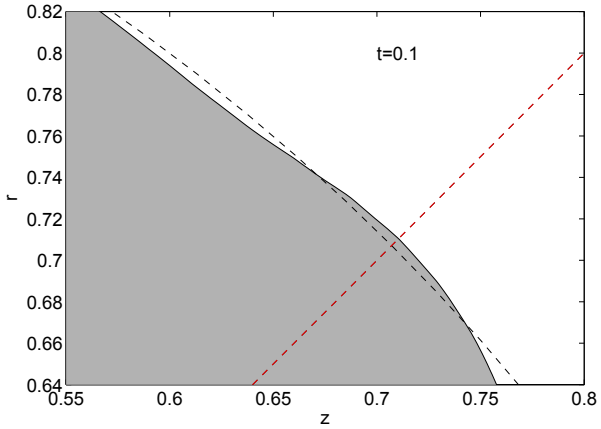


Figure 6: Capillary wave amplitude. Focused front and superimposed initial condition (dashed) at t=1. The dashed red line is the line $y = x$

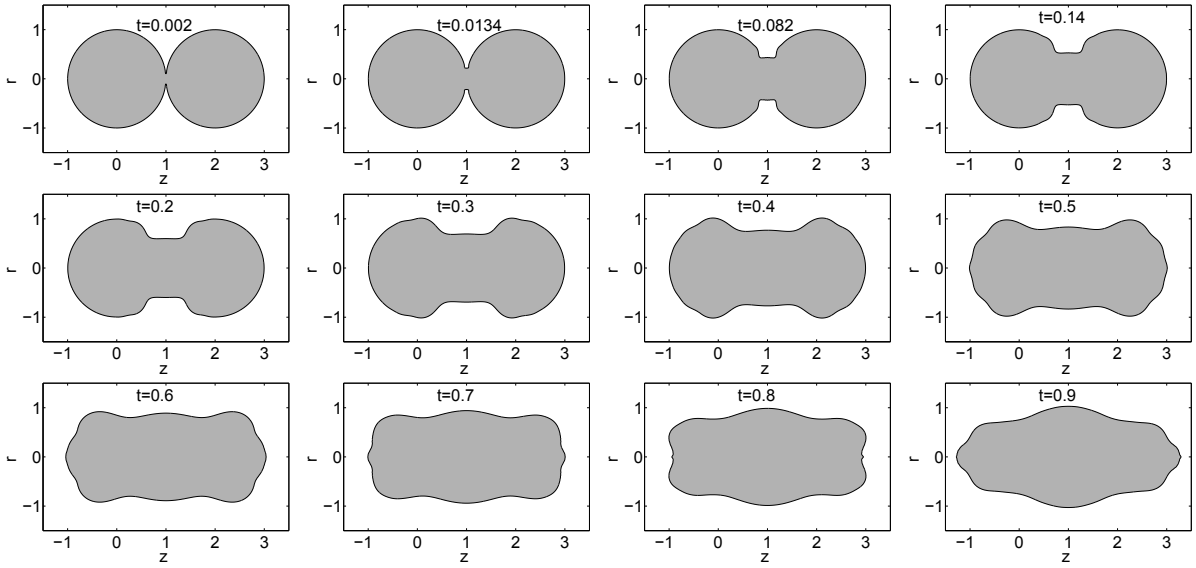


Figure 7: Fronts of two equal drops coalescing at indicated times

5 CONCLUSIONS

- In this paper we have used a three-dimensional algorithm with axial symmetry, based on the Boundary Integral-level set coupling to approximate hydrodynamic problems with free boundaries. Using the level set/extended potential model singular flow events, such as fluid break-up or merging, are easily handled and the computations can go past these singular times.

- In particular we have presented here the numerical results obtained in the case of two drops of same size coalescing. With our numerical technique there is no need to start the calculations with the artificial initial bridge needed in previous numerical simulations. Moreover, the scaling laws obtained, drop profiles and time occurrences are in very good agreement with published laboratory experiments.

REFERENCES

[1] Garzon, M., Adalsteinsson, D., Gray, L. J. and Sethian, J. A. A coupled level set-boundary integral method for moving boundary simulations. *Interfaces and Free Boundaries* (2005) **7**: 227-302.

[2] Garzon, M. and Gray, L. J. and Sethian, J. A. Numerical simulation of non-viscous liquid pinch-off using a coupled levelset-boundary integral method. *J. Comput. Phys.* (2009) **228**: 6079-6106.

[3] Garzon, M., Gray, L. J. and Sethian, J. A. Simulation of the droplet-to-bubble transition in a two-fluid system. *Phys. Rev. E* (2001) **83**: 046318.

[4] Garzon, M., Gray, L. J. and Sethian, J. A. Axisymmetric boundary integral formulation for a two-fluid system. *Int. J. Numer. Meth. in Fluids* (2012) **69**: 1124-1134.

[5] Garzon, M., Gray, L. J. and Sethian, J. A. Numerical simulations of electrostatically driven jets from nonviscous droplets. *Phys. Rev. E* (2014) **89**: 033011.

[6] Duchemin, L., Eggers, J. and Josserand C. Inviscid coalescence of drops. *J. Fluid Mech.* (2003) **487**: 167-178.

[7] Paulsen, J. D. *et all.* Coalescence of bubbles and drops in an outer fluid. *Nature Communications* (2014) **5**: 3182.

[8] Menchaca-Rocha, A., Martinez-Davalos, A., Nunez, R., Popinet, S. and Zaleski, S. Coalescence of liquid drops by surface tension. *Phys. Rev. E* (2001) **63**, 046309.

[9] Sprittles, J. E. and Shikhmurzaev, Y. D. Coalescence of liquid drops: Different models versus experiment. *Physics of fluids* (2012) **24** 122105.

[10] Paulsen J. D. *et al.* The inexorable resistance of inertia determines the initial regime of drop coalescence *Proc. Natl Acad. Sci. USA* (2012) **109**: 68576861.

[11] Thoroddsen T. S., Takehara K.,Etoh T. G. The coalescence speed of a pendant and sessile drop *J. Fluid Mech.* (2005) **527**: 85-114.

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis (Eds)

FLUID-THERMAL-STRUCTURE COUPLED ANALYSIS OF GRID FINS FOR HYPERSONIC FLIGHT VEHICLE

SHENGZE LI*, ZHENYU JIANG*, WEIHUA ZHANG* AND KE PENG*

* College of Aerospace Science and Engineering
National University of Defense Technology
109 Deya Road, 410073 Changsha, China
e-mail:lsz86591989@163.com

Key words: Grid Fin, Hypersonic, Fluid-Thermal-Structure Coupling, Angle of Attack, Thermal Protection.

Abstract. The aim of this work is to provide a numerical method of fluid-thermal-structure coupled simulation of grid fins in hypersonic flows. Hypersonic flight vehicle with grid fins offers many advantages under high speed condition. The main advantages are small hinge moments, efficient packaging and capability to produce effective aerodynamic force at high angles of attack over wide Mach number ranges. The structure, however, sustains complicated loads caused by aerodynamic forces and heating during hypersonic flight. The large aggregations of experimental studies are limited to parts of physical characteristic due to the difficulty in the data transmission. This study takes the full influence of fluid flow into account as well as heat transfer on structure strength and stiffness. The following work is done. Firstly, Reynolds-Averaged approach and Spalart-Allmaras turbulence model are employed to solve the problem of flowing fluid. The simulation is achieved by using second order upwind scheme and hexahedral mesh grids. Validation has done by comparing the computed normal force coefficient with wind tunnel data for a Mach number from 1.8 to 3.5 and different angles of attack. Secondly, computations are performed at free stream Mach number 6 at five angles of attack from 0 to 15 degrees. The data of aerodynamic forces and heat flux is obtained after fluid-thermal coupled calculation. The strength and stiffness calculations are carried out by taking the aerodynamic forces and heat flux as boundary condition of solid domain. Finally, two types of structure form are discussed for long time thermal protection. One is niobium alloy, the other is C/SiC composites. The article shows that different thermal and mechanical predictions are affected by various conditions such as angle of attack, material characteristics. The structure response takes on a whole different status under different angles of attack. This allows engineers to choose appropriate size and type for hypersonic thermal protection materials.

1 INTRODUCTION

With the comprehensive and integrated development of engineering system, researches on some problems are no longer confined to a single discipline, and the multi-physics coupling analysis is a universal example. The analysis often requires different processes simulations for the same region in order to get some results of two or more field interactions. Therefore, it's a

great challenge for the research of multi-field couplings and its application in engineering problems, and the coupling mechanism has become a research focus of many disciplines.

Grid fin is a nonconventional aerodynamic lifting and control surface consisting of an outer frame with internal lattice[1]. In the 1940s, scientists from the Soviet Union first carried out theoretical and experimental researches of the grid fin[2, 3] in aerodynamics, structure, strength, quality, and manufacturing process, but because of lack of knowledge of various features of such wing surface, the grid fin was not widely used. In recent years, the grid fin has again attracted the attention of the world. Russia, United States, Germany and China have strengthened theoretical researches and experimental development of grid fin and have successfully applied it to satellites, rockets and missiles. After that, many advantages of grid fin have been proved by the studies such as low hinge moment, efficient packaging, suitability of tube launching and good lifting characteristics in supersonic or hypersonic flows[4-7]. Naturally, with the hypersonic techniques development, the researches of grid fin are also extended to the field of hypersonic applications to achieve much better performance than the conventional wing configurations.

However, it's not quite easy for grid fin using in hypersonic flying which will be affected by the flow field, thermal field and structure field at the same time. The air flows opposite to the wings and provides lift force for the flight vehicle, while the wings stressed and deformed conversely affects the distribution of the flow field. Air friction changes the surface temperature of the wings, and the temperature conversely affects the physical properties of structure and fluid. During the flight, it may encounter any or all of the following conditions: 1) extremely hot surface temperature; 2) large temperature gradients; 3) transient heating. In response to these extreme conditions, the grid fin may experience any or all of the following conditions: 1) possibly large thermal stress superimposed on aerodynamic pressure; 2) changes in material properties such as strength and elastic modulus; 3) high temperature creep. In addition, it's really difficult to obtain reliable experimental structural response with a complex fluid-thermal superposition boundary condition. Hence, there are many significant challenges in predicting structural response in such an extreme environments, including: 1) the simultaneous response of fluctuating pressures, aerodynamic heating and structure deflection over long durations; 2) material nonlinearity; 3) spatial variation of material and structure properties; 4) uncertainty in loads, material properties, geometry, and boundary conditions; and 5) the computational cost and complexity of large, coupled models.

Traditionally, in order to tackle those complex issues occurring in design process, an aerodynamicist will predict the surface pressures and heating rates by assuming a rigid isothermal body. These aerodynamic heating rates are used by a structural heat transfer analyst to predict the structure temperature distribution. Finally, a structural analyst uses the temperature distribution and aerodynamic pressure to predict the structural deformations and stresses. Such traditional independent approaches require several interactions between the different analysis methods and analysts. The approach is relatively inefficient and untrustworthy. Meanwhile, early theoretical calculations cannot fully reflect the coupling between the three physical fields of grid fin[8]. Even though a system of partial differential equations describing such problem is obtained, if it is not simplified, it's also difficult to get solved manually. In fact, in the early 1900s, Lamb et al. proposed the problem that such structure had an impact on other media[9]. After that, Whitehead[10] proposed a classic semi-empirical analysis method, *actuator disk*, for the numerical analysis of the flow field in 1959.

Whereas, because of other restrictions, not until the late 1960s when the numerical methods of finite element and boundary element came into being did the detailed analysis appear to solve the problem. In *Applied Mechanics Review* published by the American Society of Mechanical Engineers, there was no term of fluid-structure interaction from 1970 to 1980, and related terms were only aeroelasticity, tremor, etc. Not until 1981~1983 did the term fluid-structure interaction (FSI) began to come into being. With the establishment of the entry FSI, in these three years, the number of abstracts under this entry was up to 49, 73 and 50, which also shows certain development of fluid-structure interaction in the 1980s. In 1991, Bendiksen[11] was the first person to propose that fluid and structure domains could be solved through a unified equation, i.e. two-way media interaction. It meant that while the structure was solved, the force to the structure generated by the fluid should be explicitly solved. This forced decoupling method shows that people then began to use some effective ways to solve the problem of fluid-structure interaction. In 1992, Takashi[12] used Arbitrary Lagrange-Euler (ALE) for the description to solve the uniform hybrid equation and applied semi-implicit and semi-explicit method to the fluid and finally solve the N-S equation. In 2001, Choi[13] separated variables to solve the hybrid equation of fluid-structure interaction proposed by predecessors, who introduced the symmetrical pressure equation and applied the four-step method and ALE for the solution. From 2001 to 2004, Hubner[14-16] adopted the space finite element with discontinuous stabilized time to discretize the model equation, establish a single system of equations and solve the problem of flow around of two-dimensional viscous fluid. In 2005, Namkoong[17] proposed a method to directly solve the hybrid equation and address the problem of two-dimensional channel flow of two-dimensional vertical plate division. At the same time, in order to solve the problem of buckling and large deformation of the three-dimensional structure, Zhang and Hisada[18] jointly developed a program of FSI that provided a method to follow to solve fluid-structure interaction, who at the same time applied ALE to solve the fluid-structure interaction problem of two-dimensional heart beat and three-dimensional large deformation. Then, with the requirement of developing advanced hypersonic flight vehicle, the thermal influence was added in the fluid-structure coupling problem in order to synthesis the three fields interaction to support the structure design and the thermal protection system design for the hypersonic flight vehicle. In 2010, Culler[19] carried out the fluid-thermal-structure simulation on the hypersonic flight vehicle and validate the aerodynamic force and heating results with some wind tunnel test data. It shows goodness fit with the experiments and represent the future trends of flight vehicle design methodology in using fluid-thermal-structure integration method. Obviously, this method also suggests a new way of thinking about the grid fin designing and its application in hypersonic flight vehicle. In this paper, we presents a numerical approach to study the fluid-thermal-structure coupling problems of the grid fin. The influence factors of the attack angle on different field effect will be numerically discussed, and the structure response with two different material will also be revealed. The objective of this work is to provide investigators and designers of the grid fin with a relatively systematic reference.

2 NUMERICAL MODEL

2.1 Geometric model

In order to facilitate the verification of the reliability of the simulation method, this paper

established the geometric model of grid fin in accordance with reference literature[1]. The geometrical configuration and dimensions are shown in Figure 1. The grid fin is at the tail of the missile, which can work within a time of 200s, at a height of 20km and a flight speed of 6Ma.

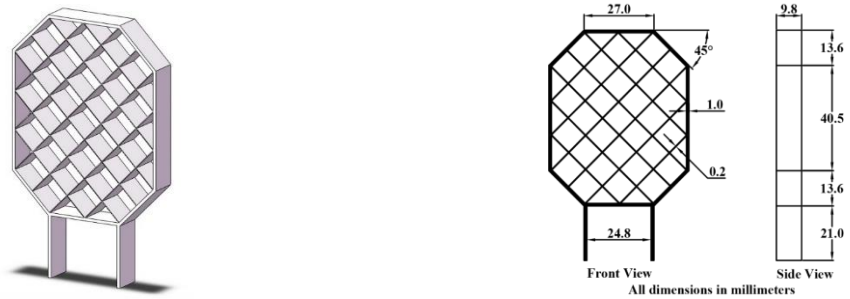


Figure 1: Geometric model and dimensions of grid fin

2.2 Flow field model

Figure 2 and 3 shows the mesh of flow field in ICEM. The minimum distance between the far field and the grid fin is 20 times of the chord length to ensure that the border will not interfere with the flow field around the grid fin. To facilitate the application of hexahedral mesh, the flow field is divided into multiblocks. On the contour of the fin, there are 118,883 nodes distributed, and after swept along the flow direction, a total of 13,465,840 units are generated. The grids on the boundary layer are densified.

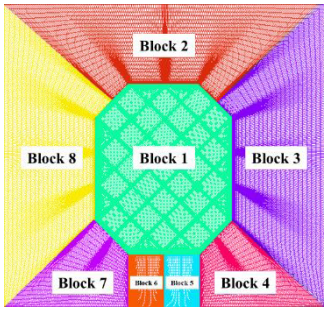


Figure 2: Local meshes distributed on flow direction

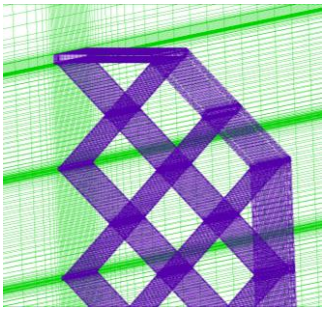


Figure 3: Local meshes distributed on grid fin surface

After the geometric model and mesh division are completed, the coupled areas need to be marked so that the corresponding coupling surface can be selected for the coupling calculation. The independent set is respectively named for inlet, outlet, far-field boundary of the flow field and the contour of grid fin. Reynolds-Averaged approach and Spalart-Allmaras turbulence model are employed to solve the problem of flowing fluid, and the second order upwind scheme is used to achieve the simulation.

2.3 Model validation

This paper makes reference to the simulation results of the grid fin in literature[1] and

makes comparison between the angle of attack and normal force coefficient curve to verify the correctness of the established geometry and grid of the flow field. The numerical model of grid fin is validated with the flow field parameters: 1) free incoming flow Mach number Ma=1.8, 2.5 and 3.5; 2) incoming flow temperature T=293.15K; 3) Reynolds number Re=2.3×10^6, 3.0×10^6 and 4.7×10^6; 4) angle of attack α=0~15°and interval of 5°; 5) angle of sweepback δ=0, -15°. Through the numerical simulation at multiple angles of attack of grid fin, the relationship of change between the normal force coefficient and angle of attack is obtained, as shown in Figure 4 and 5. The grid fin's normal force coefficient increases gradually with the growth of the angle of attack. It can be seen from the comparison between the measured data and model values that at the same angle of attack the normal force coefficient of the grid fin is very close and the minimum error is less than 18% (δ=0°) and 22% (δ=-15°), therefore, the parameter settings and grid distribution of the flow field model are reasonable.

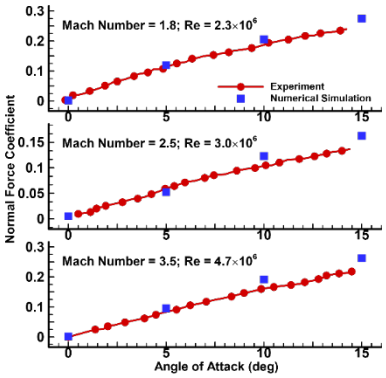


Figure 4: Normal force coefficient versus angle of attack, δ=0°

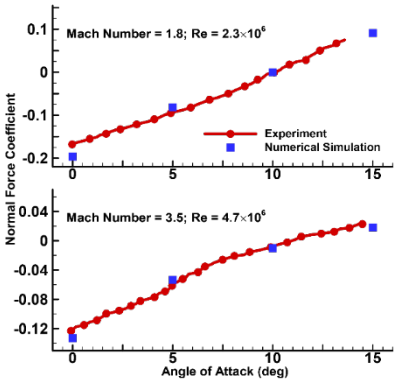


Figure 5: Normal force coefficient versus angle of attack, δ=-15°

3 NUMERICAL CALCULATION AND ANALYSIS

3.1 Coupling calculation strategy

For the coupling calculation of heat transfer between the flow field and the structure field, data exchange occurs on the wall contacting the flow field, called the coupling interface. Data repeatedly exchange on the coupling interface to satisfy the boundary conditions of the next calculation of the flow field and the structure field. The flow field requires new temperature boundary conditions provided by the coupling interface, while the structure field needs new heat flow boundary conditions. The coupling scheme in this paper is as follows:

(1) At the time t=0s, a steady-state solution of the flow field is first calculated under the wall boundary conditions at a constant temperature, and then the heat flow value on the coupling interface is obtained (heat flow transferred from the flow field to the wall), represented by q_f (f represents the flow field).

(2) Without wall radiation effect considered, according to energy conservation, the heat flow transferred from the wall into the structure field is equal to that from the flow field to the wall. So the heat flow boundary condition (s represents the structure field) to calculate the

heat transfer of the structure field can be as follows:

$$k_s\left(\frac{\partial T_x}{\partial x}n_x+\frac{\partial T_y}{\partial y}n_y+\frac{\partial T_x}{\partial z}n_z\right)=q_f \tag{1}$$

(3) The above formula is taken as the boundary condition of the structure field for solving the heat transfer equation, and after a time of Δt (the coupling time of this model is 1s), the temperature T_s of the re-distributed field structure can be obtained (the structure field is in the uniform temperature state initially).

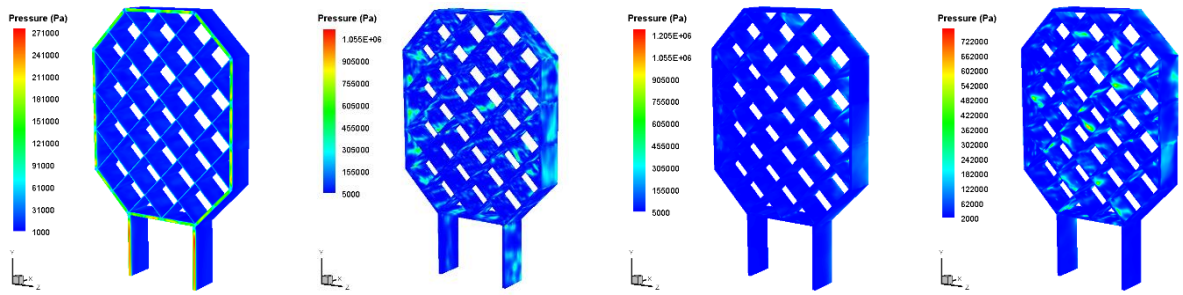
(4) After the calculation of a time step Δt (FEM) of the structure field, the temperature boundary condition T_s of the next calculated of the flow field can be obtained, i.e. the new wall temperature.

(5) At the temperature boundary T_s the transient flow field calculation is carried out. After the time Δt, the new wall heat flow value q_f can be obtained. Then this variable is extracted from the flow field and loaded into the structure field as a boundary condition, and the next calculation is carried out.

(6) The cycle is repeated until the end of the calculation.

3.2 Fluid thermal coupled

Actually, the validation model is used in the wind tunnel test and too small to be applied for the authentic flight vehicle. Hence, with considering the practical application, a major difference is that the calculation model is 10 times larger than the validation model. Meanwhile, the model is calculated in Ma=6 of free incoming flow at the angles of attack from 0 to 15°. By the calculation, the pressure and heat flux is obtained. Figure 6 shows the nephogram results of surface pressure after 200s flights in different angles of attack. It can be found that, in Figure 6(a), the high pressure area on the grid fin surface are mainly distributed in the region of the leading edge stagnation point, where the maximum pressure is 0.27 MPa. However, with the increasing of angle of attack, the high pressure area are no longer staying in the leading edge region due to the complex shock and expansion wave interaction between the inner lattice wings. Moreover, in Figure 6(c), the result shows that the area occurs in the leeside of gird fin which is totally different from the patchy distributions of angle 5° and 15°.

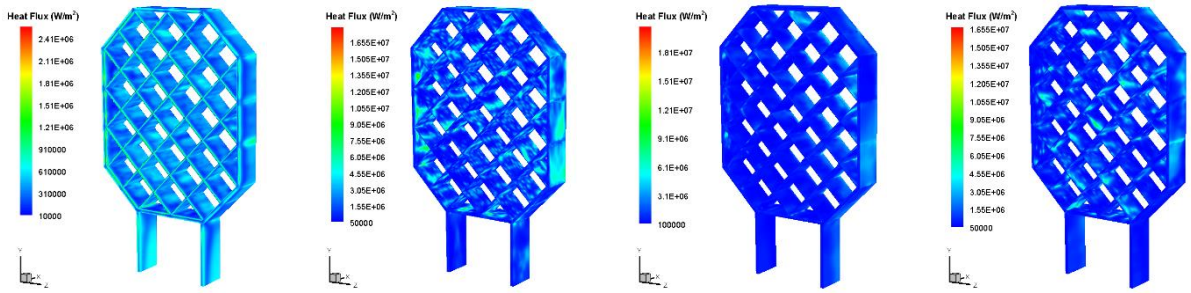


(a) α=0° (b) α=5° (c) α=10° (d) α=15°

Figure 6: Surface pressure nephogram in different angles of attack

Figure 7 shows the nephogram results of heat flux in different angles of attack. The

distributing characteristics are displayed similar to Figure 6 and most area is graded in megawatt per square meter. As the Figure 6 and 7 show, it cannot be simply asserted that the leading edge of the grid fin will be the most severe region of aerodynamic heating as well as the design focus of thermal protection of the structure which just only happened in 0° attack angle. Due to the complex flow separation, the thermal protection design should synthesize each kind of situation. In addition, it should also be noted that the high pressure and heat flux environment will further promote heat exchange between the flow field and the structure field, therefore, to get the further structure response will be the focus of study on the use under hypersonic conditions.



(a) α=0° (b) α=5° (c) α=10° (d) α=15°

Figure 7: Heat flux nephogram in different angles of attack

3.3 Structure response

During a long time flying with the combination of extreme pressure and heat, general metal alloy is difficult to bear, so that, some high temperature materials are taken into account. Firstly, the niobium alloy which could provide an excellent high specific strength from 1093°C to 1427°C is introduced in the coupled calculation with some considerations of its good mechanical property and a broad application scope of temperature. Some properties are shown in the following table:

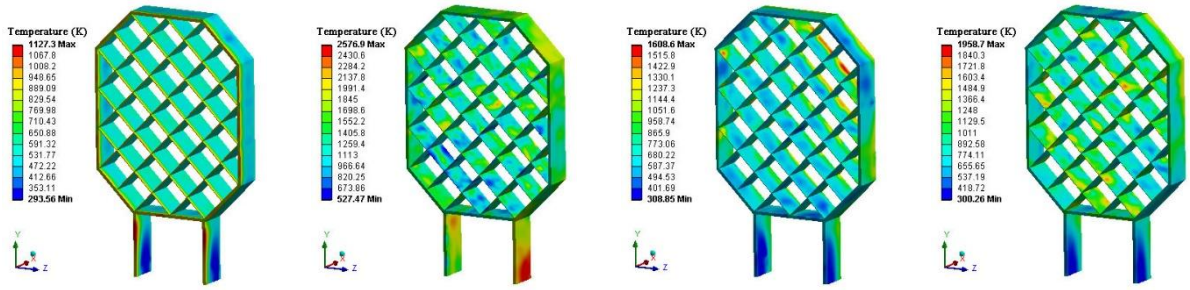
Table 1: Material properties of niobium alloy

Temperature K	Density kg/m^3	Thermal conductivity W/(m·K)	Elasticity modulus GPa	Poisson ratio	Linear expansion coefficient $\times10^{-6}$(K^{-1})	Specific heat J/(kg·K)
1150	8800	37.4	103	0.3	8	273
1580	8800	42.4	112	0.3	10	327.6

About the boundary condition, the displacement restrictions through x, y and z orientation are applied to the bottom holder of grid fin, the initial temperature of the structure is set as 295.15K. The wall boundary of grid fin is treat as the coupling area for exchanging and transferring variable data, the data of coupling is the heat flux variable which will be extracted through the fluid field calculation, at the same time, the temperature variable also could be picked up after the structure field calculation for the next coupling step.

Figure 8 shows the temperature distribution of the grid fin in using niobium alloy after 200s simulation. As the figure shows, in 0° angle of attack, the highest temperature is up to

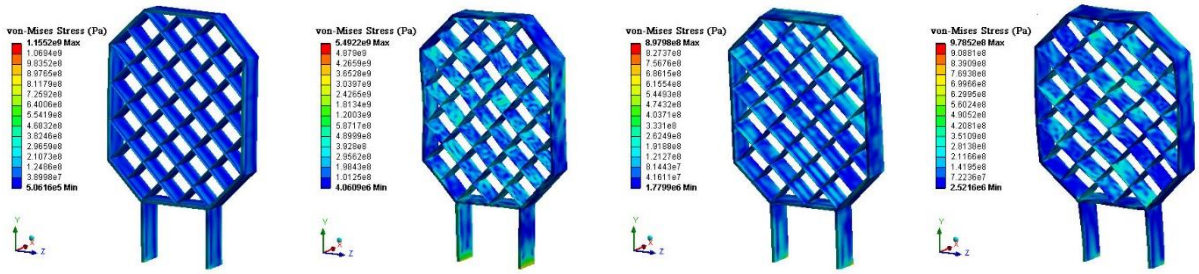
1127K and rapidly declines along the longitudinal direction; in 5°, the temperature of bottom holder sharply climbed to more than 1500K, some area is even as high as 2500K; in 10°, the trailing edge occurs extreme high temperature as the heat flux distribution of Figure 6(c); in 15°, the leading edge becomes the most serious area again as Figure8(a), but the temperature distribution of bottom holder is mostly the same as Figure 6(c). More generally, the temperature distribution is inhomogeneous caused by the differential of heat flux and most region has reached more than 500K, some local area has risen by over 1500K which is the maximum working temperature of niobium alloy.



(a) α=0° (b) α=5° (c) α=10° (d) α=15°

Figure 8: Temperature nephogram of niobium alloy in different angles of attack

The drastic temperature gradients will further effect the strength and stiffness of the structure. As the Figure 9 shows, the average stress level increases with the increasing of attack angle. It should be also noted that, with the existence of the sharp corner, significant stress concentration appears especially for the bottom holder which is one of the weak parts of the main structure and need to be further optimized.

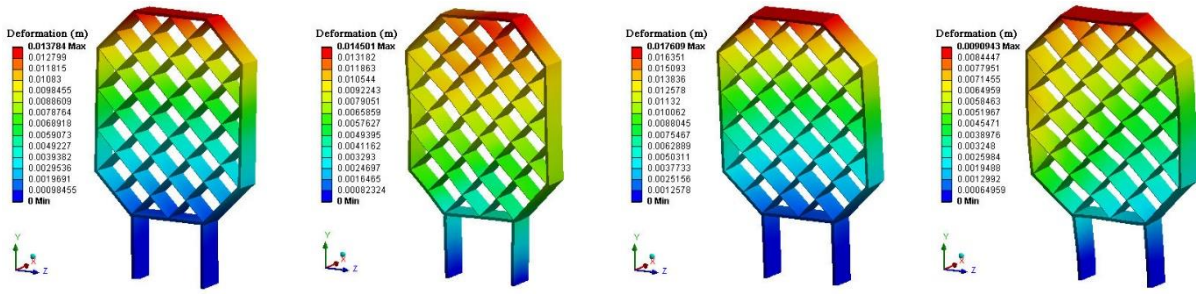


(a) α=0° (b) α=5° (c) α=10° (d) α=15°

Figure 9: Von-Mises stress nephogram of niobium alloy in different angles of attack

As the displacement nephogram shows in Figure 10 which is got from the calculation results, an interesting phenomenon is found that the displacement is not simply enlarged with the increasing of attack angle, in 5°, an obvious extrusion takes place at the left and right edge of the outer frame and in 15°, by contrast, the expansion appears. Furthermore, the maximum displacement reaches 14.5mm that severely impact the aerodynamic configuration for supplying lifting force. Actually, in this model, some more detail properties (e.g. failure strength, fracture strength, etc.) are not added in the calculation when the working temperature exceed to 1500K in order to reduce the intolerable time cost. The structure will

be cracked in such an extreme environment under the alternate loading over the long period of time.



(a) α=0° (b) α=5° (c) α=10° (d) α=15°

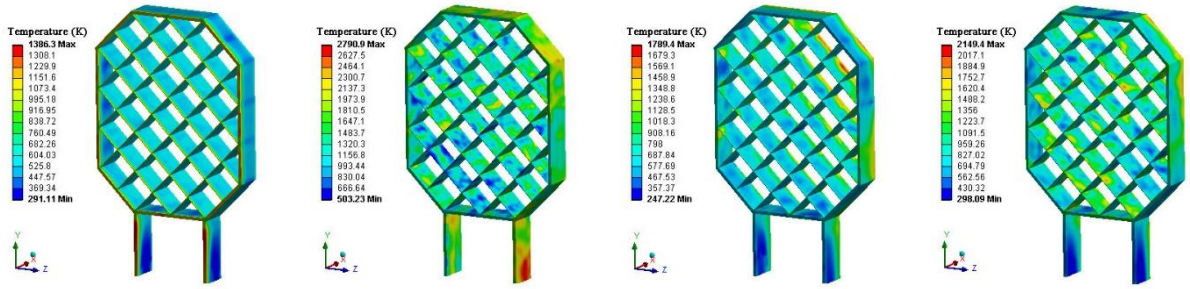
Figure 10: Displacement nephogram of niobium alloy in different angles of attack

It is with these thoughts in mind that the niobium alloy will no longer take into account due to the excess of its working temperature and large displacement caused by the high pressure and temperature environment, so some other materials should be sought to satisfy this environment. Based on the preliminary calculation results before, the C/SiC composites is used to replace the niobium alloy. It can almost meet all the need of hypersonic application: long life (under 1950K), finite life (under 2300K), and transient life (under 3100K). Here are some basic material properties listed in the following table:

Table 2: Material properties of C/SiC composites

Temperature K	Density kg/m^3	Thermal conductivity W/(m·K)	Elasticity modulus GPa	Poisson ratio	Linear expansion coefficient $\times 10^{-6}$(K^{-1})	Specific heat J/(kg·K)
295.15	2000	8	124	0.06	0.38	800
300	2000	8	112	0.06	3.09	900
600	2000	8	102	0.06	3.72	1200
900	2000	8	83	0.06	3.41	1500
1100	2000	8	91	0.06	4.49	1700
1300	2000	8	68	0.06	3.68	1700
1390	2000	8	33	0.06	4.01	1800

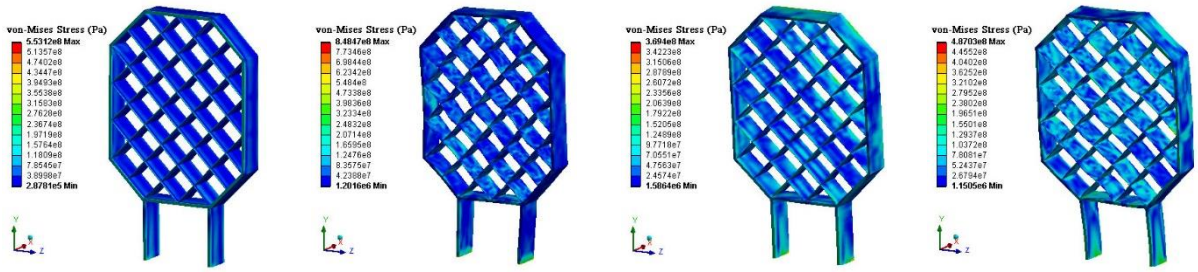
Through the comparison of temperature distribution between Figure 8 and 11, it was found that the maximum temperature of C/SiC composites is about 259K, 214K, 180K and 190K higher than the niobium alloy. It's mainly because of their closely thermal conductivity and a lower heat capacity compared with niobium alloy.



(a) α=0° (b) α=5° (c) α=10° (d) α=15°

Figure 11: Temperature nephogram of C/SiC composites in different angles of attack

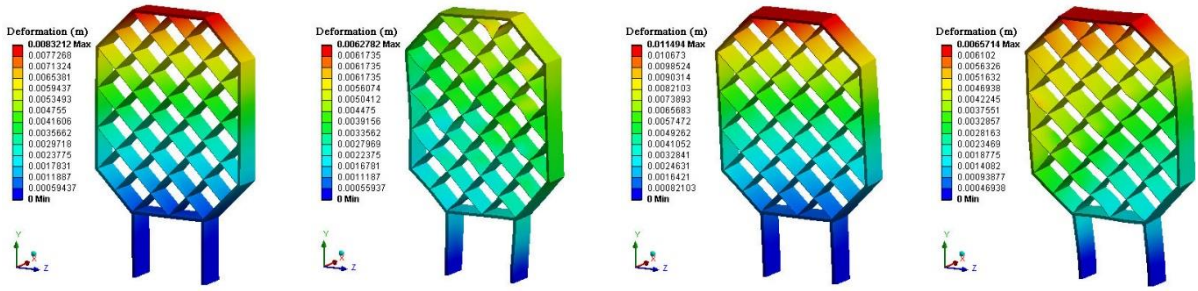
Figure 12 shows the final stress distribution of the C/SiC composites structure. Although the elasticity modulus of C/SiC composites has fallen by almost 70% in comparison to the property under ambient temperature, however, the slow growth of expansion coefficient with temperature rise will prevent the increase of thermal stress and displacement, hence, the stress level in Figure 12 is about half lower than the niobium alloy.



(a) α=0° (b) α=5° (c) α=10° (d) α=15°

Figure 12: Von-Mises stress nephogram of C/SiC composites in different angles of attack

The displacement in Figure 13 exhibits that the deformation of C/SiC composites structure is 40%, 57%, 35% and 27% lower than niobium alloy. In addition, it is noticed that, in 5° and 15° attack angle, the bottom holder happens some lateral displacement influenced by the mixture load, the supporter should be enhanced to deter the lateral load.



(a) α=0° (b) α=5° (c) α=10° (d) α=15°

Figure 13: Displacement nephogram of C/SiC composites in different angles of attack

4 CONCLUSIONS

This paper has analyzed the characteristics of fluid-thermal-structure interaction of grid fin under hypersonic conditions and makes comparison with former researches to verify the reliability of the calculation scheme. Following conclusions can be reached:

(1) The value of normal force coeffecient is very close to the calculation results of the literature, and the reliability and effectiveness of the model is verified by simulations.

(2) With the increasing of attack angle, the most serious region of aerodynamic pressure and heat flux no longer concentrate on the stagnation point of leading edge, the temperature, stress and displacement distribution changed followed by. The maximum value of temperature, stress and deformation occurs at the 5° attack angle.

(3) The deformation and stress level of C/SiC composites are better than the niobium alloy, however, in this situation, the composites not fit for long time using under a fix attack angle.

(4) The current configuration exhibits good resistance to bending, meanwhile, some lateral area should be paid attention especially for the root of bottom holder.

The results provide a reference for the design of the heat-resistant structure of grid fin.

REFERENCES

[1] Washington, W.D. and Miller, M.S. *Grid fins - a new concept for missile stability and control*. Reno: 31st Aerospace Sciences Meeting & Exhibit, (1993).

[2] Belotserkovskiy, S.M., Odnovol, L.A., Safin, Y.Z. and et al. *Wings with internal framework*. Mashinostroeniye, Vol. I, (1987).

[3] Sun, Y. and Khalid, M. *A cfd investigation of grid fin missiles*. Cleveland: 34th AIAA/ASME/SAE/ASEE Joint Propulsion Conference and Exhibit, (1998).

[4] Theerthamalai, P. Aerodynamic characterization of grid fins at subsonic speeds. *J. Aircr.* (2007) **44**:694-697.

[5] Simpson, G.M. and Sadler, A.J. *Lattice controls: A comparison with conventional, planar fins*. Sorrento: the RTO AVT Symposium, (1998).

[6] Reynier, P., Reisch, U., Longo, J. and et al. *Numerical study of hypersonic missiles with lattice wings using an actuator disk*. Missouri: 20th AIAA Applied Aerodynamics Conference, (2002).

[7] Schülein, E. and Guyot, D. *Novel high-performance grid fins for missile control at high speeds: Preliminary numerical and experimental investigations*. Neuilly-sur-Seine: RTO-MP-AVT, (2006).

[8] Wang, D. and Yu, Y. *Numerical study on drag reduction for swept-back, swept-front , delta grid fin with blunt and sharp leading edges*. National Harbor: AIAA Modeling and Simulation Technologies Conference, (2014).

[9] Mathews, I.C. Numerical techniques for three‐dimensional steady‐state fluid‐structure interaction. *J. Acoust. Soc. Am.* (1986) **79**:1317-1325.

[10] Whitehead, D.S. Vibration of cascade blades treated by actuator disc methods. *P. I. Mech. Eng.* (1959) **173**:555-574.

[11] Bendiksen, O. *A new approach to computational aeroelasticity*. Baltimore: 32nd Structures, Structural Dynamics, and Materials Conference, (1991).

[12] Takashi, N. and Hughes, T.J.R. An arbitrary lagrangian-eulerian finite element method for interaction of fluid and a rigid body. *Comp. Meth. Appl. Mech. Eng.* (1992) **95**:115-138.
[13] Choi, H.G. and Joseph, D.D. Fluidization by lift of 300 circular particles in plane poiseuille flow by direct numerical simulation. *J. Fluid Mech.* (2001) **438**:101-128.
[14] Hubner, B., Walhorn, E. and Dinkier, D. *Strongly coupled analysis of fluid-structure interaction using space-time finite elements*. Cracow: 2nd European Conference on Computational Mechanics, (2001).
[15] Hubner, B., Walhorn, E. and Dinkler, D. *Simultaneous solution to the interaction of wind flow and lightweight membrane structures*. Warsaw: Proceedings of International Conference on Lightweight Structures in Civil Engineering, (2002).
[16] Hübner, B., Walhorn, E. and Dinkler, D. A monolithic approach to fluid–structure interaction using space–time finite elements. *Comp. Meth. Appl. Mech. Eng.* (2004) **193**:2087-2104.
[17] Namkoong, K., Choi, H.G. and Yoo, J.Y. Computation of dynamic fluid–structure interaction in two-dimensional laminar flows using combined formulation. *J. Fluid Struct.* (2005) **20**:51-69.
[18] Zhang, Q. and Hisada, T. Studies of the strong coupling and weak coupling methods in fsi analysis. *Int. J. Numer. Meth. Eng.* (2004) **60**:2013-2029.
[19] Culler, A.J. *Coupled fluid-thermal-structural modeling and analysis of hypersonic flight vehicle structures*. PhD: The Ohio State University, (2010)

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis(Eds)

MESOSCOPIC MODELING OF THE RTM PROCESS FOR HOMOGENIZATION

C. Dammann*, R. Mahnken†

*, †Chair of Engineering Mechanics (LTM)
University of Paderborn
Warburger Str. 100, 33098 Paderborn, Germany
*e-mail: dammann@ltm.upb.de
†e-mail: mahnken@ltm.upb.de, web page: http://www.ltm.upb.de/

Key words: Curing, shrinkage, thermal-mechanical-chemical coupling, homogenization, multi-scale modeling, finite elements

Abstract.

Intrinsic hybrids can be manufactured in a modified resin transfer molding (RTM) process for fibre reinforced polymers. Our work concentrates on mesoscopic modeling for temperature-dependent visco-elastic effects accompanied by curing within the RTM process. During hybridization and later thermal loading the periodic mesostructure defined by resin and fibres is taken into account as a representative volume element (RVE) subjected to thermo-mechanical loading. Homogenization leads to results on the less resolved macroscale. In the examples we illustrate the characteristic behavior of the mesoscopic model, such as shrinking due to curing and temperature dependence and simulate the RTM process as well as thermal loading of the cured composite with the finite-element-method.

1 INTRODUCTION

Nowadays, polymeric materials find their applications in carbon- and glass fibre-reinforced epoxy laminates. The production process of polymeric materials is mainly characterised by thermal loading and curing. Here, in the initial uncured state the mixture of resin and curing agent exhibits a viscous liquid behavior allowing no more than hydrostatic pressure. With evolving curing polymer chains form and cross-link to each other such that the viscosity of the liquid resin, its molecular weight and the stiffness increase, see e.g. [1]. The process is highly temperature dependent and influences strongly the mechanical, thermal and chemical properties of the final composite.

In [2] a phenomenological thermo-viscoelastic curing model for finite strain deformations is proposed. The formulation is based on process dependent viscosities.

Furthermore, [3], [4] and [5] formulate ad hoc assumptions for the bulk heat-dilatation coefficient dependent on curing, whereas the bulk shrinking-dilatation coefficient is not dependent on the degree of curing.

A micro-macro strategy suitable for modeling the mechanical response of heterogeneous materials at large deformations and non-linear history dependent material behaviour is presented in [6]. [7] investigates an algorithm for the computation of homogenized stresses by volume averaging and the overall tangent moduli of microstructures undergoing small strains is derived. In [8], the influence of the RVE size on the residual stresses created during the curing process of a continuous fibre-reinforced polymer matrix tow is investigated to calculate the response of woven fibre textile composites. Additionally, mechanical loading for varying RVE size is investigated, not including the effect of curing stresses, or in other words not including the process history.

Our work enhances these investigations, concentrating on the macroscopic anisotropic thermal expansion coefficient including the history of a modfied RTM process. Within a thermodynamic framework we use an additive ternary decomposition of the logarithmic Hencky strain tensor into mechanical, thermal and chemical parts, as in [4]. Based on the concept of stoichiometric mass fractions [9] for resin, curing agent and solidified material the bulk compression modulus as well as the bulk heat- and shrinking dilatation coefficients are derived. An RVE simulating the RTM process is used to determine the macroscopic strains by homogenization. Mesoscopic as well as macroscopic residual strains can be observed for the fully cured material. Based on this foundings, the macroscopic anisotropic thermal extension is determined.

Notations

Square brackets $[\bullet]$ are used throughout the paper to denote 'function of' in order to distinguish from mathematical groupings with parenthesis $(\bullet)$.

2 A GENERAL FRAMEWORK FOR FIBRE REINFORCED POLYMER CURING

Within the framework of large strain theory $\mathcal{B}_0 \subset \mathcal{R}^3$ in Figure 1.b denotes the periodic mesostructure of a fibre reinforced composite in the reference configuration. It is related to the homogenized macrocontinuum in the reference configuration $\overline{\mathcal{B}}_0 \subset \mathcal{R}^3$, as shown in Figure 1.a. $\mathcal{V}_0$ in Figure 1.c denotes the RVE associated with the structure $\mathcal{B}_0$. The volume $\mathcal{V}_0$ consists of the resin part $\mathcal{B}_{0resin}$ and the fibre part $\mathcal{B}_{0fibre}$, respectively, each regarded as solid constituents. Thus, we consider the decompositions of the RVE and its surface $\mathcal{V}_0 = \mathcal{B}_{0resin} \cup \mathcal{B}_{0fibre}$. We denote by $P_i \subset \mathcal{B}_{0i}$, a spatial point for which either $i = resin$ or $i = fibre$ holds. Its material counterpart at time t in the current configuration of the mesostructure $\mathcal{B}_i,\ i = resin,\ fibre$ in the space-time-domain $\mathcal{B}_0 \times\ ]-\infty, T[$ is denoted by $p_i,\ i = resin,\ fibre$, where T is the total time of interest.

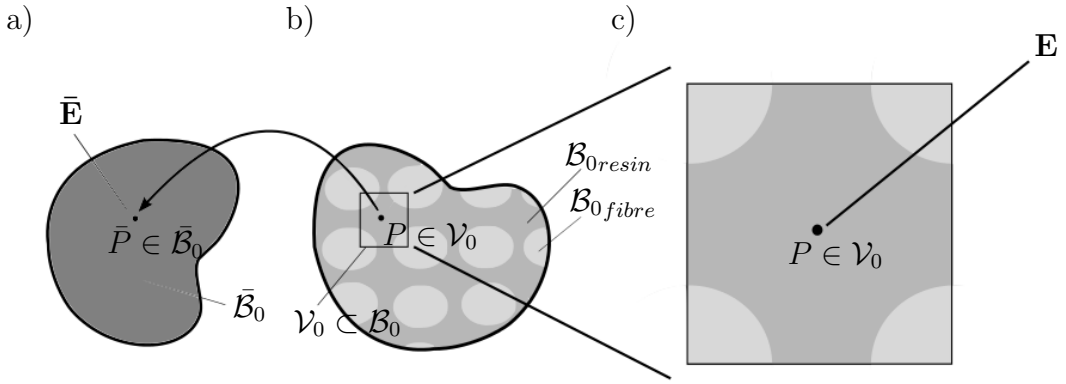


Figure 1: Macrocontinuum with a mesostructure: a) Homogenous macrocontinuum $\bar{\mathcal{B}}_0$ and strain $\bar{\mathbf{E}}$ at point $\bar{P}$, b) periodic mesostructure $\mathcal{B}_0$ and RVE $\mathcal{V}_0$, c) RVE with strain $\mathbf{E}$ at point P.

Additionally, the deformation gradient $\mathbf{F}_i,\ i = resin,\ fibre$ is introduced at each material point p_i. It maps line segments $d\mathbf{X}_i$ of the reference configuration $\mathcal{B}_{0i}$ to line segments $d\mathbf{x}_i$ of the current configuration $\mathcal{B}_i$. Additionally, we introduce its Jacobian $J_i = \det \mathbf{F}_i$, mapping a volume element dV_i of the reference configuration $\mathcal{B}_{0i}$ to a volume element dv_i of the current configuration $\mathcal{B}_i$:

$$1.\ d\mathbf{x}_i = \mathbf{F}_i \cdot d\mathbf{X}_i, \qquad 2.\ dv_i = J_i dV_i, \qquad i = resin,\ fibre. \tag{1}$$

A key point in constructing a framework of finite plasticity is the definition of the total Hencky strain tensor $\mathbf{E}_i$ in a logarithmic form, see e.g. [10, 11].We assume this strain measure to be a function of the right Cauchy-Green tensor

$$1.\ \mathbf{C}_i = \mathbf{F}_i^t \cdot \mathbf{F}_i, \qquad 2.\ \mathbf{E}_i = \frac{1}{2} \ln \mathbf{C}_i, \qquad i = resin,\ fibre. \tag{2}$$

The volumetric and deviatoric parts of the Hencky strain tensor are defined as

$$1.\quad \mathbf{E}_i^{vol} = \frac{1}{3} \mathrm{tr}\mathbf{E}_i\ \mathbf{1}, \qquad 2.\quad \mathbf{E}_i^{dev} = \mathbf{E}_i - \mathbf{E}_i^{vol}, \qquad i = resin,\ fibre. \tag{3}$$

3 MESOSCOPIC MODELING

The models for the composite constituents are described in the following subsections 3.1 and 3.2. The resin is modeled as visco-elastic and the fibre is considered as a thermo-linear-elastic solid. For notational benefits, the index $i = resin,\ fibre$ is omitted.

3.1 Modeling of resin: visco-elasticity coupled to curing

Based on the concept of stoichiometric mass fractions as discussed extensively in [9] for resin, curing agent and solidified material the *bulk compression modulus* κ, the *bulk heat-dilatation coefficient* α and the *shrinking dilatation coefficient* β are derived. For subsequent analysis, we assume that α is dependent on the pressure p, the temperature $\theta[t]$ and the *degree of cure* $z[t]$. Additionally we assume that the *bulk curing-dilatation coefficient* β is dependent on the pressure p and the temperature $\theta[t]$.

3.1.1 Mass fractions and the degree of cure

We assume a homogeneous mixture with three constituents resin, curing agent and solidified material at each instant of time t. As a consequence there are neither concentration gradients nor diffusion effects. With the time-dependent variables $dm_r[t], dm_{ca}[t]$ and $dm_{sol}[t]$ for the masses of resin, curing agent and solidified material, respectively, the conservation of mass during the curing reaction requires

$$dm_r[t] + dm_{ca}[t] + dm_{sol}[t] = dm_0 = \text{const}, \tag{4}$$

where the constant dm_0 is the total mass of the mixture [9]. Dividing Eq.(4) by dm_0 and introducing the mass fractions of resin, curing agent and solid

$$\zeta_r[t] = \frac{dm_r[t]}{dm_0}, \quad \zeta_{ca}[t] = \frac{dm_{ca}[t]}{dm_0}, \quad \zeta_{sol}[t] = \frac{dm_{sol}[t]}{dm_0}, \tag{5}$$

we obtain the balance relation

$$\zeta_r[t] + \zeta_{ca}[t] + \zeta_{sol}[t] = 1. \tag{6}$$

As explained in [9] the number of variables can be reduced by taking into account the stoichiometry of the mixture. To this end the *degree of cure* $0 \leq z[t] \leq 1$ is introduced, see also [1], such that the mass fractions of the three components of the mixture are written in the form

$$\begin{aligned} &1.\ \zeta_r[t] &&= n(1 - z[t]) \\ &2.\ \zeta_{ca}[t] &&= (1 - n)(1 - z[t]) \\ &3.\ \zeta_{sol}[t] &&= z[t]. \end{aligned} \tag{7}$$

The initial state with $z[0] = 0$ corresponds to the uncured, viscous mixture with $\zeta_r[0] = n$, $\zeta_{ca}[0] = n - 1$, $\zeta_{sol}[0] = 0$ as initial conditions. Consequently, n and $1 - n$ are the mass fractions of the resin and the curing agent at the beginning of curing, respectively. The final fully cured state at time $t \to \infty$ with $z[\infty] = 1$ corresponds to the solidified material at the end of the reaction, i.e. $\zeta_r[\infty] = \zeta_{ca}[\infty] = 0$ and $\zeta_{sol}[\infty] = 1$.

In the subsequent exposition, occasionally an index $i = r, ca, sol$ referring to resin, curing agent and solidified material will be used. Then, Eq.(6) and the connectivity for all three constituents renders the following balances at each body point $\mathbf{X} \in \mathcal{B}_0$ and for all times $t > -\infty$:

$$1.\ \sum_{i=1}^{3} \zeta_i = 1, \qquad 2.\ \zeta_i \geq 0, \quad i = r, ca, sol. \tag{8}$$

Furthermore, from Eq.(7) we obtain the functional relation

$$\zeta_i = \zeta_i[z], \qquad i = r, ca, sol, \tag{9}$$

which means, that the mass phase fractions ζ_i are independent of temperature and deformation.

We also assume, that the mixture for the resin is homogeneous, i.e. all phases are equally distributed. Then, the (bulk) densities ρ_0 and ρ of the mixture with respect to the reference and the current configurations $\mathcal{B}_0$ and $\mathcal{B}$ are respectively defined as

$$1. \quad \varrho_0 = \frac{dm}{dV}, \qquad\qquad 2. \quad \varrho = \frac{dm}{dv}. \tag{10}$$

Here, according to Eq.(1) dv and dV are the volume differentials of the mass differential dm at the reference and the current configurations, respectively.

Within the volume dv, let the i^{th} constituent have its volume dv_i and its mass dm_i, $i = r, ca, sol$. Then, the mass phase fraction ζ_i and the density of the i^{th} constituent are defined by

$$1. \quad \zeta_i = \frac{dm_i}{dm}, \qquad\qquad 2. \quad \rho_i = \frac{dm_i}{dv_i}. \tag{11}$$

The equations (10) and (11) imply the assumption, that the quantities ρ and ρ_i at a body point $P \in \mathcal{B}_0$ are defined by a limit process with volumes contracting to this point. Using Eq.(10), the relation $dv = \sum_{i=1}^{3} dv_i$ as well as Eq.(11.2) and Eq.(11.1) renders the following mixture rule for the inverse of the bulk density ρ

$$\frac{1}{\rho} = \frac{dv}{dm} = \frac{\sum_{i=1}^{3} dv_i}{dm} = \sum_{i=1}^{3} \frac{dm_i}{\rho_i dm} = \sum_{i=1}^{3} \frac{\zeta_i}{\rho_i}. \tag{12}$$

For the subsequent analysis, we assume that the density of each constituent is solely dependent on the pressure p and the temperature, that is

$$\rho_i = \rho_i[p, \theta]. \tag{13}$$

Note, that the functional relations (9) and (13) combined with the mixture rule Eq.(12) imply the functional relation

$$\rho[t] = \rho[p[t], \theta[t], z[t]], \tag{14}$$

i.e. the density change is induced by pressure p, temperature θ and/or by curing z.

3.1.2 Volume changes due to density changes

We assume that the total volume change J in Eq.(1.2) is multiplicatively decomposed into an effective part (induced by pressure p), a thermal part (induced by temperature θ) and a chemical part (induced by curing z as a result of shrinkage due to polymerization), that is

$$J = \frac{dv}{dV} = \frac{\rho_0}{\rho} = J^p \cdot J^\theta \cdot J^z. \tag{15}$$

Using the relation (12) and exploiting the functional relation (13), the time derivative of J for the third part in Eq. (15) is expressed as follows:

$$\dot{J} = \frac{d}{dt}\left(\frac{\rho_0}{\rho}\right) = \rho_0 \frac{d}{dt}\left(\sum_{i=1}^{3} \frac{\zeta_i}{\rho_i}\right) = -\sum_{i=1}^{3} \rho_0 \frac{\zeta_i}{\rho_i^2}\frac{\partial \rho_i}{\partial p}\ \dot{p} - \sum_{i=1}^{3} \rho_0 \frac{\zeta_i}{\rho_i^2}\frac{\partial \rho_i}{\partial \theta}\ \dot{\theta} + \sum_{i=1}^{3} \frac{\rho_0}{\rho_i}\frac{\partial \zeta_i}{\partial z}\ \dot{z}. \quad (16)$$

Next we define the isothermal compressibility κ, the (uni-directional) bulk heat-dilatation coefficient α and the (uni-directional) bulk curing-dilatation coefficient β:

$$\begin{aligned}
&1.\quad \kappa[p,\theta,z] = \sum_{i=1}^{3} \kappa_i[p,\theta]\zeta_i, \quad \text{where} \quad 2.\quad \kappa_i[p,\theta] = \frac{\rho_0}{\rho_i^2}\frac{\partial \rho_i}{\partial p}\\
&3.\quad \alpha[p,\theta,z] = \sum_{i=1}^{3} \alpha_i[p,\theta]\zeta_i, \quad \text{where} \quad 4.\quad \alpha_i[p,\theta] = \frac{-\rho_0}{3\rho_i^2}\frac{\partial \rho_i}{\partial \theta}\\
&5.\quad \beta[p,\theta] = \sum_{i=1}^{3} \beta_i[p,\theta]\frac{\partial \zeta_i}{\partial z}, \quad \text{where} \quad 6.\quad \beta_i[p,\theta] = \frac{\rho_0}{3\rho_i}.
\end{aligned} \quad (17)$$

For the mass fractions of the three constituents in Eq.(7), the bulk heat-dilatation coefficient in Eq.(17.3) is

$$\alpha[p,\theta,z] = \alpha_r[p,\theta]\ n(1-z) + \alpha_{ca}[p,\theta]\ (1-n)(1-z) + \alpha_{sol}[p,\theta]\ z, \quad (18)$$

where $\alpha_r[p,\theta]$, $\alpha_{ca}[p,\theta]$ and $\alpha_{sol}[p,\theta]$ are defined according to Eq.(17.4). In general, for increasing temperature θ the densities ρ_i decrease, i.e. $\partial\rho_i/\partial\theta$ is negative, such that $\alpha_i[p,\theta], i = r,\ ca,\ sol$ are positive. The bulk curing-dilatation coefficient β in Eq.(17.5) is

$$\beta[p,\theta] = -n\beta_r[p,\theta] - (1-n)\beta_{ca}[p,\theta] + \beta_{sol}[p,\theta], \quad (19)$$

where $\beta_r[p,\theta]$, $\beta_{ca}[p,\theta]$ and $\beta_{sol}[p,\theta]$ are defined according to Eq.(17.6). In general we have $\rho_{sol} > \rho_r$ and $\rho_{sol} > \rho_{ca}$, such that $\beta[p,\theta]$ is negative.

3.1.3 Strains and stresses

The thermal and the curing strains

$$\begin{aligned}
&1.\quad \mathbf{E}_{th} = \int_{s=-\infty}^{t} \mathbf{E}_{th} = \int_{s=-\infty}^{t} \alpha[s]\frac{d\theta}{ds}ds\mathbf{1}\\
&2.\quad \mathbf{E}_{cur} = \int_{s=-\infty}^{t} \beta[s]\frac{dz}{ds}ds\mathbf{1},
\end{aligned} \quad (20)$$

are governed by α and β which are in accordance with Eqn.(17.3,17.5). In [4], the effective strain tensor is obtained with $\mathbf{E}$ from Eq.(2.2):

$$\mathbf{E}^p = \mathbf{E} - \mathbf{E}_{th} - \mathbf{E}_{cur}. \quad (21)$$

Additionally, we define the Hill-stress tensor $\mathbf{T} = \mathbf{T}_{vol} + \mathbf{T}_{dev}$ with volumetric and deviatoric parts

$$\begin{aligned} 1. \quad \mathbf{T}_{vol} &= \int_{s=-\infty}^{t} \frac{1}{\kappa[s,t]} \frac{d}{ds} \mathrm{tr}[\mathbf{E}^p[s]] \; \mathbf{1} \; ds \\ 2. \quad \mathbf{T}_{dev} &= \int_{s=-\infty}^{t} 2G[s,t] \frac{d}{ds} \mathbf{E}^{dev}[s] ds. \end{aligned} \tag{22}$$

The compressibility κ in Eq.(22.1) is obtained in the same fashion as α in Eq.(18), and G Eq.(22.2) is the shear modulus, in detail described in [4].

3.2 Modeling of fibre: thermo-linear-elasticity

The fibre is considered as a thermo-linear-elastic solid. Consequently it can be modeled as a special case with the equations of subsection 3.1. For example, taking into account Eq.(21) with $\zeta_{sol} = 1$ and $\mathbf{E}_{cur} = 0$ in Eq.(20.2), Eq.(21) reduces to $\mathbf{E}^p = \mathbf{E} - \mathbf{E}_{th}$.

4 MESO TO MACRO TRANSITION

The mesostructure variables are related to the homogenized macrocontinuum $\bar{\mathcal{B}}_0 \subset \mathcal{R}^3$ by use of the Volume Averaging Theorem, as formulated in [12]. Based on the decomposition $\mathcal{V}_0 = \mathcal{B}_{0resin} \cup \mathcal{B}_{0fibre}$ in Section 2 we define the overall macro-strain $\bar{\mathbf{E}}(t)$ of the mesostructure $\mathcal{B}_0$ in relation to the strains from Eq.(2.2) as well as the general homogenization operator $\bar{\bullet}$ by

$$1. \;\; \bar{\mathbf{E}} = \frac{1}{|\mathcal{V}_0|} \int_{\mathcal{V}_0} \mathbf{E} dV, \qquad 2. \;\; \bar{\bullet} = \frac{1}{|\mathcal{V}_0|} \int_{\mathcal{V}_0} \bullet dV. \tag{23}$$

5 NUMERICAL SIMULATION

In this section an RVE is simulated in two loading sequences. The first sequence simulates the RTM process. This leads to residual strains on the mesoscale. Macroscopic strains are determined by homogenization. In the second sequence, based on the process history of the fully cured composite from sequence one, the *macroscopic anisotropic thermal expansion coefficient* is determined.

5.1 Thermal-mechanical-chemical coupling in an RVE during hybridization

In this example some features on the coupling of temperature, curing and visco-elasticity of the model proposed in Section 3 are illustrated for a mesoscopic problem and its homogenized macroscopic answer occuring in the RTM process. To model the three dimensional RVE shown in Figure 2 and representing the composite material on the mesoscale, stiff fibres (bright) are embedded in a soft resin matrix (dark), thus defining $\mathcal{B}_{0fibre}$ and $\mathcal{B}_{0resin}$. The cube-shaped RVE has a fibre-volume-fraction of 19.6 %. The

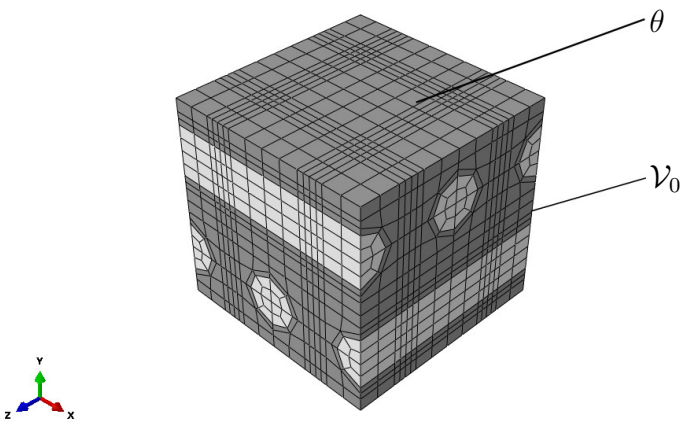


Figure 2: Statically determined RVE $\mathcal{V}_0$: Geometry, FE-discretization and thermal loading θ.

fibres are assumed thermo-elastic, while the matrix material is modeled visco-elastic, see Section 3. Fictive material parameters are chosen. The boundary conditions for the RVE are statically determined to allow thermal expansion, or shrinkage, respectively. Since temperature gradients occuring in the RVE are neglected, the RTM process is assumed to be a pure homogenous thermal loading θ of the RVE denoted as HEAT, CURE and COOL in Figure 2. In the first phase the temperature is increased with constant rate from the initial value $\theta = 25\ °C$ up to $\theta = 120\ °C$. During the second phase, the temperature is kept fixed. In the third phase the temperature is decreased at constant rate to the initial value of $\theta = 25\ °C$.

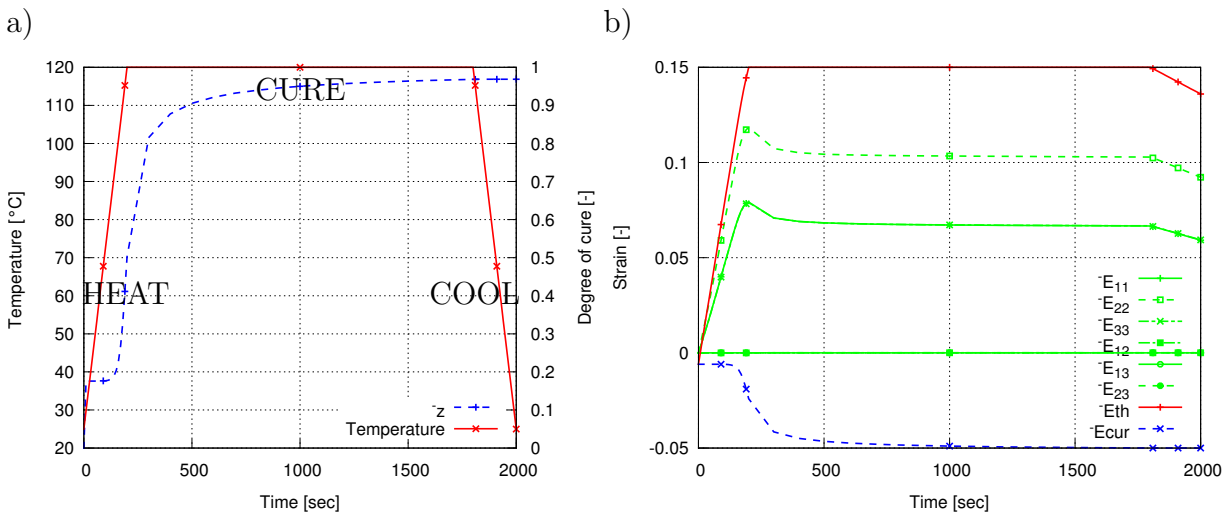


Figure 3: RVE during the RTM process: a) Thermal loading and homogenized degree of cure vs. time, b) components of homogenized strain tensor $\bar{\mathbf{E}}$ and homogenized strains $\bar{\mathbf{E}}_{th}$ and $\bar{\mathbf{E}}_{cur}$ vs. time.

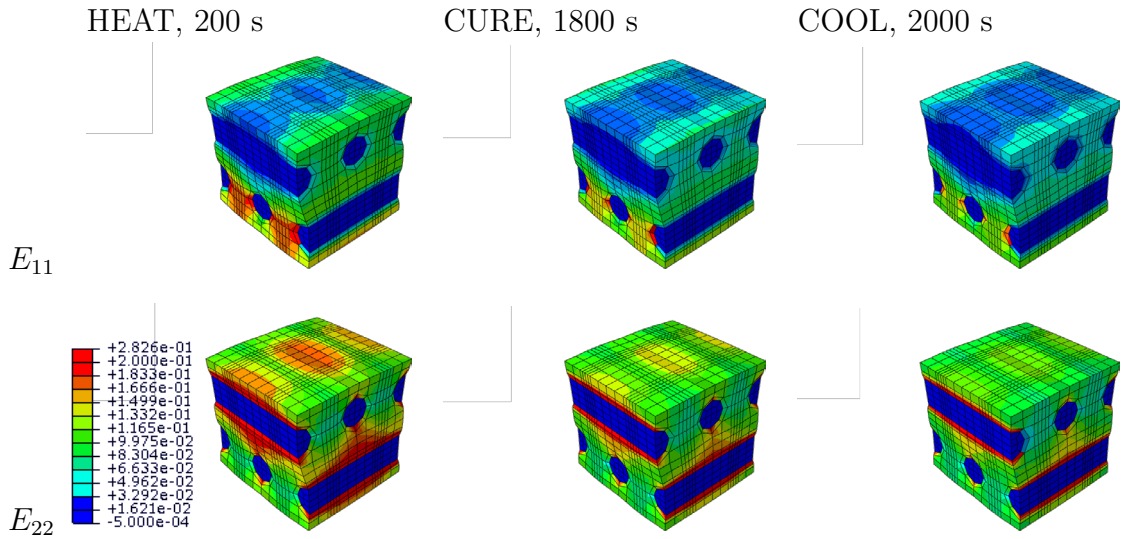


Figure 4: RVE during the RTM process: Contourplots showing strains E_{11} and E_{22} during the three phases of the loading program in Figure 3.a

As mesoscopic strains show symmetries for the x- and z-direction, it is sufficient showing the contourplots in Figure 4 only for the unsymmetric x- and y-directions. The strain contourplots offer residual strains as a remaining strain state for the fully cured composite at the end of phase COOL. This is due to the change of the thermal-expansion coefficient as a result of curing, see Eqn.(18,20.1).

In Figure 3.a-b we summarize some relevant macroscopic quantities resulting from the mechanical-thermal-chemical coupling. The influence of curing and temperature on the homogenized total strains $\bar{\mathbf{E}}$ Eq. (23.1) can be observed. The strains offer a macroscopic orthotropic deformation behavior with a symmetry in x- and z-direction and residual strains at the end of phase COOL. Applying Eq. (23.2), homogenization leads to results for the degree of cure $\bar{z}$ as well as for the uniaxial thermal and shrinking strains $\bar{\mathbf{E}}_{th}$ and $\bar{\mathbf{E}}_{cur}$ shown in Figure 3.a-b. The reason for an initial value for the homogenized degree of cure $\bar{z}_0 > 0$ is that, in contrast to the resin, the fibres initial condition on the mesoscale is $z_0 = 1$. We want to remark, that during the COOL phase linear strain behavior can be observed in Figure 3.b because curing is almost completed at 1600 s. In contrast, in the preceding phases curing governs nonlinear strains.

5.2 Macroscopic thermal expansion coefficient of the fully cured composite

The aim of this example is the determination of the macroscopic thermal expansion coefficient. Thus, the fully cured composite offers a strain state related to residual strains, as shown in Section 5.2. The subsequent loading program shown in Figure 5.a is divided into two phases HOLD and HEAT2 following the three phases in Section 5.2. During the fourth phase the temperature is kept fixed. In the fifth phase the temperature is increased by $\Delta\theta = 1\ ^\circ C$. The motivation for the HOLD phase is found in relaxation of strains of

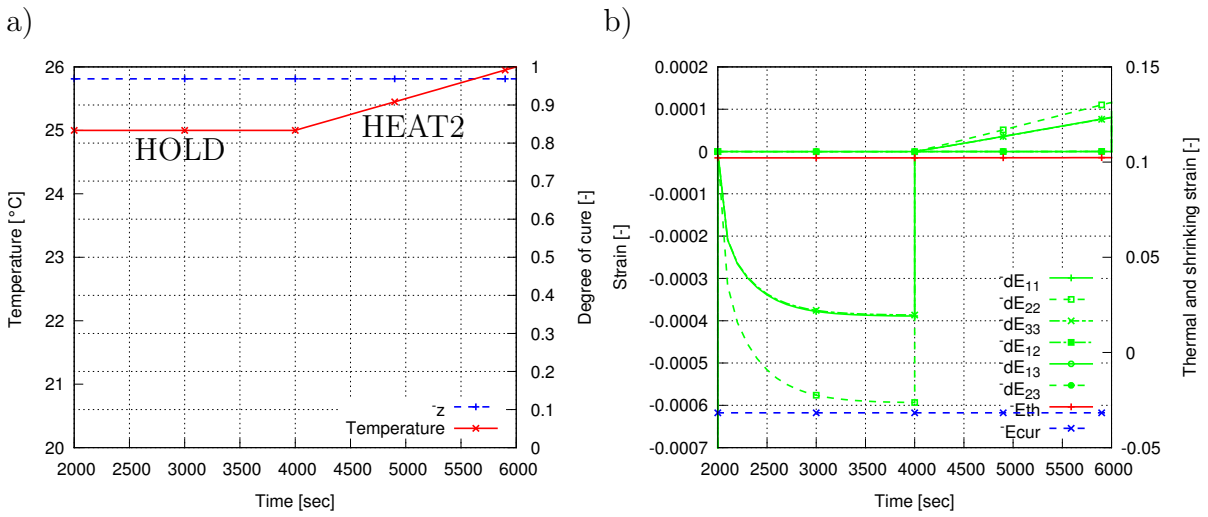


Figure 5: Thermal loading of an RVE: a) Thermal loading and homogenized degree of cure vs. time, b) tensor of homogenized strain increment per phase $\Delta\bar{\mathbf{E}}$ and homogenized strains $\bar{\mathbf{E}}_{th}$ and $\bar{\mathbf{E}}_{cur}$ vs. time.

the visco-elastic material on the mesoscale. To observe this small effect, the mesoscopic strain state at every material point of the mesoscale beyond the COOL phase is divided from the mesoscopic strain state beyond the HOLD phase, resulting in the incremental strains on the mesoscale. Homogenization then leads to

$$\Delta\bar{\mathbf{E}}_{\mathrm{HOLD}} = \bar{\mathbf{E}}_{\mathrm{HOLD}} - \bar{\mathbf{E}}_{\mathrm{COOL}}. \tag{24}$$

$\bar{\mathbf{E}}_{\mathrm{HOLD}}$ and $\bar{\mathbf{E}}_{\mathrm{COOL}}$ are quantities obtained by use of Eq.(23.1) at times $t = 4000$ s and $t = 2000$ s, respectively. As the strain increments $\Delta\bar{\mathbf{E}}$ show asymptotic behavior beyond 3500 s, the HEAT2 phase is started at $t = 4000$ s. Due to very slow heating and small temperatur variation of $\Delta\theta = 1\,°C$ we exclude viscoelastic effects. Thus, the homogenized strain increment $\Delta\bar{\mathbf{E}}$ is directly related to the *anisotropic macroscopic thermal expansion coefficients*

$$\Delta\bar{\mathbf{E}}_{\mathrm{HEAT2}} = \bar{\mathbf{E}}_{\mathrm{HEAT2}} - \bar{\mathbf{E}}_{\mathrm{HOLD}} = \bar{\alpha}. \tag{25}$$

The coefficients of $\bar{\alpha}$ can directly be obtained from Figure 5.b as

$$\bar{\alpha}^T = [8.09 \cdot 10^{-05}, 1.16 \cdot 10^{-04}, 8.02 \cdot 10^{-05}, 1.84 \cdot 10^{-08}, -1.38 \cdot 10^{-11}, 1.86 \cdot 10^{-08}]. \tag{26}$$

As expected, they show orthotropic behavior for the x- and z- components and, except numerical inexactness, shear components xy, xz, $zy = 0$. Thus, the strains due to the HEAT2 phase are very small. Contourplots for the HEAT2 phase displaying strain increments on the mesoscale are illustrated in Figure 6. We want to remark, that the displayed strains $\Delta\mathbf{E}$ are therefore residual strain free, in contrast to those of Figure 4.

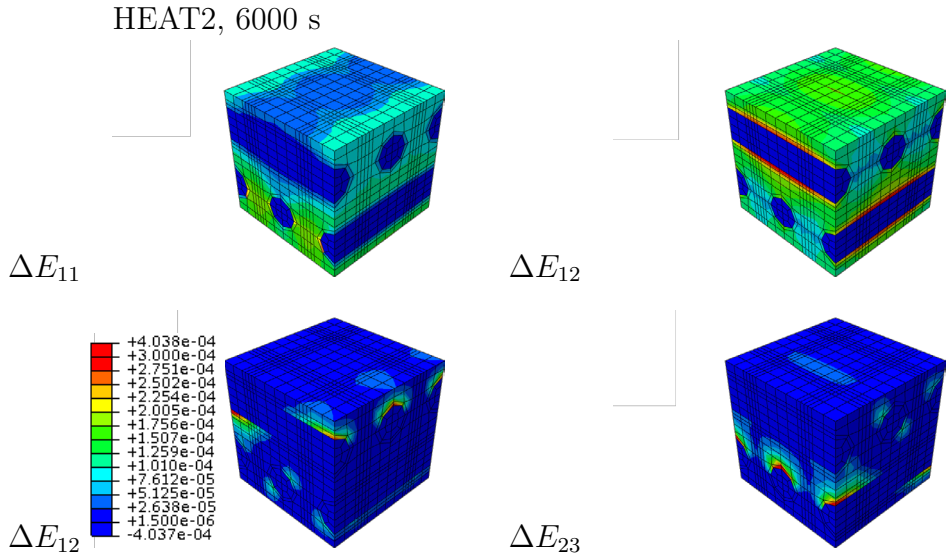


Figure 6: Thermal loading of the RVE: Contourplots showing strain increments ΔE_{11} , ΔE_{22} , ΔE_{12} and ΔE_{23} due to thermal loading HEAT2 following on the HOLD phase according to Figure 5.a

6 CONCLUSIONS

In this paper, we have developed a mesoscopic model for temperature-dependent viscoelastic effects accompanied by curing of fibre reinforced composites, which are important phenomena in production processes. The logarithmic Hencky strain tensor constitutes the basis for the large strain formulation and is additively decomposed into a mechanical, a thermal and a chemical part for the resin as for the fibre. Based on the concept of stoichiometric mass fractions for resin, curing agent and solidified material the bulk heat- and shrinking dilatation coefficients as well as a linear dependence with respect to the degree of curing for the compressibility are derived. The meso to macro transition is treated by use of the Volume Averaging Theorem. In the examples we illustrate the anisotropic shrinking and thermal expansion due to curing as well as temperature dependence. The hybridization process as a modified RTM process is simulated with the finite-element-method for an RVE. Residual process dependent strain states are observed on the mesoscopic as well as on the macroscopic scale for the fully cured composite, which effect later use. Based on the process history, the macroscopic thermal expansion coefficient is determined by further thermal loading of the fully cured RVE. The coefficient mainly shows thermo-elastic orthotropic deformation behavior.

Concerning further extensions, taking a Navier-Stokes equation for the liquid phase of the forming process into account is an area of future research work in the field of fluid-structure interaction with phase transitions.

ACKNOWLEDGEMENT

This work is based on investigations of the "SPP 1712 - Intrinsische Hybridverbunde für Leichtbautragstrukturen", which is kindly supported by the Deutsche Forschungsgemeinschaft (DFG).

REFERENCES

[1] Kiasat, M. *Curing shrinkage and residual stresses in viscoelastic thermosetting resins and composites*. PhD Thesis, TU Delft, Netherlands (2000).

[2] Hossain, M., Possart, G., Steinmann, P. A finite strain framework for the simulation of polymer curing. *Comput. Mech.* (2009) **44**:621–630.

[3] Ruiz, E., Trochu, F. Thermomechanical properties during cure of glass-polyester RTM composites: elastic and viscoelastic modeling. *J. Compos. Mater.* (2005) **39**:881–916.

[4] Mahnken, R. Thermodynamic consistent modeling of polymer curing coupled to visco-elasticity at large strains. *Int. J. Solids Struct.* (2013) **50**:2003–2021.

[5] Hill, R., Muzumdar, S.V., Lee, L.J. Analysis of Volumetric Changes of Unsaturated Polyester Resins During Curing. *Polym. Eng. Sci.* (1995) **35**(10):852–859.

[6] Kouznetsova, V., Brekelmans, W. A. M., Baaijens, F.P.T. An approach to micro-macro modeling of heterogeneous materials. *Comput. Mech.* (2001) **27**:37–48.

[7] Miehe, C., Koch, A. Computational micro-to-macro transitions of discretized microstructures undergoing small strains- *Arch. Appl. Mech.* (2002) **72**:300–317.

[8] Heinrich,C., Aldridge, M., Wineman, A.S.,Kieffer, J,Waas, A.M. and Shahwan,K. The influence of the representative volume element (RVE) size on the homogenized response of cured fiber composites. *Modell. Simul. Mater. Sci.* (2012) **20**:75007.

[9] Lion, A., Höfer, P. On the phenomenological representation of curing phenomena in contiuum mechanics. *Archive of Mechanics* (2007) **59**:59–89.

[10] Hencky, H. Über die Form des Elastizitätsgesetzes bei ideal elastischen Stoffen. *Zeitschrift für Technische Physik* (1928) **9**:215–220.

[11] Seth, B.R. *Generalized strain measure with application to physical problems*. Reiner, M., Abir, D., Second Order Effects in Elasticity, Plasticity and Fluid Dynamics (1964).

[12] Hill, R. On constitutive macrovariables for heterogeneous solids at finite strain. *Proc. R. Soc. London., A* (1972) **326**(1565):131–147.

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis(Eds)

MIGRATION OF A RIGID DISC IN COUETTE FLOW SUBJECT TO AN EXTERNAL ELECTRIC FIELD SIMULATED USING ISPH

NIMA TOFIGHI[†], MURAT OZBULUT[†], AMIN RAHMAT[†], JAMES J. FENG[‡] AND MEHMET YILDIZ[†]

[†]Faculty of Engineering and Natural Sciences (FENS), Sabanci University,
Orhanli, Tuzla, 34956 Istanbul, Turkey
e-mails: nima@sabanciuniv.edu
ozbulut@sabanciuniv.edu
arahmat@sabanciuniv.edu
meyildiz@sabanciuniv.edu

[‡]Department of Mathematics and Department of Chemical and Biological Engineering,
University of British Columbia, Vancouver, BC V6T 1Z2, Canada
e-mail: james.feng@ubc.ca

Key words: Smoothed Particle Hydrodynamics, Fluid Solid Interaction, Electrohydrodynamics, Disc Migration

Abstract. Incompressible smoothed particle hydrodynamics method has been used to simulate the migration of a two-dimensional rigid disc in Couette flow in presence of an external electric field. The electric field is perpendicular to the moving walls. The results show that the trajectory of the disc is affected by the electrical properties of fluid and solid.

1 INTRODUCTION

The interaction of a solid body with a fluid environment is one of the most common flow features in nature and industry. When exposed to an external electric field, the motion of the solid may undergo significant changes. Migration of rigid discs in Couette flow has been the subject of many studies [1–3]. On the other hand, the rotation of spherical bodies in electric field known as Quincke rotation [4], has attracted much attention [5, 6]. It is known that a particle laden fluid may exhibit changing viscosity when exposed to an external electric field [7, 8]. This behavior is attributed to changing motion of particles due to the electric field. However, a fully resolved simulation of a disc migrating in electric field has not been conducted.

In this study, a two-dimensional Incompressible Smoothed Particle Hydrodynamics (ISPH) scheme is used to simulate the motion of a rigid disc in Couette flow subject to an external electric field [9]. Solid and fluid phases are modeled as leaky dielectric materials [10] while the electric field is taken to be perpendicular to the moving walls and in-plane with the flow. Numerical simulation of a single disc placed off-center in Couette flow is carried out and the results are validated against literature data in the absence of an electric field. Then the same case at different electric permittivity and conductivity ratios is simulated and compared to the case with no electric field. The results show that when the conditions for Quincke rotation are satisfied, the disc's migration toward the channel's center is hindered, in agreement with experiments [11].

2 GOVERNING EQUATIONS

Equations governing an incompressible flow subject to an external electric field may be written in dimensionless form as

$$\boldsymbol{\nabla} \cdot \mathbf{u} = 0, \tag{1}$$

$$\frac{\mathrm{D}\mathbf{u}}{\mathrm{D}t} = -\frac{1}{\rho}\boldsymbol{\nabla} p + \frac{1}{\mathrm{Re}}\boldsymbol{\nabla} \cdot \boldsymbol{\tau} + \frac{1}{\mathrm{Ei}}\mathbf{f}_{(e)}, \tag{2}$$

where $\mathbf{u}$ is the velocity vector, p is pressure, ρ is density, t is time and $\mathrm{D}/\mathrm{D}t = \partial/\partial t + \mathbf{u}\cdot\boldsymbol{\nabla}$ represents the material time derivative. Here, $\boldsymbol{\tau}$ is the viscous stress tensor,

$$\boldsymbol{\tau} = \mu\left[\boldsymbol{\nabla}\mathbf{u} + (\boldsymbol{\nabla}\mathbf{u})^{\dagger}\right], \tag{3}$$

where μ denotes viscosity and superscript $\square^{\dagger}$ represents the transpose operation. $\mathbf{f}_{(e)}$ is the electric force vector defined as [10]

$$\mathbf{f}_{(e)} = -\frac{1}{2}\mathbf{E}\cdot\mathbf{E}\boldsymbol{\nabla}\varepsilon + q^{v}\mathbf{E}. \tag{4}$$

Here ε denotes electric permittivity, q^{v} is is the volume charge density near the interface while $\mathbf{E}$ is the electric field vector. Assuming small dynamic currents and neglecting magnetic induction effects, the electric field is irrotational [12] and may be represented by gradient of an electric potential ϕ, $\mathbf{E} = -\boldsymbol{\nabla}\phi$. Further assumption of fast electric relaxation time compared to viscous relaxation time leads to the following relations for electric potential and charge density

$$\boldsymbol{\nabla}\cdot(\sigma\boldsymbol{\nabla}\phi) = 0, \tag{5}$$

$$q^{v} = \boldsymbol{\nabla}\cdot(\varepsilon\boldsymbol{\nabla}\phi), \tag{6}$$

where σ is the electrical conductivity.

Dimensionless values are formed using the following scales

$$\begin{aligned} &\mathbf{x} = \mathbf{x}^*/a, \quad \rho = \rho^*/\rho_f, \quad \mathbf{u} = \mathbf{u}^*/Ga, \quad t = t^*G, \\ &\mathbf{E} = \mathbf{E}^*/E_\infty, \quad \phi = \phi^*/E_\infty H, \quad p = p^*/\rho_f \left(Ga\right)^2, \\ &\mathcal{R} = \rho_s/\rho_f, \quad \mathcal{M} = \mu_s/\mu_f, \quad \mathcal{P} = \varepsilon_s/\varepsilon_f, \quad \mathcal{C} = \sigma_s/\sigma_f, \quad \mathcal{D} = a/H \end{aligned} \tag{7}$$

leading to Reynolds and inverse Electroinertial numbers defined as

$$\mathrm{Re} = \frac{\rho_f G a^2}{\mu_f}, \quad \mathrm{Ei} = \frac{\rho_f \left(Ga\right)^2}{\varepsilon_f E_\infty^2}. \tag{8}$$

Here E_∞ is the undisturbed electric field intensity, H is the distance between electrodes, G is the shear rate and a denotes disc radius (figure 1-a). An asterisk marks dimensional variables whereas subscripts $\square_s$ and $\square_f$ denote fluid and solid phases, respectively.

To distinguish between different phases, a color function $\hat{c}$ is defined such that it assumes a value of zero for one phase and unity for the other. The color function is then smoothed out across the phase boundaries as

$$c_\mathrm{i} = \sum_{\mathrm{j}=1}^{J_n} \frac{\hat{c}_\mathrm{j} W_\mathrm{ij}}{\psi_\mathrm{i}}, \tag{9}$$

to ensure smooth transition between the properties of each phase when used for their interpolation. Here, $\psi_\mathrm{i} = \sum_{\mathrm{j}=1}^{J_n} W_\mathrm{ij}$, is the number density of SPH particle i, calculated as the sum of interpolation kernel of neighboring particles i and j over all neighbors of particle i, J_n. Interpolation kernel, $W(r_\mathrm{ij}, h)$, is a function of the magnitude of distance vector, $\mathbf{r}_\mathrm{ij} = \mathbf{r}_\mathrm{i} - \mathbf{r}_\mathrm{j}$, between particle of interest i and its neighboring particles j and h, the smoothing length [13, 14]. Interpolation of phase properties is carried out using Weighted Harmonic Mean (WHM),

$$\frac{1}{\chi_\mathrm{i}} = \frac{c_\mathrm{i}}{\chi_s} + \frac{1 - c_\mathrm{i}}{\chi_f}, \tag{10}$$

where χ may denote density, viscosity, permittivity or conductivity [15].

All phases are treated as liquids initially while additional rigidity constraints are imposed in solid region [9]. To this end we use the current velocity of the solid particles to compute a center-of-mass velocity and an angular velocity for the solid object:

$$\mathbf{u}_s^t = \frac{1}{J_s} \sum_{\mathrm{j}=1}^{J_s} \mathbf{u}_\mathrm{j}, \quad \mathbf{u}_s^r = \frac{1}{I_s} \sum_{\mathrm{j}=1}^{J_s} \mathbf{u}_\mathrm{j} \times \mathbf{r}_{\mathrm{j}s}, \tag{11}$$

and then assign an individual velocity to each solid particle according to rigid body motion:

$$\mathbf{u}_\mathrm{i} = \mathbf{u}_s^t + \mathbf{u}_s^r \times \mathbf{r}_{\mathrm{i}s}. \tag{12}$$

Here, $\mathbf{r}_{is} = \mathbf{r}_i - \mathbf{r}_s$ where $\mathbf{r}_s$ denotes the solid object's center of mass, J_s is the number of particles present in the solid phase and I_s is the solid object's moment of inertia about its center of mass.

A predictor-correcter scheme is employed to advance the governing equations in time using a first-order Euler approach with variable timestep according to Courant-Friedrichs-Lewy condition, $\Delta t = C_{CFL} h/u_{max}$, where u_{max} is the largest particle velocity magnitude and C_{CFL} is taken to be equal to 0.25. In the predictor step, equations (5) and (6) are solved to obtain $\mathbf{f}_{(e)}^{(n)}$ through equation (4). Then position, velocity and number density are advanced to their intermediate forms using the following relations,

$$\mathbf{r}_i^+ = \mathbf{r}_i^{(n)} + \mathbf{u}_i^{(n)} \Delta t + \boldsymbol{\delta}\mathbf{r}_i^{(n)}, \tag{13}$$

$$\mathbf{u}_i^+ = \mathbf{u}_i^{(n)} + \left(\frac{1}{\mathrm{Re}} \boldsymbol{\nabla} \cdot \boldsymbol{\tau}_i^{(n)} + \frac{1}{\mathrm{Ei}} \mathbf{f}_{(e),i}^{(n)} \right) \Delta t, \tag{14}$$

$$\psi_i^+ = \psi_i^{(n)} - \Delta t \psi_i^{(n)} \left(\boldsymbol{\nabla} \cdot \mathbf{u}_i^+ \right), \tag{15}$$

where starred variables represent intermediate values and superscript (n) denotes values at the n^{th} time step. The artificial particle displacement vector in equation (13), $\boldsymbol{\delta}\mathbf{r}_i$, is defined as in [16] and a constant value of 0.06 is used.

Using intermediate values, pressure at the next time step is found by solving the Poisson equation. This is then followed by corrections in position and velocity of the particles, completing the temporal transition:

$$\boldsymbol{\nabla} \cdot \left(\frac{1}{\rho_i^+} \boldsymbol{\nabla} p_i^{(n+1)} \right) = \frac{\boldsymbol{\nabla} \cdot \mathbf{u}_i^+}{\Delta t}, \tag{16}$$

$$\mathbf{u}_i^{(n+1)} = \mathbf{u}_i^+ - \frac{1}{\rho_i} \boldsymbol{\nabla} p_i^{(n+1)} \Delta t, \tag{17}$$

$$\mathbf{r}_i^{(n+1)} = \mathbf{r}_i^{(n)} + \frac{1}{2} \left(\mathbf{u}_i^{(n)} + \mathbf{u}_i^{(n+1)} \right) \Delta t + \boldsymbol{\delta}\mathbf{r}_i^{(n)}. \tag{18}$$

In these equations, the rigidity constraints (equations (11) and (12)) are implemented after each velocity update.

Boundary conditions are enforced through the MBT method described in [17] while the first derivative and the Laplace operator are approximated through the following expressions

$$\frac{\partial f_i^m}{\partial x_i^k} a_i^{kl} = \sum_j \frac{1}{\psi_j} \left(f_j^m - f_i^m \right) \frac{\partial W_{ij}}{\partial x_i^l}, \tag{19}$$

$$\frac{\partial^2 f_i^m}{\partial x_i^k \partial x_i^k} a_i^{ml} = 8 \sum_j \frac{1}{\psi_j} \left(f_i^m - f_j^m \right) \frac{r_{ij}^m}{r_{ij}^2} \frac{\partial W_{ij}}{\partial x_i^l}. \tag{20}$$

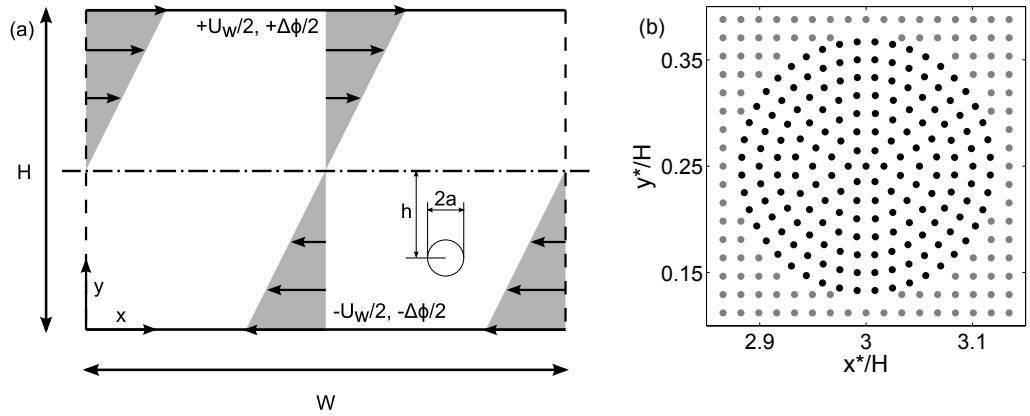


Figure 1: (a) Schematic of the test case. (b) Closeup view of initial particle distribution at the vicinity of the solid disc. Black points denote solid particles whereas gray points are fluid particles.

Here, $a_{\mathrm{i}}^{kl} = \sum_{\mathrm{j}} \frac{r_{\mathrm{ij}}^k}{\psi_{\mathrm{j}}} \frac{\partial W_{\mathrm{ij}}}{\partial x_{\mathrm{i}}^l}$ is a corrective second rank tensor that eliminates particle inconsistencies [16]. The left hand side of equation (16) is discretized as

$$\frac{\partial^2 f_{\mathrm{i}}^m}{\partial x_{\mathrm{i}}^k \partial x_{\mathrm{i}}^k} \left(2 + a_{\mathrm{i}}^{kk}\right) = 8 \sum_{\mathrm{j}} \frac{1}{\psi_{\mathrm{j}}} \left(f_{\mathrm{i}}^m - f_{\mathrm{j}}^m\right) \frac{r_{\mathrm{ij}}^k}{r_{\mathrm{ij}}^2} \frac{\partial W_{\mathrm{ij}}}{\partial x_{\mathrm{i}}^k}. \tag{21}$$

3 RESULTS

In this study, migration of a neutrally buoyant rigid disc in plane Couette flow is simulated. When released from a distance h from the middle of the channel, it is expected that the disc will migrate toward the centerline while moving with the flow for $\mathrm{Re} \leq 2$ [1, 18]. A schematic of this case is provided in figure 1-a. Computational domain consists of an 8×32 rectangle discretized by 13806 particles initially arranged in a Cartesian grid for fluid and concentric circles for the solid [9]. A close-up view of the particle arrangement at the vicinity of the solid disc is provided in figure 1-b. The confinement ratio is $\mathcal{D} = 0.125$ while the disc is placed at $h/H = 0.25$ below the centerline. Top and bottom walls abide by no-slip condition and are moving in opposite directions at a velocity of $U_w/2$ where $U_w = GH$. Periodic boundary condition is imposed in streamwise direction. Using a viscosity ratio of 100 [9] the results are compared to literature data in the absence of electric field at $\mathrm{Re} = 0.625$. Figure 2 shows the vertical position of the disc versus time. As it is seen, the results of the current simulation are in good agreement with the results found in [1–3].

When exposed to an external electric field, the motion of the disc may be altered depending on the properties of the fluid and solid. To investigate the effects, Ei is set to Re/4 to provide a large enough electric force while Re is kept equal to 0.625. A constant electric potential difference of $\Delta\phi = E_\infty H$ is imposed between the moving walls, resulting in an electric field perpendicular to the flow, while periodic boundary condition is imposed

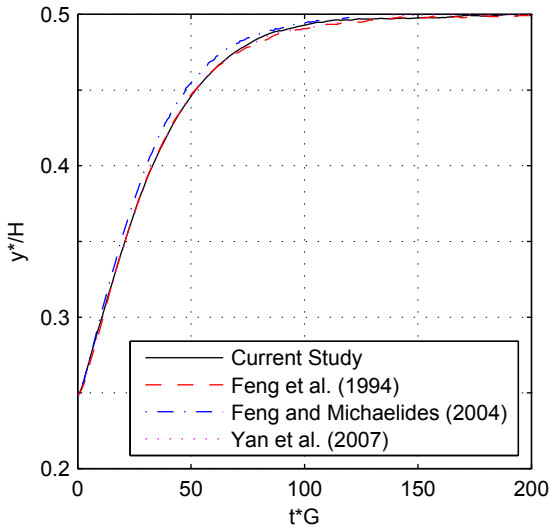


Figure 2: Comparison of the vertical position of the disc's center of mass versus time in the absence of electric field.

in streamwise direction. Permittivity and conductivity ratios are individually varied as 0.2, 0.5, 2 and 5 (table 1). When $\mathcal{P} = \mathcal{C}$, the leaky dielectric model is reduced to the perfect dielectric model. As such, equal permittivity and conductivity ratios are avoided here. Figure 3 plots the trajectory and velocity of the disc for the case without electric field and for $(\mathcal{P}, \mathcal{C})$ pairs of $(0.2, 5)$ and $(5, 0.2)$. The disc in case $(0.2, 5)$ is accelerated toward the centerline while in case $(5, 0.2)$ it is moved toward the wall. Case $(5, 0.2)$ is solved until the disc touches the bottom wall. Defining $\tau = \varepsilon/\sigma$ as the electric time scale, the necessary condition for Quincke rotation is $\tau_f < \tau_s$ [6]. The configuration of electric forces in case $(5, 0.2)$ is such that the condition for Quincke rotation is satisfied. This results in an angular moment that assists the rotation of the disc in the flow. On the other hand, case $(0.2, 5)$ has reduced angular velocity compared to the case with no electric field. Normal velocity becomes zero when disc hits bottom wall or reaches the channel centerline.

Table 1 provides the time until the disc comes to one radius distance of the bottom wall or the channel center. Bold numbers show that the disc approaches the bottom wall. The disc reaches the channel center at $t^*G = 25.94$ when no electric field is applied. When $\tau_f > \tau_s$ (lower triangle), the disc travels faster toward the channel center while when $\tau_f < \tau_s$ (upper triangle) the disc travels slower toward the channel center or moves toward the channel wall. Such observations are reported in experiments as well [11]. As it is seen, increasing the permittivity ratio at a constant conductivity ratio results in a gradual increase in the time needed for the disc to reach the channel center. At large enough permittivity ratios the trajectory is reversed and the disc migrates toward the channel wall. At this point, further increase of the permittivity ratio results in a

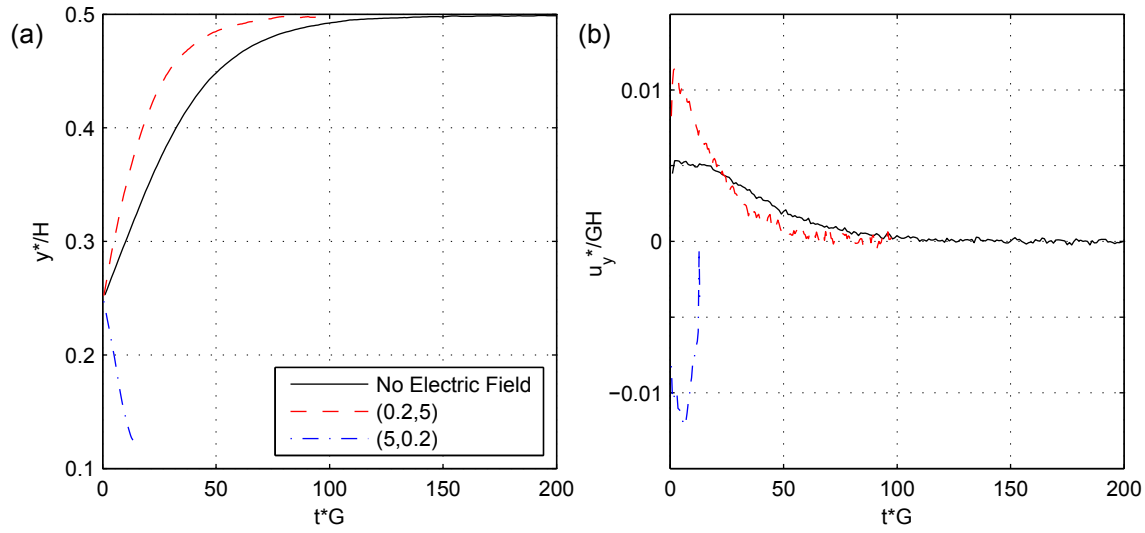


Figure 3: Comparison of the vertical position (a) and vertical velocity (b) of the disc's center of mass versus time.

Table 1: Time until the disc reaches one radius distance of channel center (regular) or bottom wall (bold). The disc reaches channel center at $t^*G = 25.94$ when no electric field is applied.

$\mathcal{P}$		0.2	0.5	2	5
$\mathcal{C}$	0.2	-	80.05	**19.18**	**13.34**
	0.5	17.79	-	**44.89**	**29.37**
	2	13.49	17.18	-	33.59
	5	13.81	16.89	24.07	-

faster migration toward the channel wall. The inverse of this trend is observed when the permittivity ratio is kept constant and the conductivity ratio is increased.

4 CONCLUSION

In this paper, we use ISPH to simulate the migration of a rigid disc in Couette flow under external electric field. The results of the case without electric field are compared to literature data and good agreement is observed. Further simulations are carried out at different permittivity and conductivity ratios and the results are compared to the case with no electric field. When the electric field augments the rotation of the disc, its motion toward the center of the channel is hindered. At large enough ratios, this results in complete reversal of the motion as the disc approaches the channel wall. Conversely the disc moves faster toward the center when its angular motion is hindered.

5 ACKNOWLEDGMENTS

The authors gratefully acknowledge financial support provided by the Scientific and Technological Research Council of Turkey (TUBITAK) for project number 112M721 and by the Natural Sciences and Engineering Research Council (NSERC) of Canada.

REFERENCES

[1] Feng, J., Hu, H.H. and Joseph, D.D. Direct simulation of initial-value problems for the motion of solid bodies in a Newtonian fluid .2. Couette and Poiseuille flows. *J. Fluid Mech.* (1994) **277**:271-301.

[2] Feng, Z.G. and Michaelides, E.E. The immersed boundary-lattice Boltzmann method for solving fluid-particles interaction problems. *J. Comput. Phys.* (2004) **195**:602-628.

[3] Yan, Y., Morris, J.F. and Koplik, J. Hydrodynamic interaction of two particles in confined linear shear flow at finite Reynolds number. *Phys. Fluids* (2007) **19**:113305.

[4] Quincke, G. Ueber rotationen im constanten electrischen felde. *Annalen der Physik* (1896) **295**:417-486.

[5] Peters, F., Lobry, L., Khayari, A. and Lemaire, E. Size effect in Quincke rotation: A numerical study. *J. Chem. Phys.* (2009) **130**:194905.

[6] Das, D. and Saintillan, D. Electrohydrodynamic interaction of spherical particles under Quincke rotation. *Phys. Rev. E* (2013) **87**:043014.

[7] Pannacci, N., Lemaire, E. and Lobry, L. Rheology and structure of a suspension of particles subjected to Quincke rotation. *Rheol. Acta* (2007) **46**:899-904.

[8] Huang, H.F., Zahn, M. and Lemaire, E. Continuum modeling of micro-particle electrorotation in Couette and Poiseuille flows-The zero spin viscosity limit. *J. Electrost.* (2010) **68**:345-359.

[9] Tofighi, N., Ozbulut, M., Rahmet, A., Yildiz, M. and Feng, J.J. Descent of a solid disk in quiescent fluid simulated using incompressible smoothed hydrodynamics. *Proceedings of WCCM XI* (2014) **5**:5310-5318.

[10] Saville, D. Electrohydrodynamics: The Taylor-Melcher leaky dielectric model. *Annu. Rev. Fluid Mech.* (1997) **29**:27-64.

[11] Boissy, C., Atten, P. and Foulc, J. On a negative rheological effect. *J. Electrost.* (1995) **35**:13-20.

[12] Hua, J., Lim, L. and Wang, C. Numerical simulation of deformation/motion of a drop suspended in viscous liquids under influence of steady electric fields. *Phys. Fluids* (2008) **20**:113302.

[13] Monaghan, J.J. and Lattanzio, J.C. A refined particle method for astrophysical problems. *Astron. Astrophys.* (1985) **149**:135–143.

[14] Monaghan, J.J. and Kocharyan, A. SPH simulation of multiphase flow. *Comput. Phys. Commun.* (1995) **87**:225–235.

[15] Tomar, G., Gerlach, D., Biswas, G., Alleborn, N., Sharma, A., Durst, F., Welch, S.W.J. and Delgado, A. Two-phase electrohydrodynamic simulations using a volume-of-fluid approach. *J. Comput. Phys.* (2007) **227**:1267-1285.

[16] Zainali, A., Tofighi, N., Shadloo, M.S. and Yildiz, M. Numerical investigation of Newtonian and non-Newtonian multiphase flows using ISPH method. *Comput. Meth. Appl. Mech. Eng.* (2013) **254**:99–113.

[17] Yildiz, M., Rook, R.A. and Suleman, A. SPH with the multiple boundary tangent method. *Int. J. Numer. Methods Eng.* (2009) **77**:1416–1438.

[18] Pan, T.W., Huang, S.L., Chen, S.D., Chu, C.C. and Chang, C.C. A numerical study of the motion of a neutrally buoyant cylinder in two dimensional shear flow. *Comput. Fluids* (2013) **87**:57-66.

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis (Eds)

MODELING AND NUMERICAL CALCULATION OF PISTON-LIKE OIL DISPLACEMENT FOR DOUBLY-PERIODIC SYSTEMS OF OIL FIELDS DEVELOPMENT

VLADIMIR I. ASTAFEV [*] **AND ANDREY E. KASATKIN** [†]

[*] Department of Oil and Gas Fields Development
Samara State Technical University
443100 Samara, Russia
e-mail: vladimir.astafev@mail.ru, web page: http://www.samgtu.ru

[†] Continuum Mechanics Department
Samara State University
443011 Samara, Russia
Email: darantion_yar@mail.ru, web page: http://www.samsu.ru

Key words: Mathematical modeling, Piston-like oil displacement, Weierstrass elliptic functions, Singular integral equation.

Summary. Prediction of the motion of the oil-water contact boundary has great importance in the problems of design of oilfield development by waterflooding: knowledge of the nature of coupled motion of oil and water, displacing oil in the reservoir allows us to optimize the system of oil field development. The simplest model of coupled filtering of oil and water is the model of "multicolored" liquids, which assumes that oil and water have the same or similar physical properties (density and viscosity).

In this paper we consider a more complex "piston-like" model of oil-water displacement, which takes into account differences in viscosity and density of the two fluids. Oil reservoir assumed to be homogeneous and infinite, fixed thickness, with constant values of porosity and permeability coefficients. It is assumed that the reservoir is developed by a group of a finite number of production and injection wells recurrent in two directions (doubly-periodic cluster). Filtration of liquids is described by Darcy's law. It is assumed, that both fluids are weakly compressible and the pressure in the reservoir satisfies the quasi-stationary diffusion equation.

Piston-like displacement model leads to the discontinuity of the tangential component of the velocity vector at the boundary of oil-water contact. Use of the theory of elliptic functions in conjunction with the generalized Cauchy integrals reduces the problem of finding the current boundaries of oil-water contact to the system of singular integral equations for the tangential and normal components of the velocity vector and the Cauchy problem for the integration of the differential equations of motion of the boundary of oil-water contact.

An algorithm for the numerical solution of this problem is developed. The monitoring of oil-water boundary motion for different schemes of waterflooding (linear row, four-point, five-point, seven-point, nine-point, etc.) is carried out.

1 INTRODUCTION

Among the methods of oil fields development the waterflooding method [1, 2] became widespread. The main objective of waterflooding is to maintain by flooding the reservoir pressure, inevitably falling in the primary field development. Simulation of the flooding process, analysis of its qualitative and quantitative characteristics for different schemes of flooding are the purpose of the present study. In this paper, the model of the piston-like oil displacement by water [2], which takes into account the difference in physical properties (density and viscosity) displaced and displacing fluids.

The task of monitoring of the line flooding motion (the line separating the displacement of the water and oil) was first considered by Muskat [3] and subsequently aroused great interest among researchers. It was noted by Leibenson [4], where the viscosity of the displacing fluid was neglected. In the paper of Danilov and Kats [5], based on the potential theory, the original problem of monitoring of the line flooding motion has been reduced to a nonlinear integro-differential equations. Danilov's method was used by Fazlyev [6] for the some scheme of the areal flooding.

2 MATHEMATICAL MODEL

Consider the plane filtration flow of a viscous compressible fluid with viscosity μ and compressibility β in an infinite horizontal reservoir with permeability k, porosity m and thickness h. For the quasi-stationary state of filtration flow the pressure in the reservoir $p(x,y,t)$ satisfies the diffusivity equation and the velocity components $V_x(x,y,t)$ and $V_y(x,y,t)$ are calculated by the Darcy law [1-3]

$$\begin{aligned} &\frac{\partial p}{\partial t} = \chi\left(\frac{\partial^2 p}{\partial x^2} + \frac{\partial^2 p}{\partial y^2}\right), \\ &V_x = -\frac{k}{\mu}\frac{\partial p}{\partial x}, V_y = -\frac{k}{\mu}\frac{\partial p}{\partial y} \end{aligned} \tag{1}$$

where $\chi = k/m\mu\beta$ - the coefficient of diffusivity.

Simulated reservoir is developed by a doubly periodic system of production and injection wells. The whole set of production and injection wells can be represented as an infinite number of repetitions of the wells in two directions. Such repetition can be described by a doubly periodic lattice, which in the complex plane $z=x+iy$ is defined by two complex periods ω_1 and ω_2. The whole set of lattice points in the complex plane is defined as $\omega = m\omega_1 + n\omega_2$ ($m,n = 0, \pm 1, \pm 2, \ldots$), the value of $\Delta = \mathrm{Im}(\bar{\omega}_1\omega_2)$ corresponds to the area of a parallelogram lattice (Figure 1).

Solution of the equation (1) in the case of doubly periodic system of production and injection wells (doubly periodic claster) has been obtained in [7-9]. The distribution of the velocity field was presented by the Weierstrass zeta function and written as follows:

$$\begin{aligned} &\bar{V}(z,\bar{z}) = V_x(x,y) - iV_y(x,y) = \\ &= -\sum_{k=1}^{n} \frac{Q_k}{2\pi h}\left(\zeta(z-z_k) + \alpha(z-z_k) - \beta(\bar{z}-\bar{z}_k)\right), \end{aligned} \tag{2}$$

where $\zeta(z)=\frac{1}{z}+\sum_{i,j=-\infty}^{\infty}{}'(\frac{1}{z-\omega}+\frac{1}{\omega}+\frac{z}{\omega^2})$ - Weierstrass zeta function, z_k - location of the k-th well in the cluster, Q_k - flow rate of the k-th well ($Q_k<0$ for the injection well and $Q_k>0$ for the production well), $\beta=\pi/\Delta$ and $\alpha=(\beta\bar{\omega}_1-2\zeta(\omega_1/2))/\omega_1$.

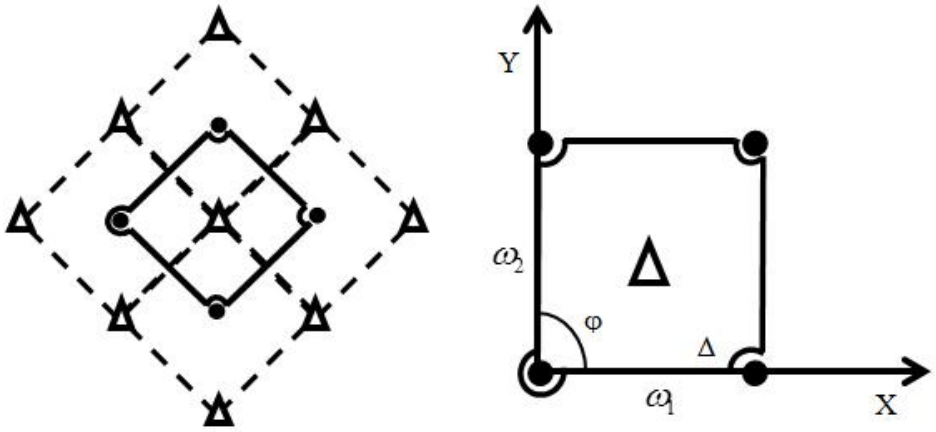


Figure 1: A five-point scheme of flooding (production wells are marked by black circles, injection – by white triangles) and doubly periodic lattice with its basic elements

In the case of the piston-like flooding (moving boundary problem in the theory of filtration) the boundary conditions on the flooding line L (figure 2) for the tangential V_t and normal V_n components of the filtration velocity and for the pressure p are the follows (the index o refers to particles of oil, and the index w - to the particles of water):

$$\begin{aligned} &V_t^{(o)}\mu_{(o)}=V_t^{(w)}\mu_{(w)};\\ &V_n^{(o)}=V_n^{(w)};\\ &p^{(o)}=p^{(w)}. \end{aligned} \tag{3}$$

Discontinuous on the line L, but doubly periodic function $\bar{V}(z,\bar{z})=V_x(x,y)-iV_y(x,y)$ in the complex plane $z=x+iy$ will be sought in the form of the Cauchy-type integral [7-9]

$$\bar{V}(z,\bar{z})=F(z,\bar{z})+\frac{1}{2\pi i}\oint_L\zeta(\tau-z)\gamma(\tau)d\tau, \tag{4}$$

where $F(z,\bar{z})$ is given by equation (2).

The left and right boundary values of this function (4) at a point $z(s)$ on the line L can be written in the form of Sokhotski–Plemel formulas [10]

$$\begin{cases} \bar{V}^{(w)}(z(s))=F(z(s))+\frac{1}{2\pi i}\oint_L\zeta(z(\sigma)-z(s))\gamma(z(\sigma))\frac{dz}{d\sigma}d\sigma+\frac{\gamma(z(s))}{2};\\ \bar{V}^{(o)}(z(s))=F(z(s))+\frac{1}{2\pi i}\oint_L\zeta(z(\sigma)-z(s))\gamma(z(\sigma))\frac{dz}{d\sigma}d\sigma-\frac{\gamma(z(s))}{2}. \end{cases} \tag{5}$$

Applying the Sokhotskiy-Plemel formulas (5), the unknown function *γ(s)* and the complex velocity $\overline{V}(z(s))$ at a point *z(s)* of the line *L* can be expressed as

$$\begin{aligned} &\gamma(s) = \overline{V}^{(w)}(s) - \overline{V}^{(o)}(s), \\ &\overline{V}(s) = (\overline{V}^{(w)}(s) + \overline{V}^{(o)}(s))/2. \end{aligned} \tag{6}$$

Taking into account the boundary conditions on the moving boundary (3), the relations (6) can be rewritten as

$$\begin{aligned} &\gamma(s) = (1 - \mu_{(w)}/\mu_{(o)}) V_t^{(w)}(s) e^{-i\alpha}, \\ &\overline{V}(s) = ((1 + \mu_{(w)}/\mu_{(o)}) V_t^{(w)}(s)/2 + iV_n^{(w)}(s)) e^{-i\alpha}, \end{aligned} \tag{7}$$

where α – the angle between the tangent line *L* and the axis *X,* i.e. $e^{i\alpha} = dz/ds$ (figure 2).

Denoting the viscosity ratio as $\kappa = \mu_{(w)}/\mu_{(o)}$, the normal and tangential velocity components of the water as $T(s) = V_t^{(w)}(s)$ and $N(s) = V_n^{(w)}(s)$, the equations (7) allow us to obtain the following singular integral equation for the unknown functions *T(s)* and *N(s)* on the line *L:*

$$\frac{1+\kappa}{2} T(s) + iN(s) = [F(s) + \frac{1-\kappa}{2\pi i} \oint_L \zeta(z(\sigma) - z(s)) T(\sigma) d\sigma] \frac{dz}{ds}. \tag{8}$$

The integral equation (8) must be supplemented by a differential equation that determines the time evolution of the line *L*. This equation has the form [1-3]

$$\begin{aligned} &m \frac{\partial \overline{z}(s,t)}{\partial t} = \overline{V}(z(s,t)), \\ &z(s,0) = z_0 + r_w e^{i\theta}. \end{aligned} \tag{9}$$

where z_0 - the center of the injection well with radius r_w, through which water is pumped into the reservoir. The initial condition (9) indicates the starting position of the point $z(s,0) = z_0 + r_w e^{i\theta}$ in the beginning of flooding, the corresponding angle θ is determined on the contour of the injection well.

Using once more the second Sokhotskiy-Plemel formula (7) and the equation (8), the equation (9) can be rewritten as

$$\begin{aligned} &m \frac{\partial \overline{z(s,t)}}{\partial t} \frac{\partial z}{\partial s} = \frac{(1+\kappa)}{2} T(s,t) + iN(s,t), \\ &z(s,0) = z_0 + r_w e^{i\theta}. \end{aligned} \tag{10}$$

where functions *T(s)* and *N(s)* for a given time *t* are obtained by solving the singular integral equation (8).

3 NUMERICAL SOLUTION OF THE MOVING BOUNDARY PROBLEM

Consider the algorithm for the numerical solution of integro-differential equations (8) and (10). For the numerical solution of singular integral equation (8) we divide the contour *L* by discrete set of points on the elements $[z_i, z_{i+1}]$, *(i = 0, 1, ... N-1)* (Figure 2). Due to the closure

of line L, the first and the last points of the partition are the same, i.e., $z_N=z_0$. Each of the points z_k corresponds to the length of the arc s_k. Let us choose and fix the point $z_k=z(s_k)$ on the contour L. Separating in equation (6) the real and imaginary parts, we obtain the following equations for the unknown values of $T_k=T(s_k)$ and $N_k=N(s_k)$ at the points $z_k=z(s_k)$:

$$\begin{cases} \dfrac{1+\kappa}{2}T(s_k) = \mathrm{Re}\left\{[F(s_k) + \dfrac{1-\kappa}{2\pi i}\oint\limits_L \zeta(z(\sigma)-z_k)T(\sigma)d\sigma]z_k^{'}\right\}; \\ N(s_k) = \mathrm{Im}\left\{[F(s_k) + \dfrac{1-\kappa}{2\pi i}\oint\limits_L \zeta(z(\sigma)-z_k)T(\sigma)d\sigma]z_k^{'}\right\}. \end{cases} \tag{11}$$

Omitting the rather cumbersome intermediate calculations, we write the approximation of the integral term in equation (11) as follows:

$$\begin{aligned} &\oint\limits_L \zeta(z(\sigma)-z_k)T(\sigma)d\sigma = \\ &= \frac{1}{2z_k'}\left[2T_k \ln\left(\frac{\Delta s_k}{\Delta s_{k-1}}\right) + T_{k+1} - T_{k-1} + \sum_{i=k+1}^{N+k-1} \zeta(z_i - z_k)T_i(\Delta s_i + \Delta s_{i-1})z_k'\right]. \end{aligned} \tag{12}$$

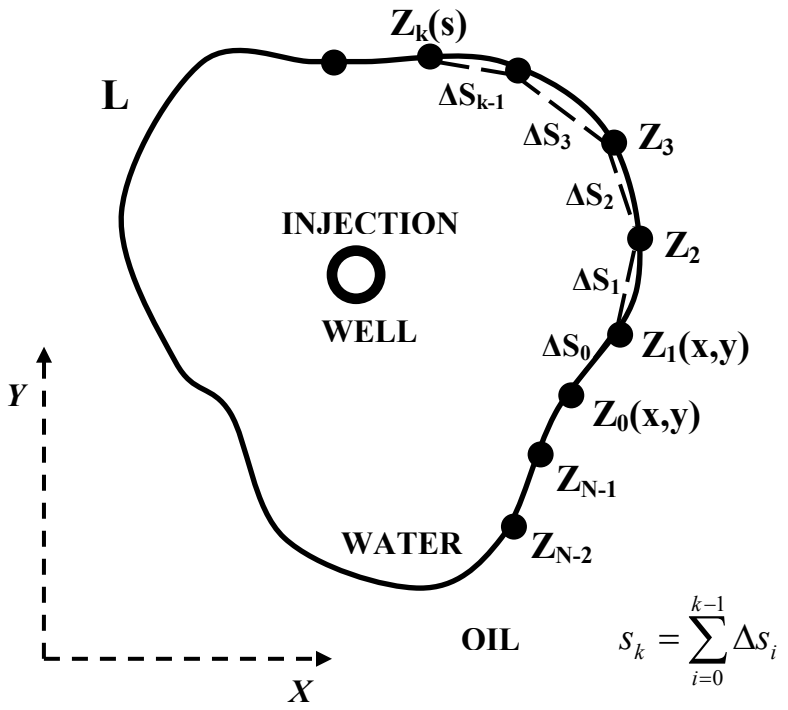


Figure 2: Parameterization scheme of line L and its discretization by points Z_k *(k =0…N-`1).*

Equations (11) show us that the real part of the singular integral equation (8) matches the finding of values of the unknown function $T_k=T(s_k)$ at a given time t_n on a discrete set of points $z_k=z(s_k)$. The imaginary part of this equation is, in fact, the formula for calculating the values of the function $N_k=N(s_k)$ at a given time t_n at the same discrete set of points.

The obtained values of T_k and N_k are then used to calculate the displacements of points $z_k=z(s_k)$ in the time interval $[t_n, t_{n+1}]$. These displacements are determined by the numerical solution of the Cauchy problem (10) with the Runge-Kutta method, modified in view of the complex nature of the differential equation (10). In the calculations was chosen the dimensionless time τ, associated with the original time t as $\tau = 10^{-4} Qt / 2\pi mh\omega_1^2$.

4 RESULTS OF CALCULATIONS

To solve this problem there was developed software system to track the evolution of the flooding front (line L) in time, as well as to quantify the effectiveness of a particular scheme of flooding (water breakthrough time to production well and waterflood sweep efficiency K_{wse}). In numerical calculations there was considered frontal row, four-point, five-point, seven-point and nine-point scheme of flooding. To determine the current position of the flooding front it was involved 180 tracers - points coming out at the initial time of the injection well.

It is known [11] that the viscosity ratio κ has a negative impact on the ultimate recovery: the decrease in the parameter κ leads to the decrease in the volume of recoverable oil due to the growing instability of oil displacement by water. In addition, when viscosity ratio κ<1, the Saffman-Taylor instability [12] occurs, which is often called as "viscous fingering". The viscous fingering effect was observed in the course of our numerical calculations. The figure 3 shows the formation of viscous fingers for five-point scheme of flooding at κ=1/5 and τ=0.045 values. For comparison, the figure 3 is supplemented by the figure from [13] (figure 6a in the cited paper) obtained experimentally for similar geometry placement of wells.

Additionally, the figures 4 and 5 show us the flooding area for seven-point and nine-point schemes of flooding. For comparison the left parts of these pictures show the flooding area with viscosity ratio $\kappa=1$ and the right parts – with $\kappa=1/4$.

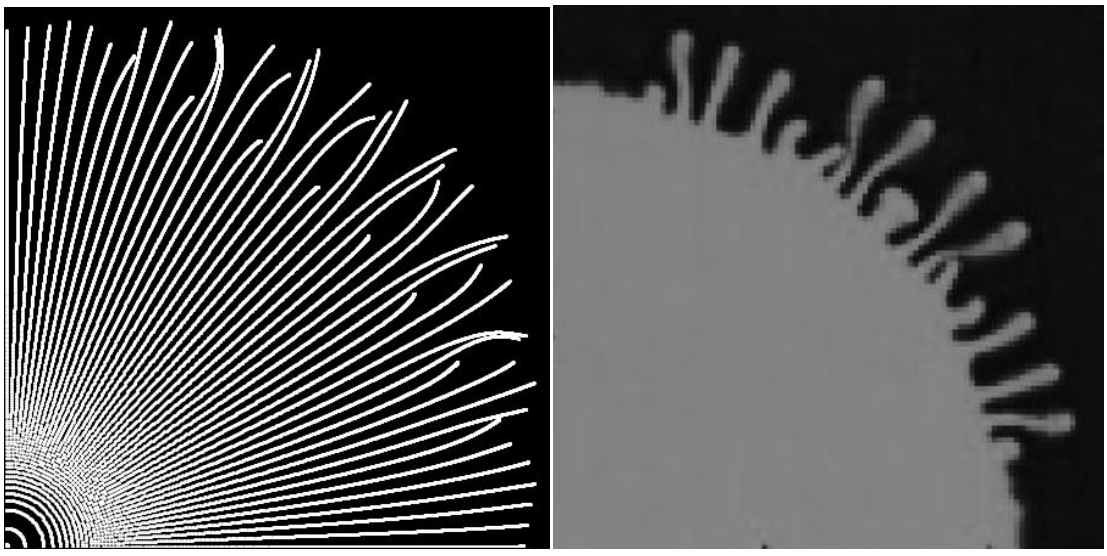
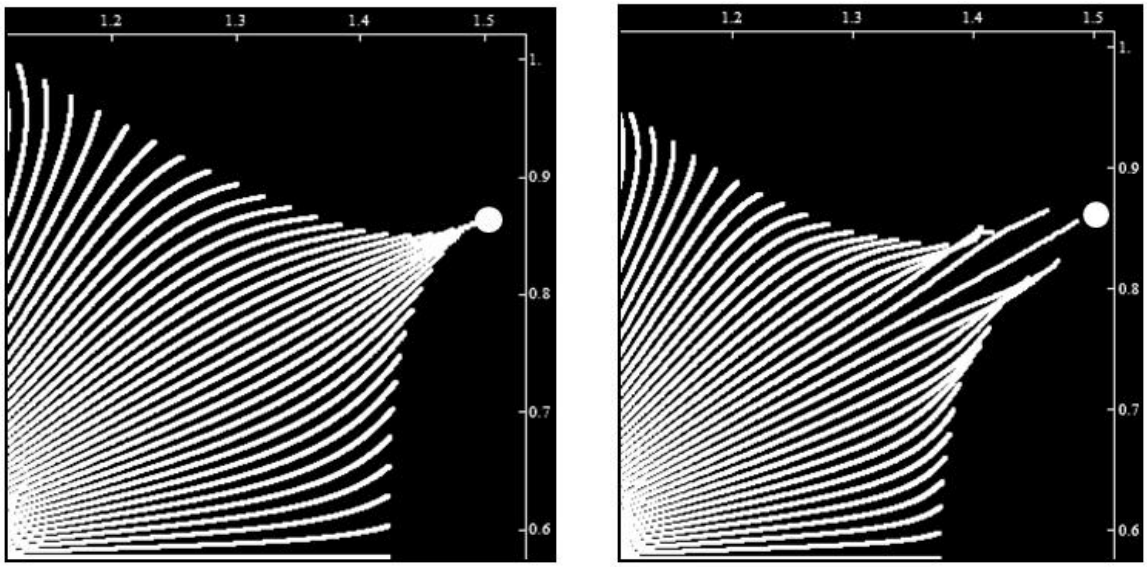
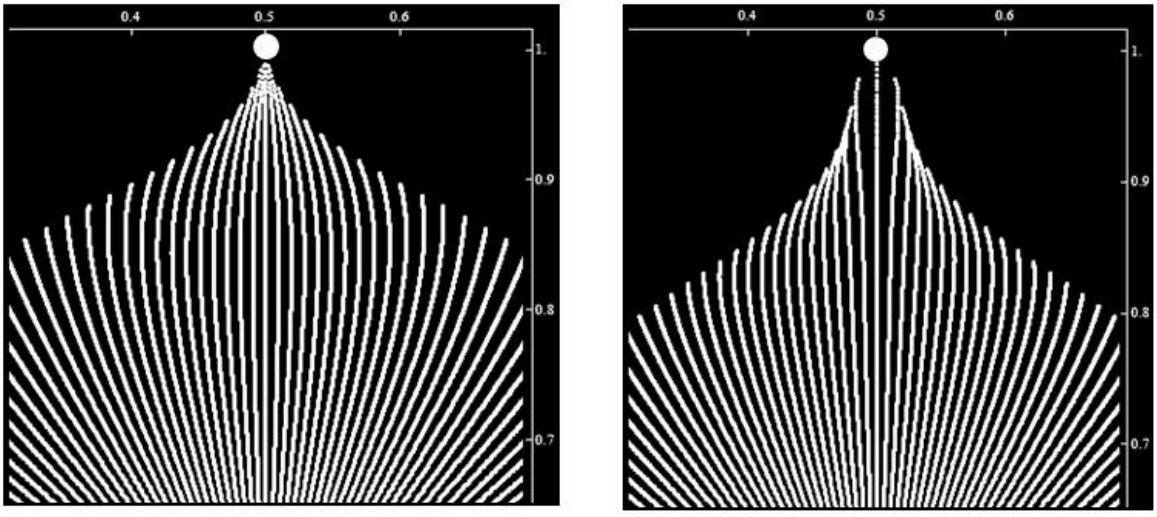


Figure 3: Saffman-Taylor instability of the front of flooding for the five-point scheme (on the left – the numerical calculation for the mobility κ=1/5 = 1/5 and τ = 0.0450, on the right - the data from [13]).



$\kappa = 1, \tau = 0.0516$ $\kappa = 1/4, \tau = 0.0439$

Figure 4: The flooding area for seven-point scheme of flooding.



$\kappa = 1, \tau = 0.0277$ $\kappa = 1/4, \tau = 0.0232$

Figure 5: The flooding area for nine-point scheme of flooding.

Further, the Table 1 shows the values of the water breakthrough time into the producing wells τ and the Table 2 - the waterflood sweep efficiency K_{wse}, calculated for four of waterflood patterns with different values of κ. Abbreviation VF (viscous fingers) indicates the appearance of the "viscous fingering" for the selected scheme of the flooding. Therefore, due to violation of the smoothness of the flooding front the count of the waterflood sweep efficiency K_{wse} for a given value of κ is not possible. For greater clarity, the columns of Table 2 are supplemented with values taken from the book by F. Craig [11] for the case of κ=1.

Table 1: The values of dimentionless time τ for different values of the mobility $\varkappa$

Scheme of flooding	Values of the dimensionless breakthrough time τ			
	κ=1	κ=1/2	κ=1/3	κ=1/4
Five-point	1152	1017	960	VF
Front row	1820	1500	VF	VF
Seven-point	516	472	452	439
Nine-point	277	252	240	232

Table 2: The values of waterflooding coverages K_{cov} for different values of the mobility $\varkappa$

Scheme of flooding	Values of the waterflood sweep efficiency K_{wse}				
	κ=1 [8]	κ=1	κ=1/2	κ=1/3	κ=1/4
Five-point	70%	72,5%	63,5%	60%	VF
Front row	58%	57,2%	47%	VF	VF
Seven-point	73%	75,2%	68,5%	65,3%	63,3%
Nine-point	55%	52,7%	48%	45,5%	44%

ACKNOWLEDGEMENTS

This work was performed as the basic part of the state problem of the Russian Ministry of Science and Education for Samara State Technical University (code project 1785). Authors are also thankful the Russian Foundation for Basic Research for financial support under grant 14-01-97041-r-Volga region-a.

REFERENCES

[1] Zheltov, Y.P. *Development of oil fields.* Moscow, Nedra, (1998, in Russian).
[2] Willhite, G. Paul. *Waterflooding.* SPE Textbook. Vol. 3. Society of Petroleum Engineers, Richardson, Texas, (1986).
[3] Muskat, M. *The flow of homogeneous fluids through porous media.* New York, McGraw-Hill, (1937).
[4] Leibenzon, L.S. *Oilfield mechanics.* Moscow, Nefteizdat, Part II, (1934, in Russian).
[5] Danilov, VL, Kats, P.M. *Hydrodynamic calculations mutual displacement of fluids in porous media*. Moscow, Nedra, (1980, in Russian).
[6] Fazlyev, R.T. *Waterflooding oilfields.* Moscow-Izhevsk, Institute of Computer Science, (2008, in Russian).
[7] Astafev, V.I., Kasatkin A.E., Roters P.V. Elliptic functions in modeling of oil recovery *ECMOR XIII – 13th European Conference on the Mathematics of Oil Recovery.* Biarritz, France, (2012). DOI: 10.3997/22144609.20143268.
[8] Astafev, V.I., Roters P.V. Analitical solution for a double-periodic multiwell reservoir systems. *Tyumen 2013 - New Geotech-nology for the Old Oil Provinces.* Tyumen, Russia, (2013). DOI: 10.3997/22144609.20142692.
[9] Astafev, V.I., Roters. P.V. Simulation and optimization of multi-well field development doubly periodic clusters. *Vestnik of SSU.* (2013) **110:**. 170-183 (in Russian).
[10] Gakhov, F.D. *Boundary value problems*. Dover Publications, New York, (1990)
[11] Craig, F.F. *The reservoir engineering aspecs of waterflooding*. SPE Monograph

Series. Vol. 3. Society of Petroleum Engineers, Dallas, (1971).
[12] Saffman, P.G. Viscous fingering in Hele-Shaw cells. *J. Fluid Mech.* (1986) **173:** 73-84.
[13] Dapira, P., Glimm, J., Lindquist, B. and McBryan, O. Polymer Floods: A Case Study of Nonlinear Wave Analysis and of Instability Control in Tertiary Oil Recovery. *SIAM Journal on Applied Mathematics* (1988) **48:** 353-373.

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis (Eds)

PARAMETER IDENTIFICATION OF CHABOCHE MATERIAL MODEL USING INDANTATION TEST DATA AND INVERSE APPROACH

MARTIN CERMAK[*], RADIM HALAMA[†], TOMAS KARASEK[*] AND JAROSLAV ROJICEK[†]

[*] IT4Innovations
VSB - Technical University of Ostrava
17. listopadu 15/2172, Ostrava, Czech Republic
e-mail: martin.cermak@vsb.cz, tomas.karasek@vsb.cz

[†] Faculty of Mechanical Engineering
VSB - Technical University of Ostrava
17. listopadu 15/2172, Ostrava, Czech Republic
e-mail: radim.halama@vsb.cz, jaroslav.rojicek@vsb.cz

Key words: Genetic Algorithm, Sensitivity Analysis, Inverse Approach, Finite Element Method, Indentation, Ratcheting

Abstract. In this paper genetic algorithm and sensitivity analysis are used to identify 6 parameters of Chaboche kinematic hardening model using repeated Finite element (FE) simulations of indentation test. Five of them are material constants of Chaboche kinematic hardening model itself. The last one represents the stiffness of the foundation and the indenter. To obtain experimental data indentation test under cyclic loading on universal tensile testing machine was performed. Because for sensitivity analysis to obtain all possible combinations of parameters and its values large number of simulation have to be performed supercomputer Anselm hosted by IT4Innovation has been used. Advantage of using supercomputer is that every simulation could use multiple cores which will reduce computational time. Moreover, since each simulation is independent, computational time could be further reduced by performing multiple simulations at the same time. It is clear from the comparison of both methods that the genetic algorithm is very good choice for the parameter estimation.

1 INTRODUCTION

Although experimental measurement bears certain level of uncertainty and is influenced by error it is still the only way how to obtain material properties of any material. To describe the stress-strain behavior of any material under cyclic loading huge number of experiments has to be performed. In practice, it is sometimes very difficult to prepare specimens of the investigated material or application of standard testing method is unrealistic. The later occurs for example in the case of thin samples. Therefore, we need a new method to obtain stress-

strain behaviour of investigated material. Numerical modelling and simulation could help to reduce number of physical experiments and also could give us better insight into the problematic

For sufficiently accurate FE modelling of cyclically loaded structures we need good representation of material behaviour. In the case of fatigue occurrence we should describe well cyclic plasticity phenomenon as Bauschinger effect, cyclic hardening/softening of material or the phenomenon called ratcheting (cyclic creep). Ratcheting can be described as the accumulation of any plastic strain component of strain tensor with increasing number of cycles. The ratcheting may occur in practice for instance in the rolling/sliding contact.

One of the first plasticity models, which can qualitatively capture ratcheting in numerical calculations, is Chaboche model [1]. Cyclic plasticity models have been extensively developed over the past three decades. The most popular kinematic hardening rules introduced into new constitutive theories are Ohno-Wang model II [2] and AbdelKarim-Ohno model [3]. For certain materials we can use Chaboche model with two backstress parts to capture all important effects in the simulation but for others it is necessary to implement a robust cyclic plasticity model into the FE code [4].

The main aim of this contribution is comparison of various approaches to cyclic plasticity model calibration from indentation tests performed on the wheel steel Class C. Two algorithms have been applied, the genetic algorithm and the sensitivity analysis to estimate 5 parameters of Chaboche model [1] and a stiffness of the indenter using experimental data from an indentation test with repeated loading. Results show very good prediction of the test using only two backstress parts in the Chaboche suporposition rule. The developed method can be advantageously used to compare various wheel steels from ratcheting point of view.

2 EXPERIMENTAL DATA

An indentation test was realized on testing machine TESTOMETRIC M500-50CT (FS_ZAZ_MR_11_009) at the VSB-Technical University of Ostrava to extract data for comparison of effectiveness of two different algorithms for material parameters identification. Our experiment was done by indeting of 5 mm steel ball into Class C wheel steel specimen. Applied force was between 5 and 2000 N increasing by force rate of 20 N/s. To obtain results for our calculations 10 cycles were performed and displacement as a function of applied force was recorded using a standard.

3 MODEL DESCRIPTION

Time-independent theory of elastoplasticity [5] was applied in this paper. Plastic behaviour is characterized by von Mises plasticity condition and could be described by the following equation

$$f = \sqrt{\frac{3}{2}(\mathbf{s}-\mathbf{a}):(\mathbf{s}-\mathbf{a})} - \sigma_{\mathrm{Y}} = 0, \tag{1}$$

where $\mathbf{s}$ is the deviatoric part of stress tensor $\boldsymbol{\sigma}$, a is deviatoric part of kinematic tensor $\boldsymbol{\alpha}$ and $\boldsymbol{\sigma_Y}$ is the yield stress.

To describe Bauschinger effect [6] pure kinematic hardening rule could be considered. Thus, no isotropic hardening was assumed for material model in this study.

Memory term introduced by Armstrong and Frederick in [7], added to Prager's bilinear kinematic rule could be written in following form:

$$\mathbf{d\alpha} = \frac{2}{3} C \mathbf{d\varepsilon_p} - \gamma \boldsymbol{\alpha} dp\,, \tag{2}$$

where C, γ are material parameters and dp is accumulated equivalent plastic strain increment. It is possible to describe only the ratcheting with steady state (constant ratcheting strain increment in every cycle) with Armstrong-Frederick model and correct stress - strain response characterization is difficult. To treat these disadvantages of Armstrong-Frederick model, Chaboche proposed a superposition rule in [8] for backstress

$$\boldsymbol{\alpha} = \sum_{i=1}^{M} \boldsymbol{\alpha}_i\,, \tag{3}$$

whereas evolution of each kinematic part is directed by Armstrong-Frederick rule

$$\mathbf{d\alpha}_i = \frac{2}{3} C_i \mathbf{d\varepsilon_p} - \gamma_i \boldsymbol{\alpha}_i dp\,. \tag{4}$$

Practically, from two to five kinematic parts are usually used. In this paper two kinematic parts are assumed, thus it is necessary to estimate material parameters C_1, γ_1, C_2, γ_2, and also the yield stress σ_Y.

4 METHODS DESCRIPTION

To perform parameter identification several methods such as random gradient, genetic algorithm, and sensitivity analysis could be used [9] [10]. In this paper genetics algorithm and sensitivity analysis are used to find material constants of Chaboche kinematic hardening model using repeated FE simulations of indentation test. Results from FE analysis are compared with experimental measurements.

Advantage of genetics algorithms is that for well-defined problem number of iteration needed to reach global minimum is usually very low. Detailed description of genetic algorithm could be found in [11]. In principle it is iterative process which is repeated until global minimum is found. In our case global minimum is defined as difference between measured and calculated values.

Sensitivity analysis due to high number of possible combinations could be seen as an inadequate approach. Number of all possible combinations for *N* number of parameters, and P values for each parameter could be calculated as

$$V_N^P = N^P. \tag{15}$$

In our case, we have 6 parameters and 5 values for each parameter, therefore number of all possible combinations is

$$V_5^6 = 5^6 = 15\,625. \tag{16}$$

The disadvantage of sensitivity analysis i.e. high number of simulations could be balanced by better understanding of how final results depend on input parameters and their combinations.
To reduce computational time needed for performing of high number of simulations supercomputers could be employed. In our case we used supercomputer Anselm hosted by IT4Innovation to run simultaneously several FE simulations.

5 NUMERICAL EXPERIMENTS

Numerical model reproducing experimental set-up was created using commercial FE package ANSYS. Numerical model of specimen was created using approximately 2000 axisymmetric structural elements (PLANE 182). The element has plasticity, hyperelasticity, stress stiffening, large deflection, and large strain capabilities. In our numerical simulation indented ball was considered as absolutely rigid. Numerical model of specimen and indented ball is shown at Figure 1.

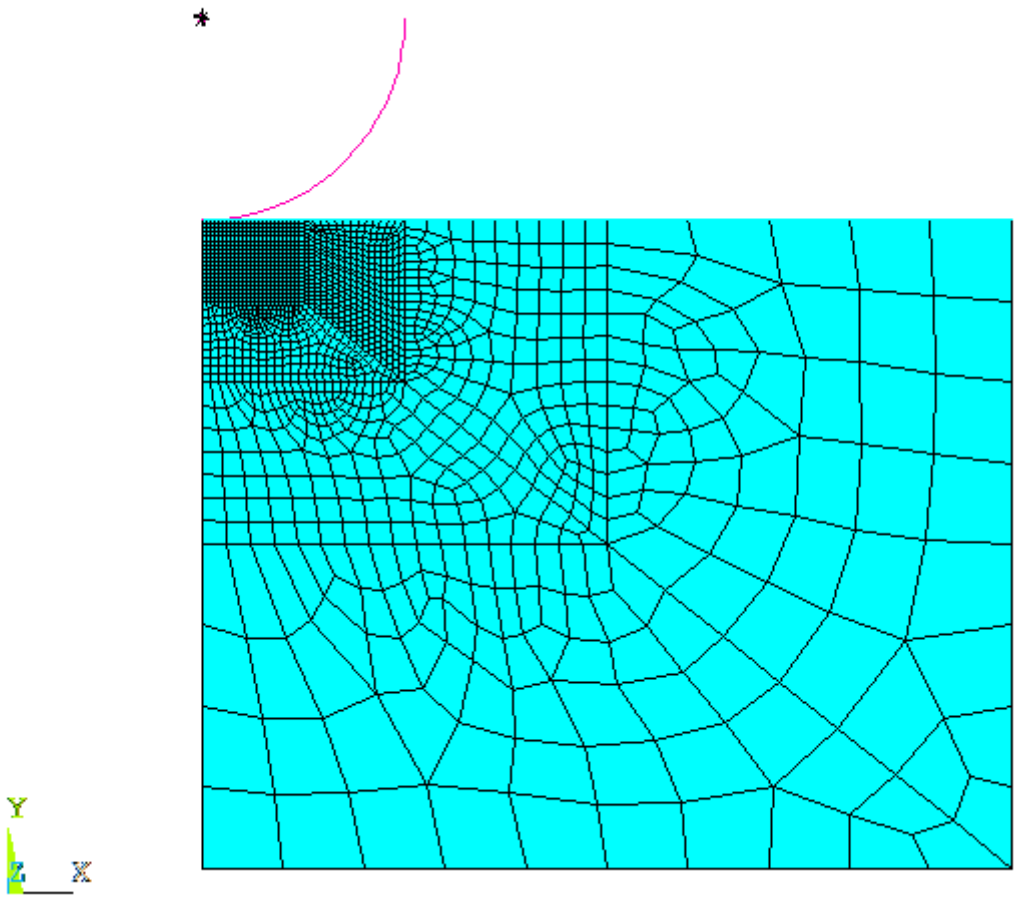


Figure 1: FE model of specimen and indented ball.

As boundary conditions axisymmetric boundary condition (displacement in x direction is set to zero) is used for left boundary and displacement in y direction was set to zero for nodes at the bottom of the specimen. Nodal forces in y direction were applied according to experiment.

The values of searched parameters for genetic algorithm are listed in Table 1, where the first iteration (second row of table) is the initial state and in the next two rows are values of parameters in 600 and 1452 steps. The convergence of genetic algorithm is depicted in Figure 2.

Tabulka 1: Values of parameters for genetic algorithmm

Iteration	σ_Y	C_1	γ_1	C_2	γ_2	rigidity	error
1	450.0	120 000	400	10 000	5.0	120 000	0.17400
600	337.0	113 601	552	7 496	5.6	104 998	0.00847
1452	354.9	107 336	582	7 146	5.5	103 920	0.00750

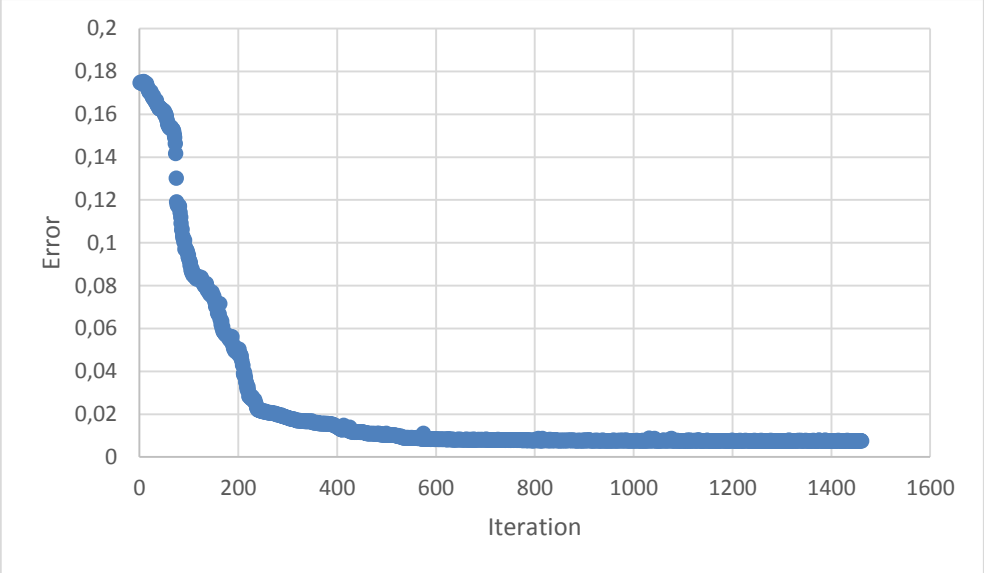


Figure 2: Convergence of genetic algorithm.

Material model used for numerical experiment is described in previous section. For sensitivity analysis values of parameters C1, C2, C3, C4, C5, and C6 are listed in Table 2.

Table 1: Values of parameters for sensitivity analysis

C1	80 000	90 000	100 000	110 000	120 000
C2	350	375	400	425	450
C3	60 000	75 000	90 000	105 000	120 000
C4	200	250	300	350	400
C5	10 000	12 500	15 000	17 500	20 000
C6	1	2	3	4	5

Since number of iteration of sensitivity analysis is apriori known we will be interested whether this method will be able to give us results which will satisfy our accuracy criteria. We decided that satisfactory results will be results for which difference between measured and calculated results is less than 1%. In case of genetics algorithm where we will run as many iterations as needed to fulfil accuracy criteria we will be interested in number of iterations as well.

For genetics algorithm 991 number of iterations in total were needed to obtain combination of parameters which leads to the solution satisfying our criteria described above. Fig. 3 shows comparison between numerical solution and experimental results. Maximal error is 0.75% and optimal values for input parameters are listed in Table 3.

Sensitivity analysis needed 15 625 iterations, as explained in previous chapter, to test all possible combinations of the parameters. The error for best combinations of the parameters is 1.0% and optimal value of input parameters are listed in Table 3. Fig. 4 shows comparison between experimental results and results obtained from numerical simulation. On Figure 5 comparison between both numerical approaches i.e. genetic algorithm and sensitivity analysis could be seen and detail is depicted in Figure 6.

Table 2: Comparison of the best value for genetic algorithm and sensitivity analysis

Parameter	Gen. Alg.	Sens. Anal.
C1	103919.9	110000
C2	354.9	425
C3	107336.7	10500
C4	582.1	400
C5	7146.3	12500
C6	5.53	5

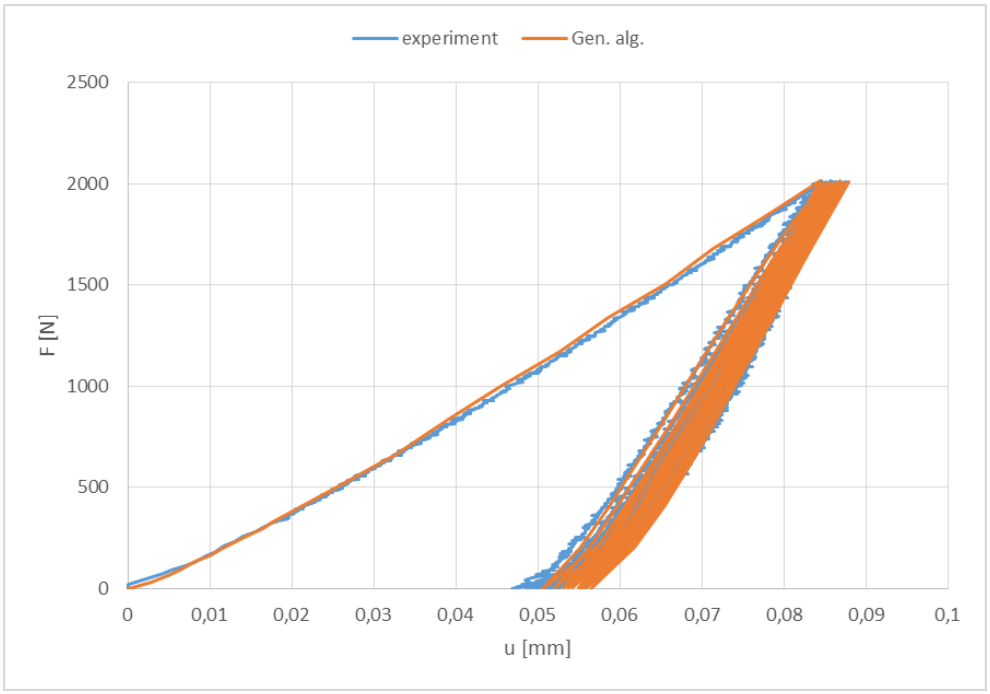


Figure 3: Comparison between numerical simulation by genetic algorithm and experimental data (F – force [N], u-depth of indentation [mm]).

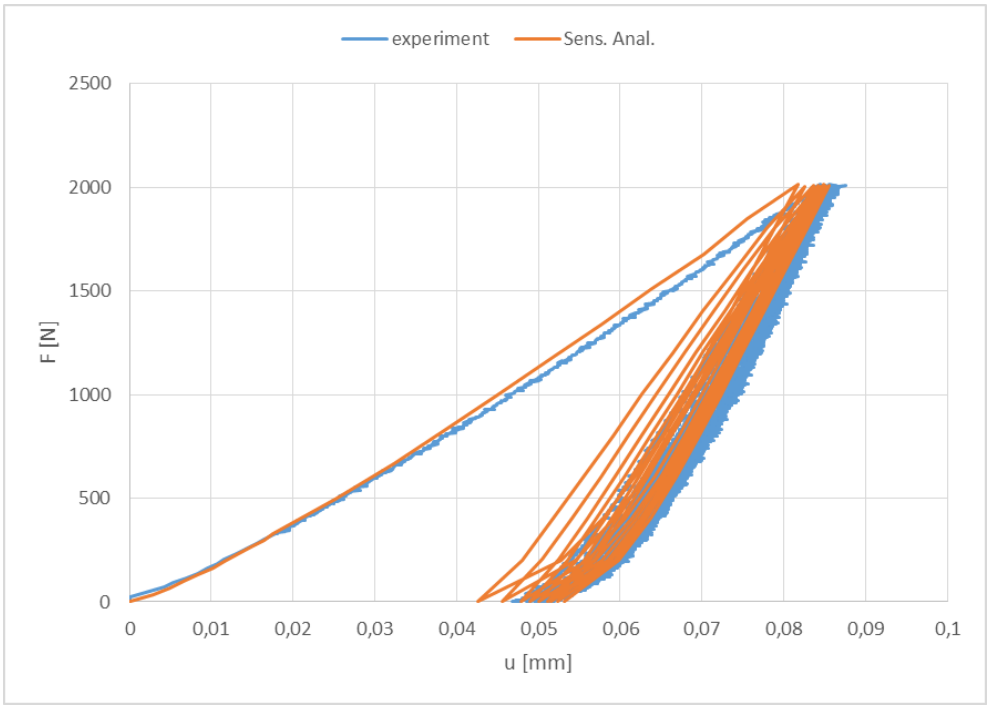


Figure 4: Comparison between numerical simulation by sensitivity analysis and experimental data (F – force [N], u-depth of indentation [mm]).

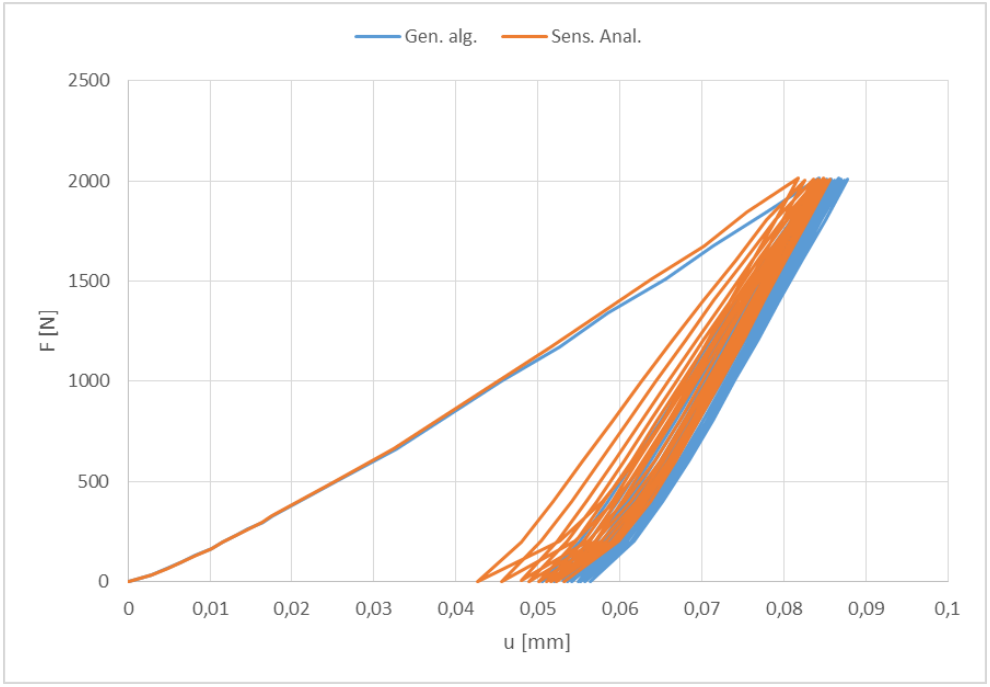


Figure 5: Comparison between numerical simulations by genetic algorithm and sensitivity analysis (F – force [N], u-depth of indentation [mm]).

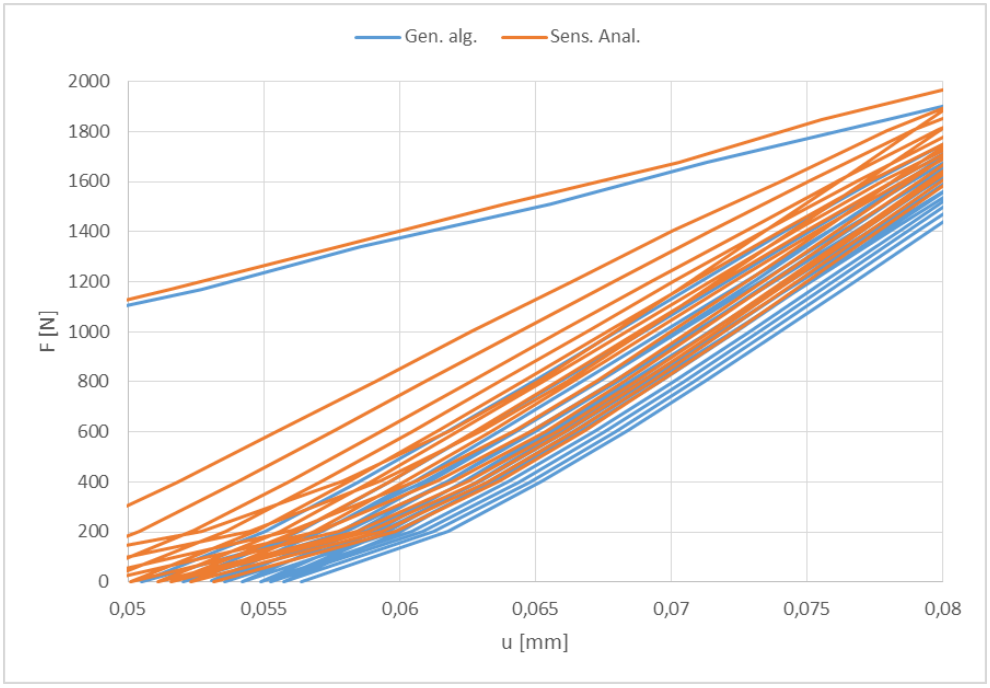


Figure 6: Detail of Figure 4 (F – force [N], u-depth of indentation [mm]).

6 CONCLUSIONS

Results presented in this paper shows that both methods i.e. genetics algorithm and sensitivity analysis are able to produce results with required precision. It is clear that genetic algorithm needs much smaller number of iterations to obtain satisfactory results. This advantage will be even more eminent if we increase number of parameters we would like to identify. A disadvantage of the sensitivity analysis is also the necessity of value interval estimation, which requires extensive experience with the used cyclic plasticity model. The paper was focused mainly on the inverse algorithm evaluation, a case study showing the correctness of estimated parameters for subsequent cyclic plasticity modelling will be discussed in a future paper.

ACKNOWLEDGMENT

The authors acknowledge the support of the project POSTDOCI II reg. no. CZ.1.07/2.3.00/30.0055 within Operational Programme Education for Competitiveness. The work was also supported the European Regional Development Fund in the IT4Innovations Centre of Excellence project (CZ.1.05/1.1.00/02.0070) and the Grant Agency of the Czech Republic (GACR) project no. 15-18274S. Authors would like to appreciate the BONATRANS GROUP a.s. Bohumin company for the help with preparation of experiments.

REFERENCES

[1] Chaboche, J.L., Dang Van, K., Cordier, G. *Modelization of The Strain Memory Effect on The Cyclic Hardening of 316 Stainless Steel*, In: 5th International Conference on Structural Mechanics in Reactor Technology, Division L11/ (Rojíček, 2010), Berlin, 13.-17. August 1979, Ed. Jaeger A and Boley B A. Berlin: Bundesanstalt fűr Materialprűfung, pp.1-10.

[2] Ohno, N. and Wang, J.D. Kinematic Hardening Rules with Critical State of Dynamic Recovery, Part I: Formulation and Basic Features for Ratchetting Behavior, International Journal of Plasticity 9, 1993, pp. 375-390.

[3] Abdel-Karim, M. and Ohno, N. *Kinematic Hardening Model Suitable for Ratchetting with Steady-State*, International Journal of Plasticity 16, 2000, pp. 225-240.¨

[4] Halama, R., Fojtík, F. and Markopoulos, A. *Memorization and Other Transient Effects of ST52 Steel and Its FE Description*, Applied Mechanics and Materials, 2013, Vol.486, p.48-53. ISSN 1662-7482.

[5] Chaboche, J.L., Lemaitre, J. Mechanics of Solid Materials. Cambridge University Press, Cambridge, 1990.

[6] Halama, R., Sedlák, J. and Šofer, M. *Phenomenological Modelling of Cyclic Plasticity, Numerical Modelling*, Peep Miidla, Ed. ISBN: 978-953-51-0219-9, InTech, 2012, pp. 329-354. Available from: http://www.intechopen.com/books/numerical-modelling/phenomenological-modelling-of-cyclic-plasticity

[7] Armstrong, P.J., Frederick, C.O. A Mathematical Representation of the Multiaxial Bauschinger Effect, G.E.G.B. Report RD/B/N, 731, 1966.

[8] Chaboche, J.L., Dang Van, K., Cordier, G., Modelization of The Strain Memory Effect

on The Cyclic Hardening of 316 Stainless Steel, Proceedings of the 5th International Conference on Structural Mechanics in Reactor Technology, Division L11/3, Berlin, Ed. Jaeger A and Boley B A. Berlin: Bundesanstalt fŏr Materialprŏfung, 1979, p.1-10.

[9] Rojíček, J. and Halama, R. *Numerical simulations of pipeline bending tests.* Appl Comp Mech 2008, Vol.2, pp.347-356.

[10] Rojíček, J. *The Set of Experiments - Analysis by Inverse Identification Method.* Journal for Theory and Practice in Metallurgy 49. (2), pp. 488-492, 2010.

[11] Cermak, M., Karasek, T., Rojicek J. *Material Parameter Identification of Modified AbdelKarim-Ohno Model*, submitted to the conference ICTCM 2015 – International Conference on Theoretical and Computational Mechanics, London, 2015.

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis (Eds)

A HIGH-ORDER FULLY COUPLED ELECTRO-FLUID-DYNAMICS SOLVER FOR MULTIPHASE FLOW SIMULATIONS

N. EMAMY*, M. KARCHER, R. MOUSAVI AND M. OBERLACK

Chair of Fluid Dynamics, Technical University of Darmstadt, Germany
*Correspondent author's email: n.emamy@fdy.tu-darmstadt.de

Key words: coupled problems, Electro-Fluid-Dynamics, multiphase flow, incompressible flow, high-order, discontinuous Galerkin method.

Abstract. *A high-order discontinuous Galerkin Finite Element solver is developed for solving electro-fluid-dynamics problems. The solver is employed to perform numerical simulations of deformation of a droplet suspended in another immiscible liquid by applying steady and oscillatory electric fields. The level set method is adopted to represent the common interface of the droplet and surrounding medium. Electrostatics equation with a jump in the dielectric property at the interface is solved to find the electric field distribution. The incompressible Navier-Stokes equations including the surface tension force are solved to find the flow field. The Electrostatics and Navier-Stokes equations are coupled through changes in the geometry because of the deformation of the droplet and the dielectrophoretic body force, which is present at the interface.*

1 INTRODUCTION

Numerical simulations of deformation of a droplet in steady and oscillatory electric fields are performed in the present study. The droplet, which is shown in figure 1, is suspended in another immiscible fluid with the same density and viscosity but a different dielectric property (permittivity). The droplet and surrounding fluid are considered as perfect dielectrics. By applying an electric field, the fluids are polarized and because of the jump in the dielectric property, the dielectrophoretic force exerts at the interface of the droplet and surrounding fluid. The droplet continues to deform until a force balance between the electric force, pressure and surface tension force is achieved and the droplet becomes a spheroid, see e.g. Torza et al. [1]. The deformation of the droplet is defined in figure 1.

A two-way coupling exists between the fluid and electric sub-problems. On one-hand, the electric force exerts at the interface of the droplet and surrounding fluid and on the other hand, the deformation of the droplet changes the geometry for the electric field computation. Therefore, an electromechanical approach is required, which includes solving the governing equations of both electric and fluid fields, computing the electric force and capturing the movement of the interface of the droplet and surrounding fluid. Supeene et al. [2] have considered a moving mesh approach to find the movement of the interface, which is suitable for small deformations. Hua et al. [3] have used a front tracking/finite volume method. In the present study, a high-order discontinuous Galerkin Finite Element method (DG) is employed and a one-fluid approach is followed, which enables us to solve one set of the governing equations for the droplet and surrounding fluid.

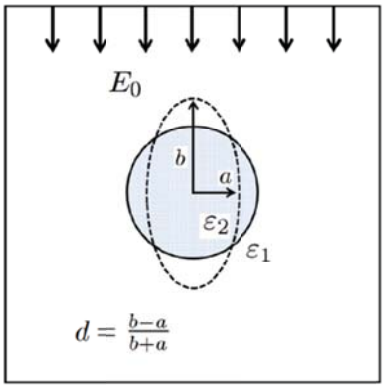


Figure 1: A droplet suspended in another immiscible liquid with a different dielectric property. Deformation of the droplet, d, in response to applying an electric field is shown.

2 METHODOLOGY

The interface is represented as the zero iso-value of a level set function, φ. The governing equations are the electrostatics, continuity, incompressible Navier-Stokes and level- set advection equations:

$$\begin{aligned}
&\nabla\cdot\left(\varepsilon(\varphi)\nabla\Phi\right)=0,\\
&\vec{\mathrm{E}}=-\nabla\Phi,\\
&\nabla\cdot\vec{\mathrm{u}}=0,\\
&\frac{\partial\vec{\mathrm{u}}}{\partial\mathrm{t}}+\vec{\mathrm{u}}\cdot\nabla\vec{\mathrm{u}}=-\nabla\mathrm{p}+\frac{1}{\mathrm{Re}}\Delta\vec{\mathrm{u}}-\frac{1}{\mathrm{We}}\delta(\varphi)\kappa\vec{\mathrm{n}}-\frac{1}{2}|\vec{\mathrm{E}}|^2\nabla\varepsilon(\varphi),\\
&\frac{\partial\varphi}{\partial\mathrm{t}}+\vec{\mathrm{u}}\cdot\nabla\varphi=0,
\end{aligned}\tag{1}$$

where ε represents the dielectric property and Φ is the electric potential. For the case of steady or slowly oscillating electric fields, the electrostatics equations are valid, where the electric field, $\vec{E}$, is defined as the gradient of the electric potential and a Laplace equation is solved to compute the electric potential. A diffuse interface model is used to regularize the jump in the dielectric property at the interface. A signed-distance level set function is used in combination with the diffuse interface model to assure the same thickness of the interface everywhere. The solution algorithm is shown in figure 2. For more information see Emamy [4].

The governing equations are discretized using the DG method within an in-house CFD code, Kummer [5]. The DG method employs a high-order local polynomial representation of the solution. However, the locality of the solution makes it discontinuous at the cell boundaries. To solve the incompressible Navier-Stokes and continuity equations, a projection scheme is employed. Considering that the inter-cell discontinuity of the solution may be large in case of low resolution (coarse grids/low polynomial degrees), the projection scheme is adapted for the DG method, Emamy [6], to provide a long-term stable and accurate scheme. Moreover, a DG weak-formulation is used to compute derivatives of the flow field variables and electric potential.

The surface tension force is modeled using the continuum surface force model (CSF), Brackbill et al. [7]. To decrease the intrinsic spurious velocities, the surface tension force is computed by using high-order polynomials for computation of the normal vector, $\vec{n}$, and curvature, κ. The normal vector is computed as the gradient of a signed-distance level set function, φ_{SD}. The curvature is the divergence of the normal vector by definition. A reinitialization equation is solved in each time step to find φ_{SD} from the level set function φ, which provides the position of the interface by solving the level set advection equation, Mousavi [8].

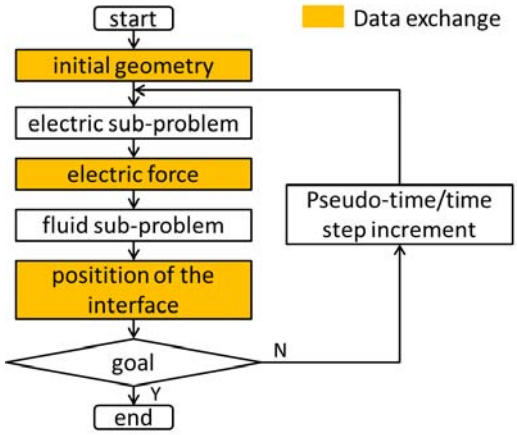


Figure 2: The solution algorithm for the coupled Electro-Fluid-Dynamics problem.

To find the electric potential, the Laplace equation is discretized using an interior penalty method (IP-MCP), which Emamy [4] has provided for the case of a regularized jump in the coefficient (dielectric property). After solving for the electric potential, the electric field and electric force are computed. The electric force is added as a body force to the Navier-Stokes equations, (1).

3 NUMERICAL RESULTS

To perform the numerical simulations, a Reynolds numbers of 1 is considered because the physical problem, which we consider, is a creeping flow. The numerical settings and boundary conditions from Karcher [9] are shown in figure 3. In figure 4, a test case with a negligible surface tension force is considered. In absence of the surface tension force, the droplet continues to deform in response to the electric force and there is no equilibrium state. The droplet preserves the shape of an ellipse while deforming. A convergence study for this test case is performed by grid refinement, which is shown in figure 5. Using polynomial degrees of 5 for the flow field variables, a convergence rate of 3.25 is achieved for the coupled problem. A reduced convergence rate is expected because of the regularization of the jump in the dielectric property at the interface.

Including the surface tension force, in figure 6, effect of the size of the computational domain and boundary conditions on the deformation of the droplet in the equilibrium state, d_∞, is studied. Dirichlet velocity (wall) and Dirichlet pressure boundary conditions are compared. For larger domains, influence of the boundary conditions becomes smaller. For

this test case, which includes the surface tension force, we have applied the capillary time step size restriction [7] and modified it for the high-order DG method as

$$\Delta t_C = \sqrt{\frac{We\, h^3}{2\pi\, k^6}}, \tag{2}$$

where, We is the Weber number, h is the element size and k is the polynomial degree for the pressure and velocity.

As a final test case, an oscillating electric potential difference with frequency of 1 is applied as the boundary condition. Deformation of the droplet versus time is shown in figure 7, which has a period of 0.5. This means that the droplet oscillates with a frequency that is twice the excitation frequency. This frequency is expected by the presence of $|\vec{E}|^2$ in the formulation of the dielectrophoretic force as a body force in the Navier-Stokes equations.

4 CONCLUSIONS

The coupled electro-fluid-dynamics simulations, predict the expected physical behavior of a perfect dielectric droplet suspended in another perfect dielectric immiscible fluid when steady and oscillatory electric fields are applied. Applying a steady electric field the droplet deforms to an ellipse for 2D simulations. In case of the negligible surface tension force the droplet continues to deform until it bursts. In presence of the surface tension force the droplet reaches an equilibrium state. If the computational domain is large enough the effect of boundary conditions on the deformation in equilibrium state is negligible. Considering the surface tension force, the capillary time step size restriction is modified for the high-order DG method. In case of an oscillatory electric field the droplet oscillates with a frequency, which is twice the excitation frequency.

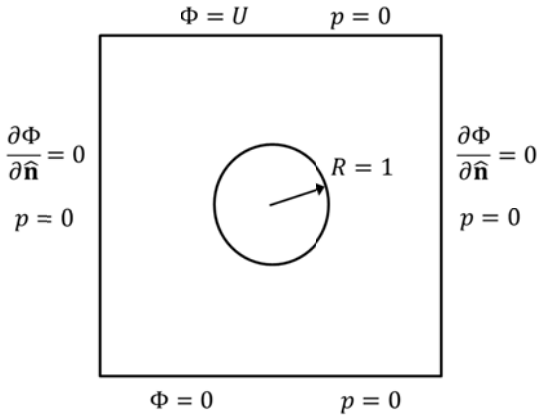


Figure 3: A droplet suspended in another immiscible liquid with a different dielectric property. Boundary conditions for the fluid and electric sub-problems are shown.

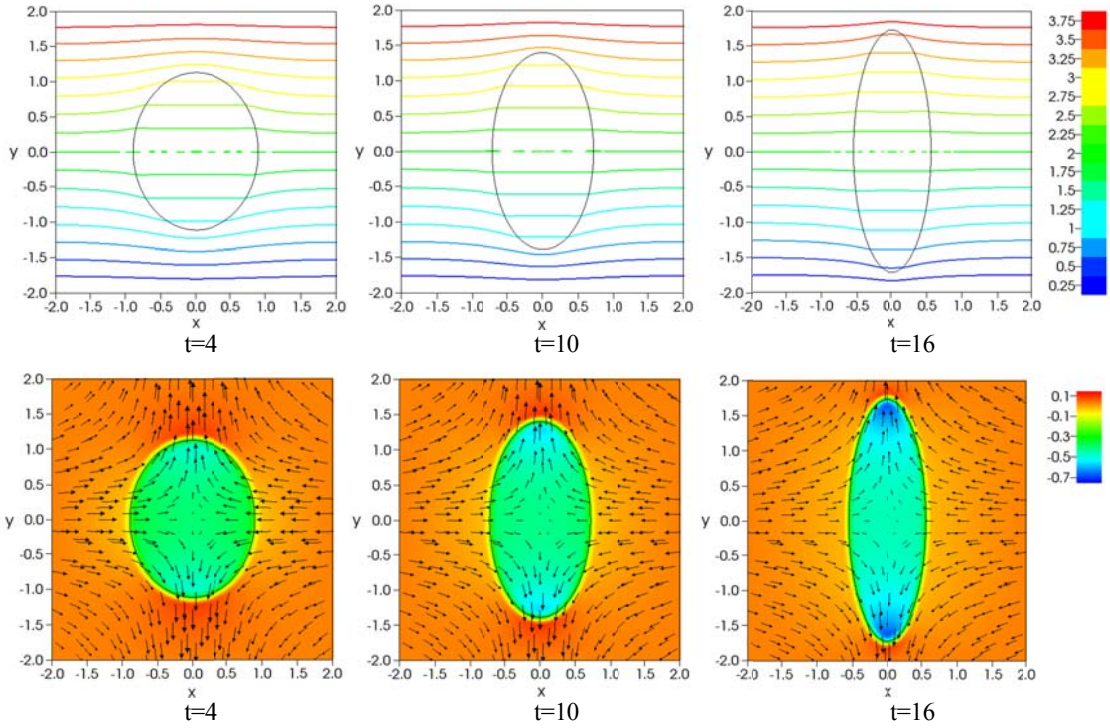


Figure 4: A droplet, which is suspended in another immiscible fluid with a jump ratio of 2 in the dielectric property at the interface, is considered. The surface tension force is neglected. Iso-lines of the electric potential are shown on the first row. Pseudo-colors of the pressure and vectors of the velocity are shown on the second row. An electric potential difference of 4 is applied. The computational domain is a [-2, 2]×[-2, 2] square. The half-thickness of the regularized interface is 1/8. A Cartesian grid of 24×24 cells is employed. Polynomial degree of 8 for the electric potential and 5 for the pressure, velocity and level set function are used. A time step size of 0.01 is used.

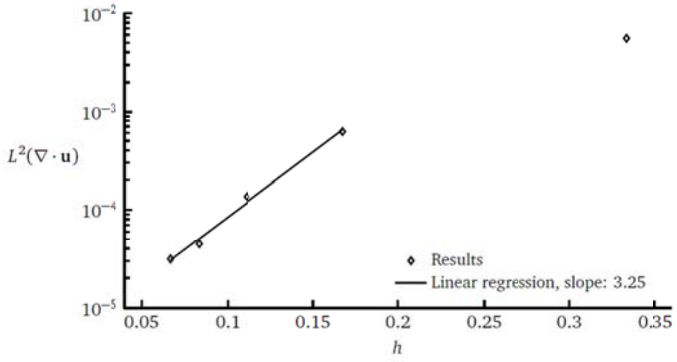


Figure 5: Convergence study is performed for the test case of a droplet suspended in another immiscible fluid with a jump ratio of 2 in the dielectric property at the interface. L2-norm of divergence of the velocity is shown, as a measure for the error, vs. the element size, h, after the first time step. The computational domain is a [0, 2]×[-2, 2] quadrilateral, where a symmetry boundary is used at x=0 axis. The half-thickness of the regularized interface is 1/7. Polynomial degree of 8 for the electric potential and 5 for the pressure, velocity and the level set function are used. A time step size of 0.01 is applied.

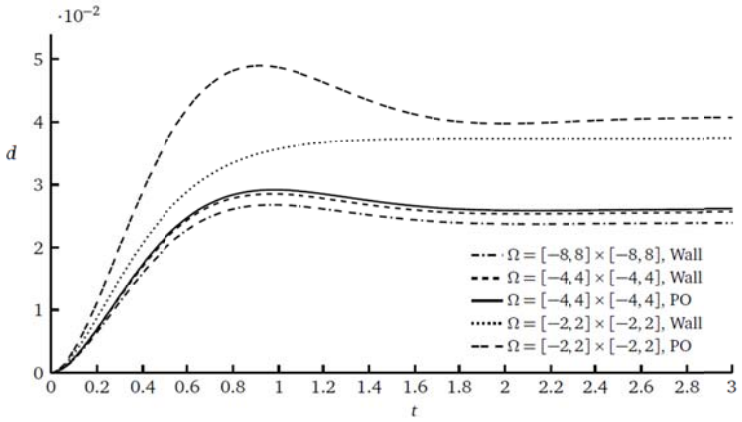


Figure 6: Deformation, d, vs. time, t, is shown for a test case with a jump ratio of 10 in the dielectric property at the interface and a Weber number of 0.1. Different domain sizes and boundary conditions are tested. Wall stands for a no-slip boundary condition for the velocity and pressure-outlet (PO) stands for a Dirichlet boundary condition for the pressure. The half-thickness of the regularized interface and the element size of the Cartesian grids are 1/8. Polynomial degrees of 8 for the electric potential and 5 for the pressure, velocity and the level set function are used. A time step size of 5×10-4 is used by considering the capillary time step size restriction.

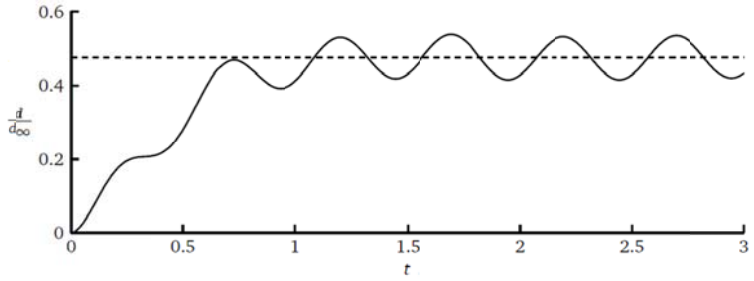


Figure 7: Relative deformation, d/d∞, is shown vs. time, t, for a test case of an oscillating droplet. An oscillating potential difference of 4cos(2πt) is applied. A jump ratio of 10 in the dielectric property at the interface and a Weber number of 0.1 are considered. The dashed line shows the relative mean value of 0.48. The computational domain is a [-2, 2]×[-2, 2] square with the wall boundary conditions. The half-thickness of the regularized interface and the element size of the Cartesian grid are 1/8. Polynomial degrees of 8 for the electric potential, 5 for the pressure and velocity and 4 for the level set function are used. A time step size of 5×10-4 is used by considering the capillary time step size restriction.

REFERENCES

[1] Torza, S., Cox, R. G. and Mason, S. G. Electrohydrodynamic Deformation and Burst of Liquid Drops. *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences* (1971) 269.1198, 295–319.

[2] Supeene, G., Koch, C. R. and Bhattacharjee, S. Deformation of a droplet in an electric field: Nonlinear transient response in perfect and leaky dielectric media. *Journal of Colloid and Interface Science* (2008) 318.2, 463–476.

[3] Hua, J., Lim, L. K. and Wang, C.-H. Numerical simulation of deformation/motion of a drop suspended in viscous liquids under influence of steady electric fields. *Physics of Fluids* (2008) 20.11, 113302.
[4] Emamy, N. *Numerical simulation of deformation of a droplet in a stationary electric field using DG*. Dissertation. TU Darmstadt, Darmstadt (2014).
[5] Kummer, F. *The BoSSS Discontinuous Galerkin solver for incompressible fluid dynamics and an extension to singular equations*. Dissertation. TU Darmstadt/Fachgebiet für Strömungsdynamik (2012).
[6] Emamy, N. *Projection scheme for solving the incompressible Navier-Stokes equations using DG*. Annual report for the DFG. TU Darmstadt/Fachgebiet für Strömungsdynamik (2013).
[7] Brackbill, J. U, Kothe, D. B and Zemach, C A continuum method for modeling surface tension. *Journal of Computational Physics* (1992) 100.2, 335–354.
[8] Mousavi, R.: *Level Set Method for Simulating the Dynamics of the Fluid-Fluid Interfaces: Application of a Discontinuous Galerkin Method*. Dissertation. TU Darmstadt (2014).
[9] Karcher, M. *Numerical 3D simulation of deformations of a droplet in an oscillatory electric field using DG*. Master's project, TU Darmstadt (2014).

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis (Eds)

ANALYSIS ON FLOW AROUND A SPHERE AT HIGH MACH NUMBER, LOW REYNOLDS NUMBER AND ADIABATIC CONDITION FOR HIGH ACCURACY ANALYSIS OF GAS PARTICLE FLOWS

T. NAGATA[*], T. NONOMURA[†], S.TAKAHASHI[†], Y.MIZUNO[†] AND K.FUKUDA[†]

[*] Tokai University, 4-1-1, Kitakaname, Hiratsuka, Japan, e-mail: 1beu2216@mail.tokai-u.jp

[†]Japan Aerospace Exploration Agency, Institute of Space and Astronautical Science (JAXA/ISAS), 3-1-1, Yoshinodai, Sagamihara, Japan, e-mail: nonomura@flab.isas.jaxa.jp

[†]Tokai University, 4-1-1, Kitakaname, Hiratsuka, Japan, e-mail: takahasi@mail.tokai-u.jp

[†]Tokai University, 4-1-1, Kitakaname, Hiratsuka, Japan, e-mail: 1beu2104@mail.tokai-u.jp

[†]Tokai University, 4-1-1, Kitakaname, Hiratsuka, Japan, e-mail: fukuda@mail.tokai-u.jp

Key words: High Mach number and Low Reynolds number flow, Drag model

Abstract. This study analyses gas particle flow around a sphere under an adiabatic condition at high Mach number and low Reynolds number by direct numerical simulation of the three–dimensional compressible Navier–Stokes equation to investigate flow properties. The calculation was performed on a boundary-fitted coordinate system with a high-order scheme of sufficient accuracy. Analysis is conducted by assuming a rigid sphere with a Reynolds number based on the diameter of the sphere, and the free-stream velocity set between 50 and 300 and a free-stream Mach number set between 0.3 and 2.0. The effect of the Mach number on the flow properties and drag coefficient are discussed. The calculation shows the following results: 1) unsteady fluctuation of the hydrodynamic force becomes smaller as the Mach number increases, 2) the drag coefficient increases along with the Mach number due to an increase in the pressure drag by the shock-wave, and 3) an accurate prediction of the drag coefficient in the supersonic regime using traditional models might be difficult.

1 INTRODUCTION

Certain acoustic phenomena are caused by fluid behavior. In particular, the exhaust gas from a rocket engine generates a strong acoustic wave. The acoustic waves reflected from the ground surface and launch facility, causes vibration of the payload in the fairing. Therefore, prediction and reduction of the acoustic level at lift-off are necessary. Traditionally, the acoustic level has been predicted by a semi-empirical method such as NASA SP-8072 [1] and subscale tests [2]. NASA SP-8072 is based on a large amount of flight data and results of static firing tests that were conducted by the organization in the United States; it does not

consider the effects of the launch pad and facility because NASA SP-8072 assumes that sound sources are distributed along a free jet. Therefore, it is necessary to create a model to modify the prediction results for each launch pad. Hence, NASA SP-8072 is not suitable as a design tool for new launch pads. Moreover, static firing tests are costly. Recent, predictions of the acoustic level using computational fluid dynamics (CFD) are required for these reasons.

The alumina particles released from solid rocket motors and water droplets introduced by water injection to the exhaust gas might attenuate the acoustic wave. However, the attenuation mechanism is not well known. Recently, the effects of water droplets have been verified [3,4]. In previous studies, Tsutsumi et al. [5,6] performed analyses in consideration of the effects of the launch facility and the flame deflector plate. In addition, an analysis considering the difference of the components of the exhaust and the atmosphere was performed by Nonomura et al. [7]. Consequently, the acoustic phenomena have been clarified. If it becomes possible to consider the effect of the particles on the analysis, it is expected that acoustic wave level prediction using CFD will become more accurate than predictions under the present method. The diameters of the alumina particles released from solid rocket motors are 30-200 μm [8], and the exhaust gas is a supersonic flow. Therefore, flow around each particle has a high Mach number and a low Reynolds number. Accordingly, a drag model under a high Mach number and a low Reynolds number condition is necessary to perform analysis that considers the influence of the particle on the jet flow. In this study, direct numerical simulation (DNS) of the flow around a sphere at high Mach number and low Reynolds number is conducted to construct a subgrid scale and body-force models that consider the influence of particles. Analyses are performed by assuming that the alumina particle is a rigid sphere; the Reynolds number based on the diameter of the sphere and the free-stream velocity is set between 50 and 300, the free-stream Mach number is set between 0.3 and 2.0 (Table 1).

2 ANALYTICAL METHOD

2.1 Computational grid

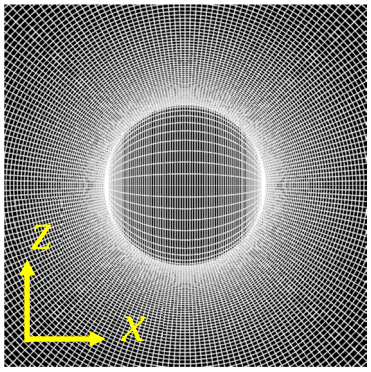
A boundary-fitted grid is adopted for this computation. The dimensions of the grid are $\xi \times \eta \times \zeta = 107 \times 48 \times 177$ points, with 909,072 grid points. The boundary condition at the sphere's surface is nonslip and adiabatic. At the boundaries of the ξ and η directions, periodic boundary conditions with three overlapped grid points are imposed (Figure 1). The diameter of the analysis region is 100 times that of the sphere. Figure 1 shows a computational grid. The sphere's diameter is 1D. The grid size of the ζ direction is stretched by 1.03 times from the minimum grid width within 15*D* of the origin, and the grid size is constant after it reaches 0.2*D*. In regions 15*D* or more away from the origin point, the grid size is stretched by 1.2 times toward the outer boundary. The minimum grid size is calculated by the following formula [9.]:

$$dr_{\min} = \frac{1.13}{\sqrt{\mathrm{Re}} \times 10.0}. \tag{1}$$

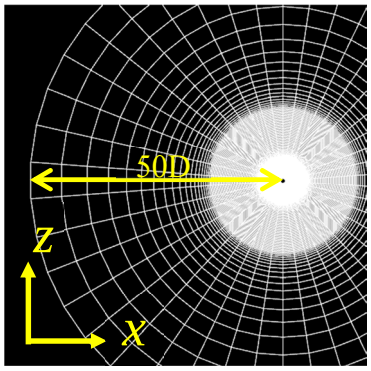
In this study, the minimum grid for all calculations is fixed with the size determined by assuming that the Reynolds number is 300.

Table 1: Analysis case

Reynolds number	Mach number	Cace.
50	0.30	Re050M030
	0.80	Re050M080
	1.20	Re050M120
	2.00	Re050M200
100	0.30	Re100M030
	0.80	Re100M080
	1.20	Re100M120
	2.00	Re100M200
150	0.30	Re150M030
	0.80	Re150M080
	1.20	Re150M120
	2.00	Re150M200
200	0.30	Re200M030
	0.80	Re200M080
	1.20	Re200M120
	2.00	Re200M200
250	0.30	Re200M030
	0.80	Re250M080
	1.20	Re250M120
	2.00	Re250M200
300	0.30	Re300M030
	0.70	Re300M080
	0.80	Re300M120
	0.95	Re300M200
	1.05	Re300M030
	1.20	Re300M080
	1.50	Re300M120
	2.00	Re300M200



a. Close view



b. Far view

Figure 1: Computational grid

2.2 Computational method

In this calculation, the Navier–Stokes equation is employed as the governing equation. The equation is non-dimensionalized by the free-stream speed of sound, the density, and the diameter of the sphere. The convection term is evaluated by the WENO-CU6-FP [10] method, the viscous term by the sixth-order central difference method, and time integration is performed by the third-order TVD Runge–Kutta method. WENO-CU6-FP is the WENO method satisfying the geometric conservation laws in the curvilinear coordinate system, which was proposed by Nonomura et al. [10]. This scheme can preserve the free-stream. In this study, the central difference component of the WENO-CU6-FP scheme is replaced by one of the splitting type [11] in order to better stabilize the calculation.

3 COMPUTATIONAL RESULT OF FLOW AROUND A SPHERE

3.1 Flow regime

3.1.1 Pressure coefficient distribution and vortex structure

Figures 2 and 3 are the snapshots of the moment field. Figure 2 shows the pressure

coefficient distribution and the stream lines. Figure 3 shows the entire view of the vortex structure visualized by the isosurface of the second invariant of a velocity gradient tensor. The threshold is set to the appropriate value for each case. From the visualization, the pressure coefficient distribution is axisymmetric and a hairpin vortex is observed in the subsonic flow. Conversely, the pressure coefficient distribution becomes almost symmetric under transonic and supersonic conditions, and a clear hairpin vortex structure disappears as the Mach number increases. The hairpin vortex structure also disappears at Mach 0.95. In addition, at Mach 0.95, the shock-wave is formed at the downstream side of the sphere.

3.1.2 Separation point and separation length

The separation point is probed by considering the velocity gradient on the sphere surface. In this study, the separation point is expressed as the degree from the x axis, as shown in Figure 4. Separation length is expressed as the distance from the sphere's surface to the separation regime termination. In this study, the separation length is probed by the u direction component of the flow velocity on the x axis of the downstream side. The end of the separation region is set to be the position where the sign of the u component velocity switches (Figure 4). Figures 5 and 6 show the separation point and separation length of the time-averaged field, respectively. The separation point location and separation length change significantly in the transonic region. The separation point moves to the upstream side at subsonic flow and to the downstream side from transonic to supersonic flow. Conversely, the separation length increases under the subsonic and transonic flows and decreases at the supersonic flow. The separation length rapidly increases at Mach 0.95. In addition, the separation length is hardly changed; therefore, the separation point is influenced by the presence or absence of the shock-wave (i.e., $M < 1$ or $M > 1$), and the separation length is gradually influenced by the effects of compressibility. Hence, there is a need to discuss the interference of the shock-wave with the wake.

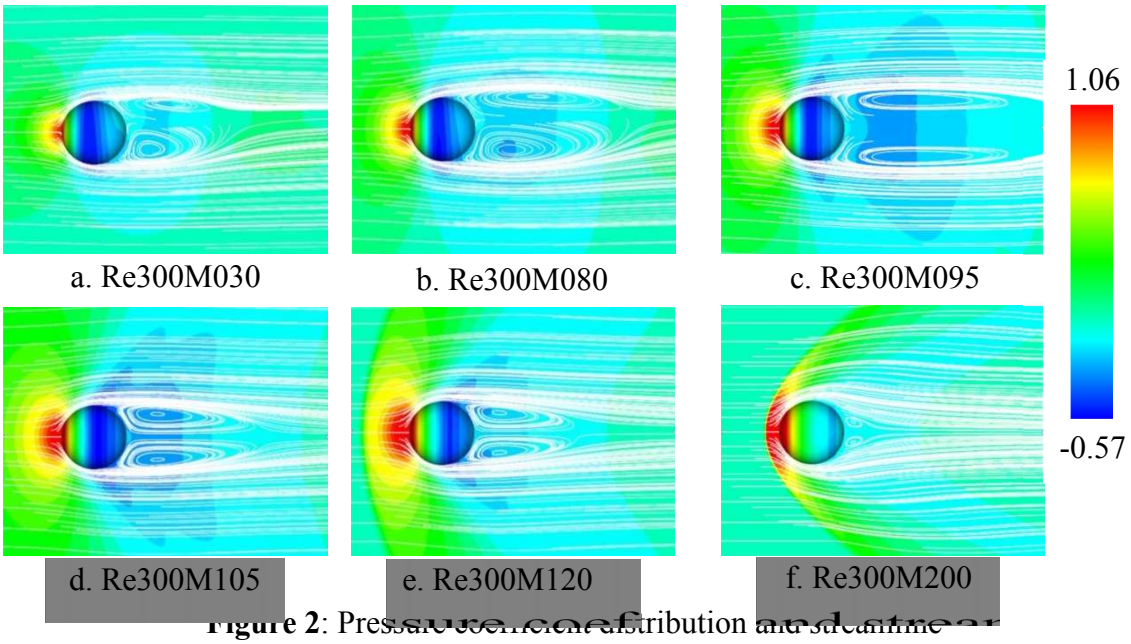
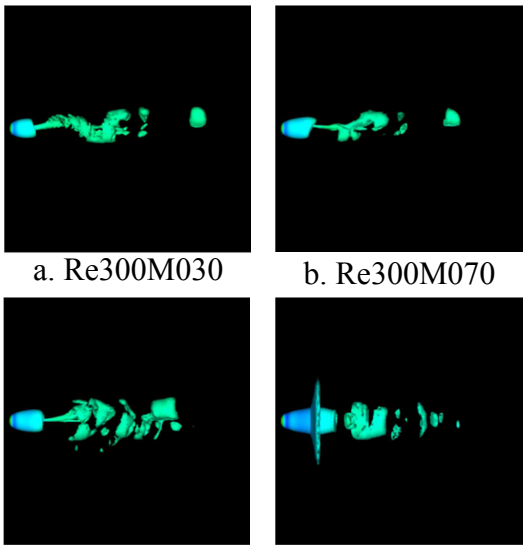


Figure 2: Pressure coefficient distribution and streamline



a. Re300M030 b. Re300M070

c. Re300M080 d. Re300M095

Figure 3: Isosurfaces of the second invariant of the velocity gradient tensor

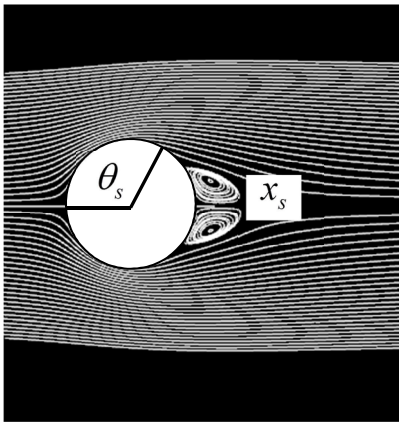


Figure 4: Separation point θ_s and separation length x_s

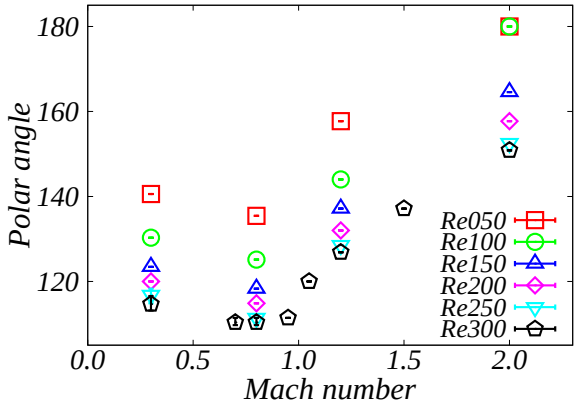


Figure 5: Separation point

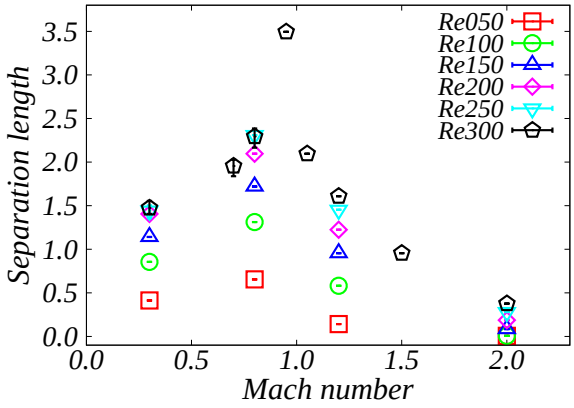


Figure 6: Separation length

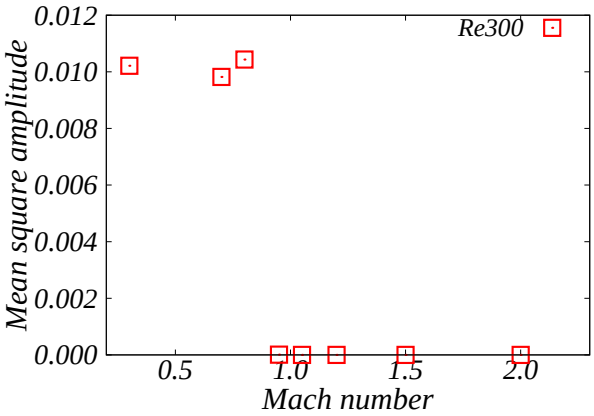


Figure 7: Mean square amplitude

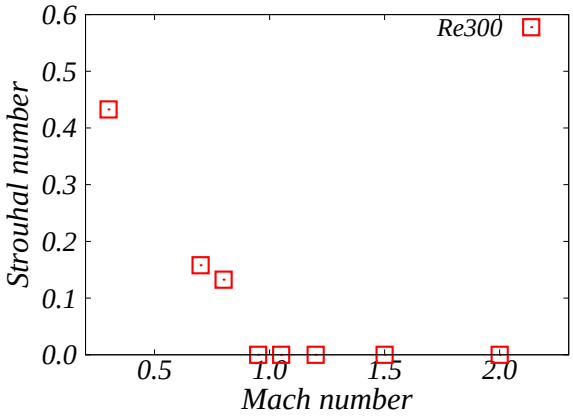


Figure 8: Strouhal number

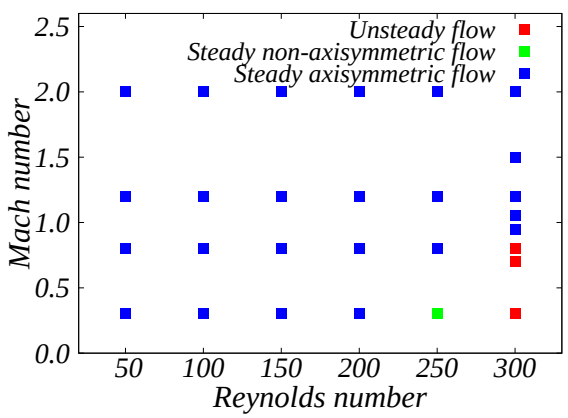


Figure 9: Flow regime

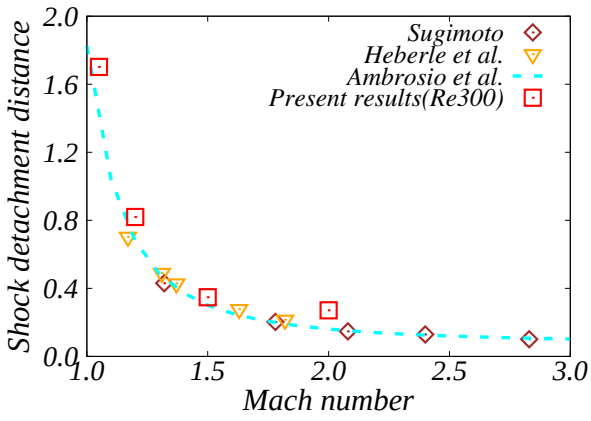


Figure 10: Shock detachment distance at Re300

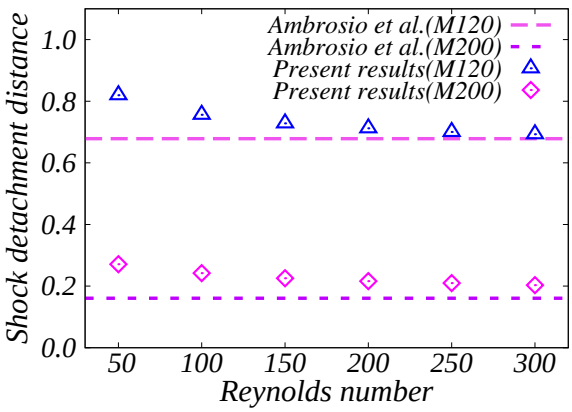


Figure 11: Influence of the Reynolds number on the shock detachment distance

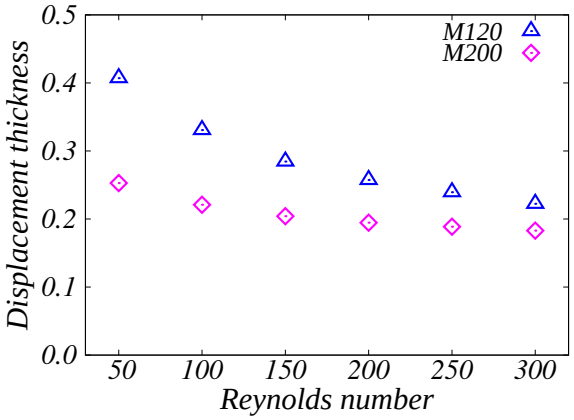


Figure 12: Displacement thickness

3.1.2 Time variation of the flow field

Figure 7 shows the root-mean-square amplitude of the lift coefficient when the Reynolds number is 300. The lift coefficient vibrates in the case of Reynolds number 300 and Mach numbers 0.3, 0.7, and 0.8. It is estimated that the phenomenon appears with the generation of the hairpin vortex. Therefore, flow fields are unsteady in these cases. Conversely, the flow fields under the transonic and supersonic conditions are steady. In addition, the root-mean-square amplitude of the lift coefficient is mostly not influenced by the Mach number at Reynolds number 300. Hence, the flow regime is influenced by the presence or absence of the shock wave. Moreover, Figure 8 shows the Strouhal number based on the lift coefficient at Reynolds number 300. The Strouhal number decreases as the Mach number increases. This shows that the fluctuation frequency of the body force is influenced by the Mach number. Figure 9 shows the flow regime of each case. The red plot shows unsteady flow, the green plot shows steady non-axisymmetric flow, and the blue plot shows the steady axisymmetric

flow. Flow becomes unsteady under the subsonic condition at Reynolds number 300. Conversely, flow becomes steady under the transonic and supersonic conditions at Reynolds number 300. In addition, when the Mach number is 0.3 and the Reynolds number is 250, there is a steady axisymmetric flow. At Mach numbers greater than 0.8, flow becomes steady and axisymmetric. Moreover, the flow regime at Mach 0.3 is the same as the results of previous studies of incompressible flow [9]. Hence, it appears that the flow field becomes steady and axisymmetric when the Mach number increases.

3.1.3 Shock detachment distance

Figures 10 and 11 show the comparisons of the shock detachment distance of the previous study and present results, respectively. Sugimoto et al. [12] and Haberle et al. [13] showed experimental results. Ambrosio et al. [14] showed a prediction model of the shock detachment distance that was constructed from the experimental data. From Figure 11, the present results agree well with Ambrosio et al. at Reynolds number 300. The model does not include the effects of the Reynolds number. The present results show that the shock standing distance increases as the Reynolds number decreases. Figure 12 shows the displacement thickness estimated by considering the flat plate boundary layer thickness. From Figure 12, the displacement thickness increases as the Reynolds number decreases. Hence, the shock detachment distance increases because the effective size of the object increases. Therefore, the fluctuation of the shock detachment distance is influenced by the thickness of the boundary layer. At low Reynolds numbers, the boundary layer is thicker than that at high Reynolds numbers. It is believed that the difference is caused by the boundary layer thickness.

3.2 Drag coefficient

In Figure 13–15, changes in the drag coefficients and their components as functions of the Mach number are shown. Figures 13, 14, and 15 show the drag coefficients, the pressure drag coefficients, and the viscous drag coefficients, respectively. Figure 12 shows that the drag coefficient increases as the Mach number increases. In particular, the drag coefficient increases rapidly in the transonic flow. Conversely, in the supersonic flow, the drag coefficient is hardly changed. In Figure 13, the pressure drag coefficient is influenced by the Mach number. In contrast, the viscous drag coefficient is hardly changed. In the below sections, the pressure and viscous drag coefficients will be discussed in detail.

3.2.1 Pressure drag coefficient

Figure 13 shows that the pressure drag coefficient increases along with the Mach number. Under the subsonic and transonic conditions, it rapidly increases. However, under the supersonic condition, it hardly changes. In addition, the pressure coefficient decreases in the low Reynolds number case of the supersonic flow. The pressure drag coefficient decrease under the subsonic condition occurs by the moving of the separation point to the upstream side. Under the transonic condition, the separation point moves to downstream side; however, the pressure drag coefficient increases. This appears to be because of the formation of the shock wave.

3.2.2 Viscous drag coefficient

Figure 14, illustrates that the viscous drag coefficient is hardly influenced by the Mach number. In an incompressible flow, the viscous drag coefficient is influenced by the behavior of the separation point. However, Figures 6 and 14 show a different trend. In addition, Figure 15 shows the trend of the increase in non-dimensional viscosity coefficients with increasing Mach number due to aerodynamic heating.

In subsonic flow, the separation point moves to the upstream side, and the viscous drag becomes smaller due to the decrease of the attached flow area. Therefore, the viscous drag coefficient does not change much because the increase in the viscosity coefficient excludes the effects of the separation point. Conversely, in supersonic flow, the separation point moves to the downstream side, and the viscosity coefficient increases, though the viscous drag coefficient hardly does so. This is because of the deceleration. Figure 16 shows the velocity gradient of the u direction component at the one layer from the sphere surface normalized by the free-stream. From Figure 16, the velocity gradient decreases in the supersonic flow. This seems to be caused by the deceleration of the fluid at the detached shock waves. For these reasons, it is estimated that the viscous drag coefficient does not significantly increase under subsonic and supersonic conditions.

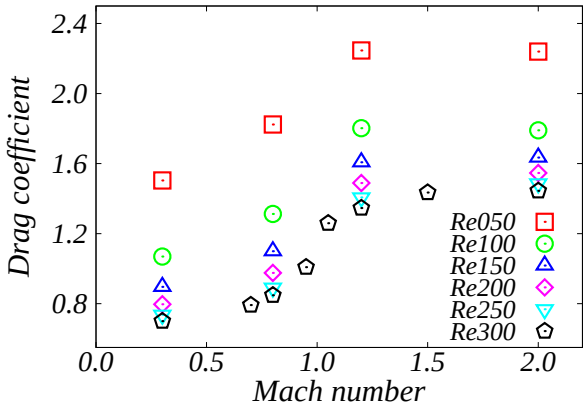


Figure 13: Drag coefficient

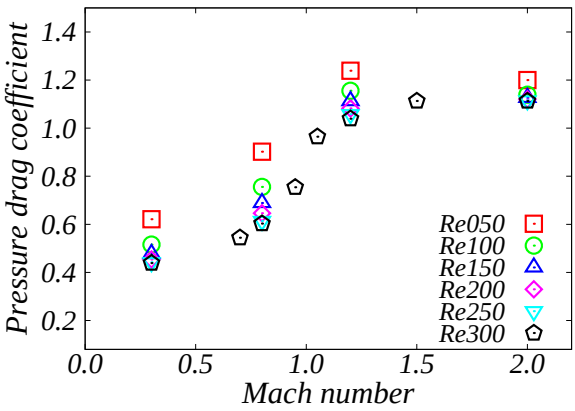


Figure 14: Pressure drag coefficient

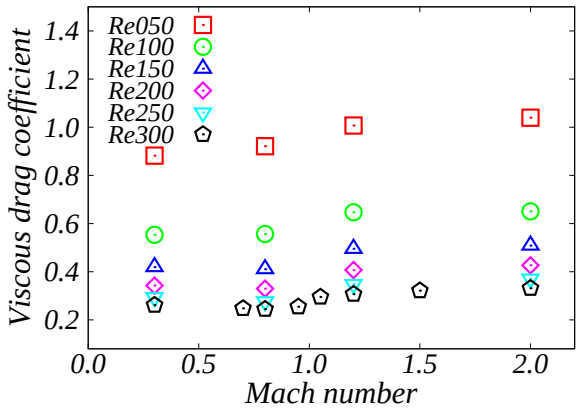


Figure 15: Viscous drag coefficient

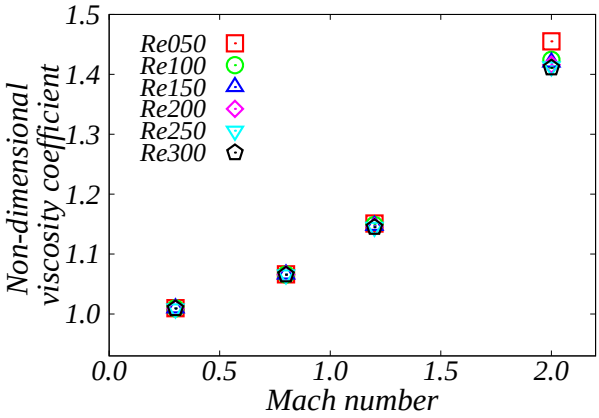


Figure 16: Non-dimensional viscosity coefficient on the sphere surface

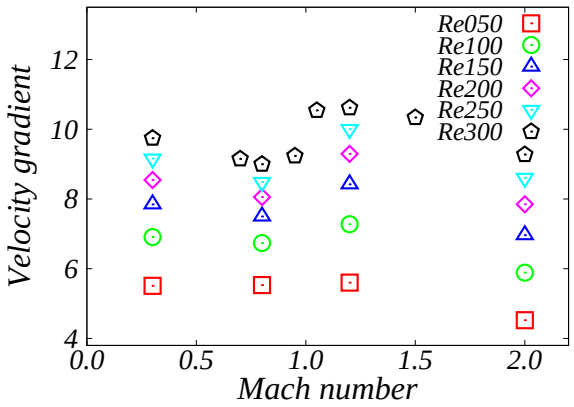


Figure 17: Mean value of the x-direction component of the velocity gradient on the sphere surface

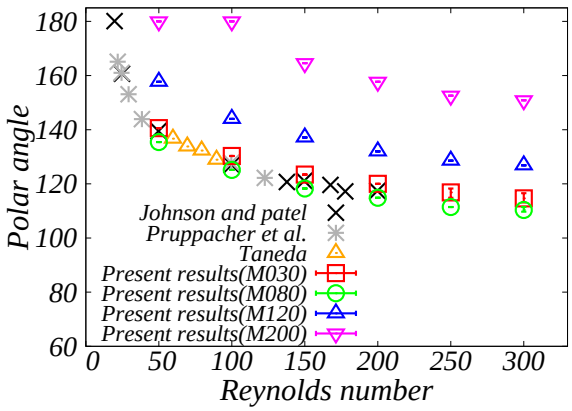


Figure 18: Separation points of the analysis results and the previous study results

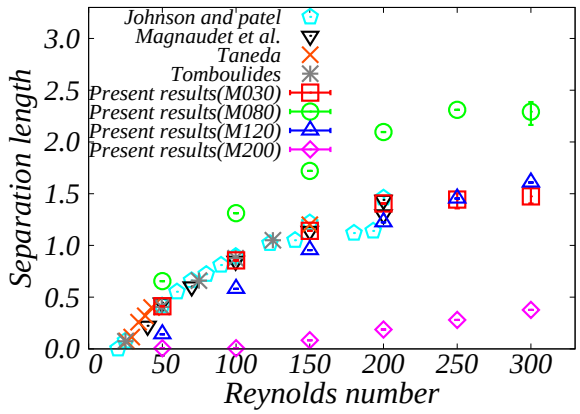


Figure 19: Separation lengths of the analysis results and the previous study results

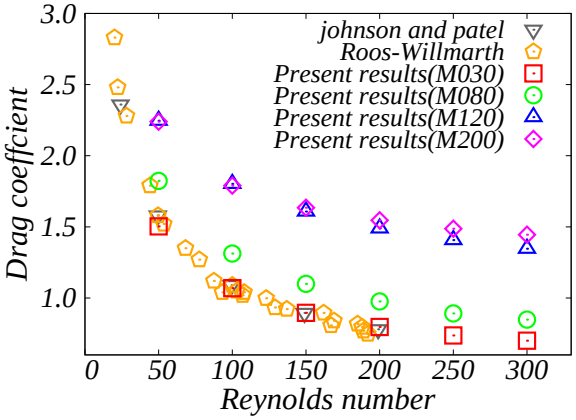


Figure 20: Drag coefficients of the analysis results and the previous study results

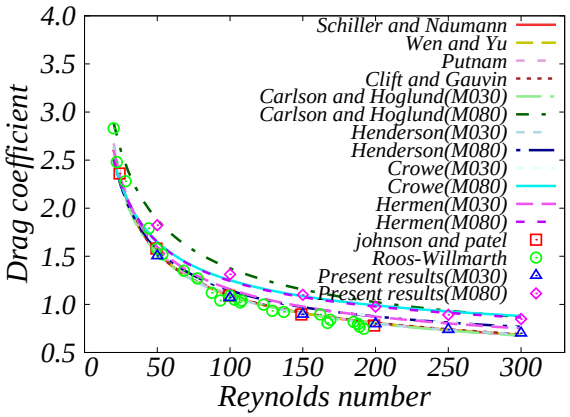


Figure 21: Analysis results and values predicted by drag models at the subsonic

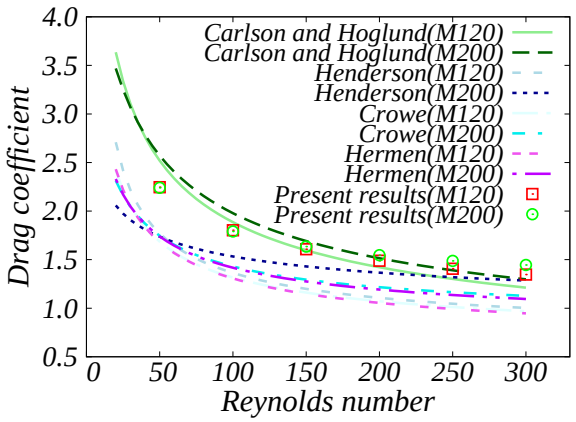


Figure 22: Analysis results and values predicted by drag models at the subsonic

3.3 Comparison with incompressible flow

3.3.1 Flow regime

Under incompressible flow, the flow regime is same if the Reynolds number is matched. However, in compressible flow, both the Reynolds number and the Mach number need to be considered. Figure 17 shows a comparison of the separation point. The present results in the subsonic condition are in good agreement with incompressible results. However, the difference in the separation point becomes larger in transonic and supersonic flows. Therefore, it is estimated that the separation point is affected by the presence and intensity of the detached shock wave. Conversely, the separation length is different at Mach number 0.8 as compared with the results of the incompressible flow (Figure 18). Therefore, separation length appears to be gradually affected by the compressibility.

3.3.2 Drag coefficient

Figures 19–21 show the drag coefficient. Figure 19 shows the results in the incompressible flow condition [9] along with the results of this study. At Mach 0.3, the present result is in good agreement with the results of incompressible flow studies. However, under supersonic flow, there are large differences; thus, it can be seen that there is the need to consider the Mach number when predicting the drag coefficient for high Mach number and low Reynolds number flows. Figures 20 and 21 show the comparison between the prediction values of the drag models [8, 15-18] and the present results. In Figure 20, the results of the experiment and calculation from previous studies are also shown; under subsonic flow, predicted values of drag models show good agreement with the results of this study. The model of Carlson and Hoglund [15] shows the best agreement of any model. In addition, the values predicted by the drag model that does not consider the Mach number effect also show good agreement with the results of this study. However, in the case of high Mach numbers, the difference of the drag coefficient becomes larger over the entire region. In Figure 21, under the supersonic flow condition, predicted values of the drag models do not show as good agreement as in the subsonic case. In particular, the difference between the predicted value of the drag models and present result becomes larger at low Reynolds numbers. In addition, there are differences in each drag model. Note that the Carlson and Hoglund model still shows good agreement in the supersonic case.

The model presented by Carlson and Hoglund does not consider the temperature ratio. However, it shows good agreement at all Mach numbers used in this analysis. The temperature ratio is the ratio between the temperatures of the particle and the free-stream. In this study, the calculation conditions are adiabatic at the sphere surface. Therefore, the temperature ratio is basically 1.0. However, at high Mach number flow, the sphere surface is heated by aerodynamic heating. Accordingly, in this study, particle temperature is decided by the average temperature of the sphere surface. Therefore, the cause of the large difference of the drag coefficient might be that the effects of the temperature distribution are not considered and there are problems with the calculation method. However, there is difference in the predicted value despite calculation under the same conditions. Therefore, it seems that the cause of the difference is something other than the temperature distribution. In addition, the difference becomes larger as Mach number increase. The target flow field has high Mach

number and low Reynolds number. Therefore, the measurement of the drag force is difficult. Hence, the drag model was built in a combination of the theoretical formula and indirect experimental data. Therefore, it seems that the prediction of the drag coefficient by the drag model is difficult under high Mach number and low Reynolds number conditions.

4 CONCLUSIONS

In this study, we performed the analyses of high Mach number and low Reynolds number flows. Consequently, we clarified the following topics: 1) As Mach number increases, unsteady fluctuation becomes weak and the separation point moves to the downstream side, 2) the drag coefficient becomes higher in the supersonic region due to the increase in the pressure drag by the detached shock-wave in the supersonic region, while the separation does not significantly affect it, and 3) the previous drag model shows good agreement with the present results in the subsonic region, while accurate prediction of the drag coefficient in the supersonic regime by traditional models might be difficult because the previous drag models are based on a model of incompressible flow and indirect experiment results.

REFERENCES

[1] Eldred, K. M. "Acoustic Loads Generated by the Propulsion System", NASA SP-8072, (1971).

[2] Ishii, T., Tsutsumi, S., Ui, K., Tokudome, S., Ishii Y., Wada, K. and Nakamura, S., "Acoustic Measurement of 1:42 Scale Booster and Launch Pad", Proceedings of Meetings on Acoustics, Vol. 18, No. 1, pp. 040009, (2014).

[3] Fukuda, K., Tsutsumi, S., Shimizu, T., Takaki, R., and Ui, K., "Examination of Sound Suppression by Water Injection at Lift-off of Launch Vehicles", 17th AIAA/CEAS Aeroacoustics Conference, (2011).

[4] Ignatius, J. K., Sathiyavageeswaran, S., and Chakravarthy, S. R., "Hot-Flow Simulation of Aeroacoustics and Suppression by Water Injection During Rocket Liftoff", AIAA Journal, Vol. 53, No. 1, pp. 235-245, (2014), DOI: 10.2514/1.J053078.

[5] Tsutsumi, S., Shimizu, T., Takaki, R., Shima, E., Arita, M. and Fujii, K, "Numerical Study of Pressure Waves Generated by H-IIA Launch Vehicle at Lift-off", AJCPP, pp. 266-271, (2008).

[6] Tsutsumi, S., Ishii, T., Ui, K., Tokudome, S., Wada, K., "Study on Acoustic Prediction and Reduction of Epsilon Launch Vehicle at Liftoff", Journal of Spacecraft and Rockets, pp. 1-12, (2014), DOI: 10.2514/1.A33010

[7] Nonomura, T., Morizawa, S., Obayashi, S. and Fujii, K., "Computational Prediction of Acoustic Waves from a Subscale Rocket Motor", Trans. JSASS Aerospace Thec. Japan, Vol. 12, No. 29, pp. Pe_11-Pe_17, (2014).

[8] Shimada, T., Daimon, Y. and Sekino, N. "Computational Fluid Dynamics of Multiphase Flows in Solid Rocket Motors", JAXA Special Publication, JAXA-SP-05-035E, (2006).

[9] Johnson, T. A and Patel, V. C., "Flow Past a Sphere Up to a Reynolds Number of 300",

Journal of Fluid Mechanics, Vol. 378, pp. 19-70, (1999).

[10] Nonomura, T., Terakado, D., Abe, Y. and Fujii, K., "A New Technique for Freestream Preservation of Finite-Difference WENO on Curvilinear Grid", Computers & Fluids, Vol. 107, pp. 242-255, (2015).

[11] Pirozzoli, S., "Stabilized Non-dissipative Approximations of Euler Equations in Generalized Curvilinear Coordinates", Journal of Computational Physics, Vol. 230, No. 8, pp. 2997-3014, (2011).

[12] Sugimoto, M., Giken-Zappo, No. 355(1944).

[13] Heberle, W. J., Wood, P. G., and Gooderum, B. P., N.A.C.A. Tech. Note 2100 (1950).

[14] Ambrosio, A. and Wortman, A., "Stagnation Point Shock Detachment Distance for Flow Around Spheres and Cylinders", ARSJ, Vol. 32, No. 281, pp. 875-875, (1962).

[15] Carlson, J. D. and Hoglung, F. R., "Particle drag and heat transfer in rocket nozzles", AIAA Journal, Vol. 2, No. 11, pp. 1980-1984, (1964).

[16] Henderson, C. B., "Drag Coefficients of Spheres in Continuum and Rarefied Flows", AIAA Journal, Vol. 14, No. 6, pp. 707-708, (1976).

[17] Crowe, T. C., Babcock, R. W., Willoughby, G. P. and Carlson, L. R., "Measurement of Particle Drag Coefficients in Flow Regimes Encountered by Particles in a Rocket Nozzle", UTC 2296-FR, United Technology Center, (1968).

[18] Hermsen, R. W., "Review of Particle Drag Models", JANAF Performance Standardization Subcommittee 12th Meeting Minutes, CPIA, Vol. 113, (1979).

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis (Eds)

Coupled Simulation of Shock Waves in Gas-Particle Mixtures Introducing Motion Equations

YUSUKE MIZUNO*, SHUN TAKAHASHI*, TAKU NONOMURA† , TAKAYUKI NAGATA* AND KOTA FUKUDA*

* Tokai Univesity, 4-1-1 Kitakaname, Hiratsuka, Kanagawa, JAPAN,
e-mail: 1beu2104@tokai-u.jp

* Tokai Univesity, 4-1-1 Kitakaname, Hiratsuka, Kanagawa, JAPAN,
e-mail: takahasi@tokai-u.jp

† ISAS/JAXA, 3-1-1 Yoshinodai, Chuo-ku, Sagamihara, Kanagawa, JAPAN,
e-mail: nonomura@flab.isas.jaxa.jp

* Tokai Univesity, 4-1-1 Kitakaname, Hiratsuka, Kanagawa, JAPAN,
e-mail: 1beu2216@tokai-u.jp

* Tokai University, 4-1-1 Kitakaname, Hiratsuka, Kanagawa, JAPAN,
e-mail: fukuda@tokai-u.jp

Key words: Three-dimensional compressible Navier-Stokes equation, direct numerical simulation, moving particles, shock wave

Abstract. In this work, direct numerical analyses for flow around particles passing a shock wave was carried out to predict effects of small particles in rocket plumes. A flow solver based on three-dimensional compressible Navier-Stokes equations is developed for the purpose of high accurate prediction of the acoustic field around rocket plumes. This flow solver is capable of analysing a flow around moving multiple particles and motion equations was introduced. The flow field and the drag coefficient after the shock wave passage were validated by comparing with the drag models at shock Mach number 1.2-2.8. The result was in good agreement with the drag models. In the flow around multiple particles, the interference between particles was confirmed.

1 INTRODUCTION

The acoustic waves from rocket plumes are sufficiently strong to damage the satellites inside the fairing of a rocket. These waves are assessed by empirical prediction methods in Europe, the United States and Japan [1]. These methods were based on the enormous launching results and test data in the United States. But, the low accuracy of these methods renders them unsuitable for a new rocket and a new launch site and more accuracy prediction of the acoustic environment and the quantitative evaluation of acoustic waves is required. Therefore various studies by the numerical analysis have been conducted. The acoustic prediction was enabled in various assumptions qualitatively [2-4]. In real large scale liquid launch vehicles, the generated acoustic waves are suppressed by water injection. For the high accuracy prediction of acoustic waves, examination of detail of the effect of water droplets and the physical phenomenon is required and it is not considered at present numerical simulation.

Fukuda et al. showed that acoustic waves are primarily attenuated by interactions between particles and turbulence [5] from theoritical analysis. These particles which influence turbulence are alumina particles exhausting from the rocket nozzle or water droplets surrounding the rocket. However, the mechanism of the phenomenon is not well understood.

Therefore, in this study, the direct numerical analysis is carried out in order to understand the behavior of gas-particle mixtures in high Mach number turbulent flowfield. The purpose of this study is to analyze the flowfield in which multiple particles pass a shock wave in order to understand the interference of those particles and turbulence in the rocket plumes. The flow of the particle passing shock wave has been studied as the multiphase flow or the shock tube problem [6-8]. However, those studies are not numerical analysis for the three-dimensional flowfield and understanding of the flow is not enough. This analysis employed in this study is direct numerical analysis for the flow in which multiple purticles pass a shock wave using the immersed boundary method.

2.1 Governing Equations

In the present study, flows are governed by three-dimensional compressible Navier-Stokes equations.

$$\frac{\partial Q}{\partial t}+\frac{\partial E}{\partial x}+\frac{\partial F}{\partial y}+\frac{\partial G}{\partial z}=\frac{\partial E_v}{\partial x}+\frac{\partial F_v}{\partial y}+\frac{\partial G_v}{\partial z}$$

$$Q=\begin{bmatrix}\rho\\ \rho u\\ \rho v\\ \rho w\\ \rho e\end{bmatrix},E=\begin{bmatrix}\rho u\\ \rho u^2+p\\ \rho uv\\ \rho uw\\ (\rho e+p)u\end{bmatrix},F=\begin{bmatrix}\rho v\\ \rho vu\\ \rho v^2+p\\ \rho vw\\ (\rho e+p)v\end{bmatrix},G\begin{bmatrix}\rho w\\ \rho wu\\ \rho wv\\ \rho w^2+p\\ (\rho e+p)w\end{bmatrix}, \tag{1}$$

$$E_v=\begin{bmatrix}0\\ \tau_{xx}\\ \tau_{xy}\\ \tau_{xz}\\ \tau_{xx}u+\tau_{xy}v+\tau_{xz}w+\kappa T_x\end{bmatrix},F_v=\begin{bmatrix}0\\ \tau_{yx}\\ \tau_{yy}\\ \tau_{yz}\\ \tau_{yx}u+\tau_{yy}v+\tau_{yz}w+\kappa T_y\end{bmatrix},G_v=\begin{bmatrix}0\\ \tau_{zx}\\ \tau_{zy}\\ \tau_{zz}\\ \tau_{zx}u+\tau_{zy}v+\tau_{zz}w+\kappa T_z\end{bmatrix}$$

where E, F, and G are the inviscid fluxes in the x-, y-, and z- directions, respectively E_v, F_v, and G_v are the corresponding viscous fluxes, and Q contains the conservative variables. The pressure p is related to the total energy e per unit mass by the equation of state:

$$\rho e=\frac{p}{\gamma-1}+\frac{1}{2}\rho(u^2+v^2+w^2) \tag{2}$$

Here the stress tensor components are given as

$$\begin{aligned}
\tau_{xx} &= \frac{2}{3}\mu(2u_x - v_y - w_z) \\
\tau_{yy} &= \frac{2}{3}\mu(2v_y - w_z - u_x) \\
\tau_{zz} &= \frac{2}{3}\mu(2w_z - u_x - v_y) \\
\tau_{xy} &= \tau_{yx} = \mu(u_y + v_x) \\
\tau_{yz} &= \tau_{zy} = \mu(v_z + w_y) \\
\tau_{zx} &= \tau_{xz} = \mu(w_x + u_z)
\end{aligned} \tag{3}$$

All variables are nondimensionalized by the freestream conditions of density, sound speed, and unit length. The above equations are discretized on an equally spaced Cartesian mesh with a cell-centered arrangement. To eliminate additional numerical dissipation everywhere, except in the vicinities of shock waves and potential flows, the inviscid terms are computed by a hybrid scheme that combines the monotone upstream-centered scheme for conservation laws (MUSCL)-Roe scheme [9] and the second-order pseudo skew-symmetric scheme [10]. This cord has been validated in a two-dimensional flowfield [11].

2.2 Boundary Representation

The boundary is defined by the level set method and ghost-cell method. The level set function is determined in whole cells as assigned distance from the object boundary. In the present study, flows around multiple moving objects are solved by extending the level set method to multiple level set functions based on simple minimum distance approach. All cells are classified into three categories: fluid cell, ghost cell, and object cell. The ghost cells behave as guard cells between the fluid and object regions and are assigned in two layers under the present definition as follows:

$$\begin{aligned}
&d_{PC} > 0 \\
&d_{GC} \le 0 \, and \, d_{GC} \ge \sqrt{3}\Delta x \\
&d_{OC} < \sqrt{3}\Delta x
\end{aligned} \tag{4}$$

The present ghost-cell method is adopted the image point approach. Ghost cells are used for imposing boundary condition. An image point set in the region of fluid cells is used to collect flow information for a ghost cell. The image point is located at the edge of a probe that extends from a ghost cell through the object boundary in the direction normal to the surface. The probe length is 1.75 times the mesh size. The probe is set to be longer then √3 times the mesh size to avoid a recursive reference. The nodes enclosed are classified only as fluid cells. The primitive variables on the image point are interpolated by the trilinear function based on the surrounding cells. Finally, the value of ghost cell is defined by the value of the image point.

3 COMPUTATIONAL RESULT OF FLOW AROUND A PARTICLE PASSED A SHOCK

3.1 Computational Condition

Figure 1 shows a computational domain. In this study, the shock wave is fixed in the coordinate system. A particle passes the shock from the low pressure side to the high pressure side. The Reynolds number, based on the particle diameter and freestream values (including viscosity), is fixed at 300, while the shock Mach number (freestream, initial particle velocity) are varied as Mach 1.2, 1.5, 2.0, 2.5, and 2.8. Mesh size is 0.1D, where the diameter D of the particle. The calculation domain is 20D at x-direction and 10D at y- and z- direction. The initial temperature of the particle surface is the same as the temperature of the shock wave front flow. After passing through it varies depending on the temperature of the shock wave wake. Dirichlet conditions are imposed on all flow variables at the inflow boundary and on density alone at the outflow boundary. Neumann conditions are imposed at the top and bottom boundaries for all variables. Table 1 shows the initial condition normalized by speed of sound and characterstic length.

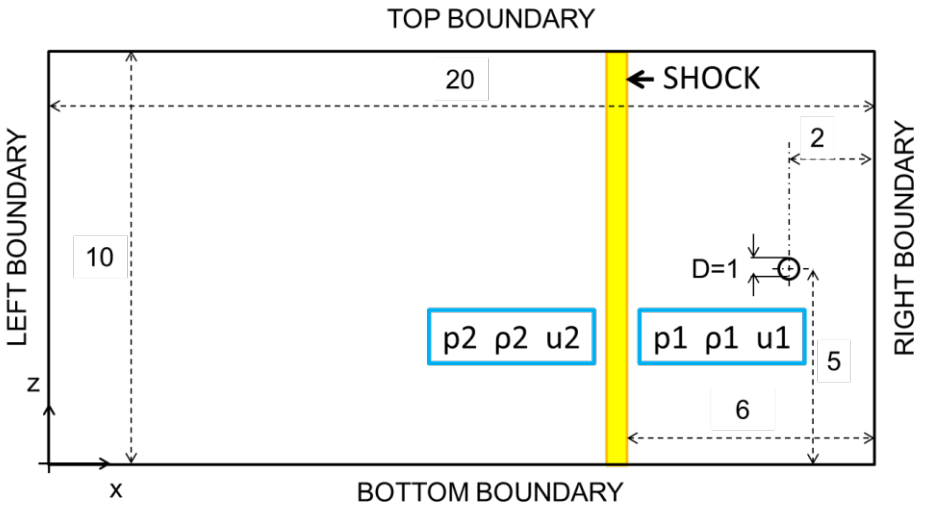


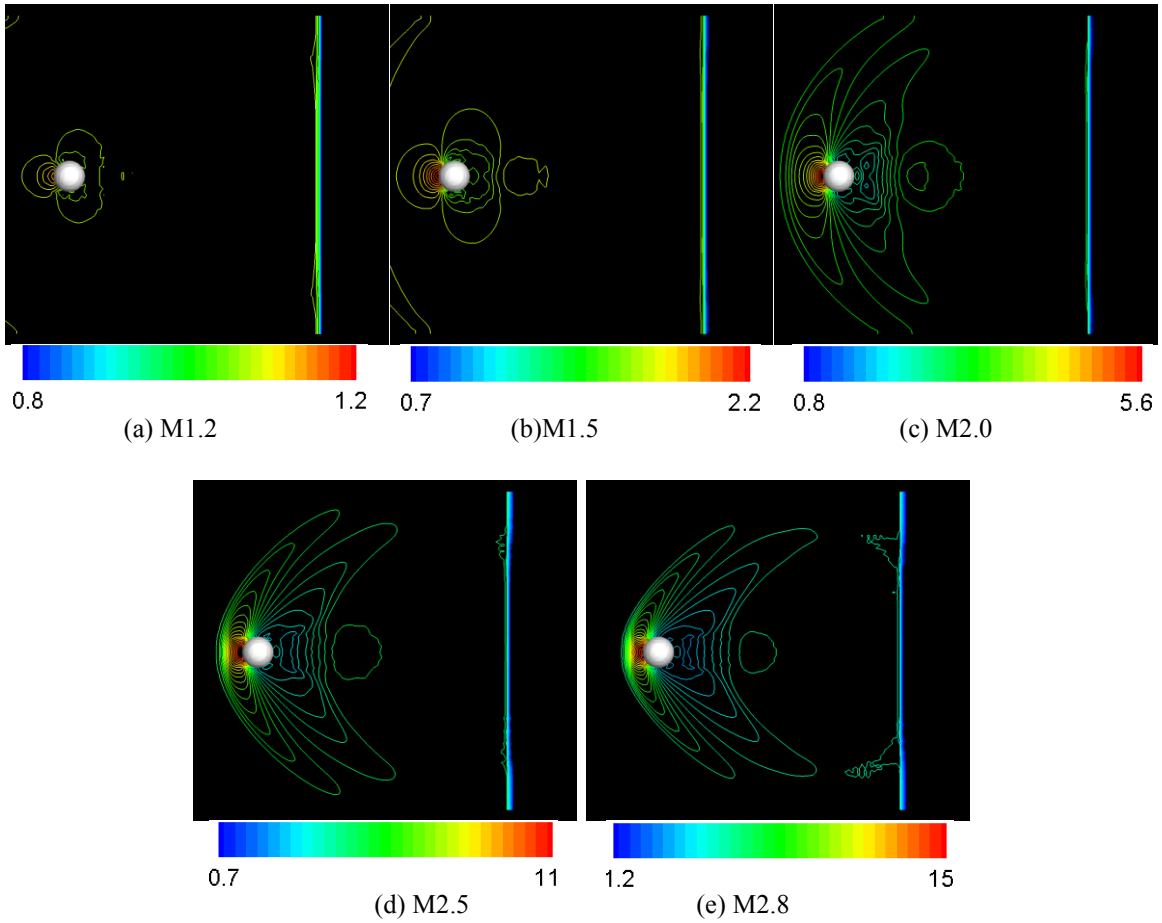
Figure 1 : Computational domain

Table 1 : Initial conditions

Reynolds	Shock Mach Number	p1	ρ1	u1	p2/p1	ρ2/ρ1	u2/u1
300	1.2	0.71	1.0	1.2	1.51	1.34	0.75
	1.5			1.5	2.46	1.86	0.54
	2.0			2.0	4.50	2.67	0.38
	2.5			2.5	7.13	3.33	0.30
	2.8			2.8	8.98	3.66	0.27

3.1 Computational result

Figure 2 plots the pressure distributions at $x = 6$. The shock wave is not formed in front of a particle at M1.2 and M1.5, but it is formed at M2.0, M2.5, and M2.8 after the particle passes the shock wave. For this reason, the particle velocity exceeds speed of sound between M1.5 – M2.0 in the domain of the shock wave slipstream.



(a) M1.2 (b)M1.5 (c) M2.0

(d) M2.5 (e) M2.8

Figure 2 : Pressure distributions ($x = 6$)

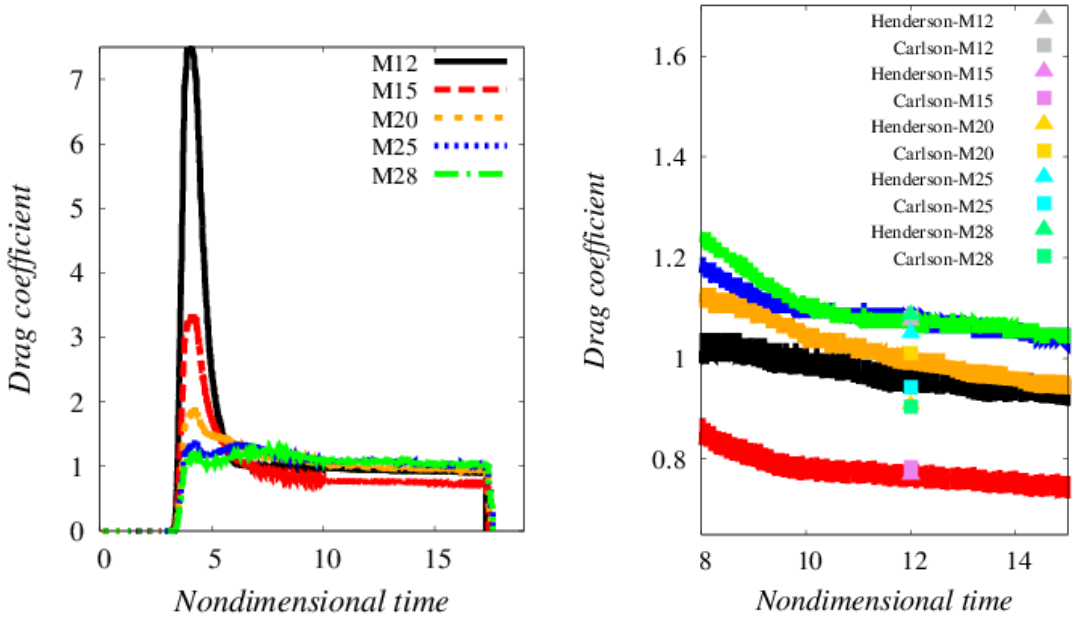
Figure 3 shows drag coefficient. When a particle passes a shock wave at $t = 3$, the value of the drag coefficient increases sharply. The maximum value becomes large as the low Mach number. After the passage, the value decreases. The drag coefficient indicates zero because the particle goes out of the computational domain at t = 18.

To compare with the drag model, the value of the drag coefficient of Henderson's model [12] and Carlson and Hoglund's model [13] that calculated with the number of practice. Reynolds number at $t = 12$. Henderson has relatively close at M1.5 – M2.8. On the othe hand, Carlson has the same quantity at M1.5 and M2.0. Table 2 shows error (%) of the drag coefficient. Both models are the same value, however, the present result is different value at M1.2. Since this error occurred when the flow around a stationary particle is analyzed, this analysis code has an error in this case. Henderson' calculation formula contains the

temperature ratio effect of fluid and particle. On the other hand, Carlson and Hoglund's calculation formula does not consider this temperature ratio. The influence of this temperature ratio is samll, but this influence increases as the Reynolds number increases. Therefore the present result is significantly different from Carlson and Hoglund in the high Mach number cases.

Figure 4 plots the particle velocity. It is normalized in the particle velocity. The particle is decelerated after the shock wave passes. As the case in which particle velocity increases, the particle velocity suddenly decelerated. This is due to the wave resistance.

Figure 5 shows the turbulence kinetic energy that is nondimensionalized by the sound speed behind the shock wave. Since the shock wave is $x = 14$, the value is rapidly changed. Furthermore the particle is $x = 6.0$, the value increases the wake of the particle. The increment becomes large as the Mach number increases.



(a) Analysis results (b) Prediction value by drag models

Figure 3 : Drag coefficient

Table 2 : Drag coefficient error(%)

M_{shock}	Henderson	Carlson and Hoglund
1.2	11	12
1.5	0.7	2.6
2.0	9.5	1.7
2.5	2.4	14
2.8	1.5	18

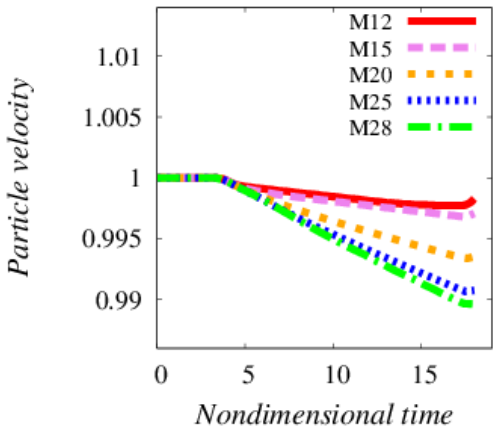


Figure 4 : Particle velocity

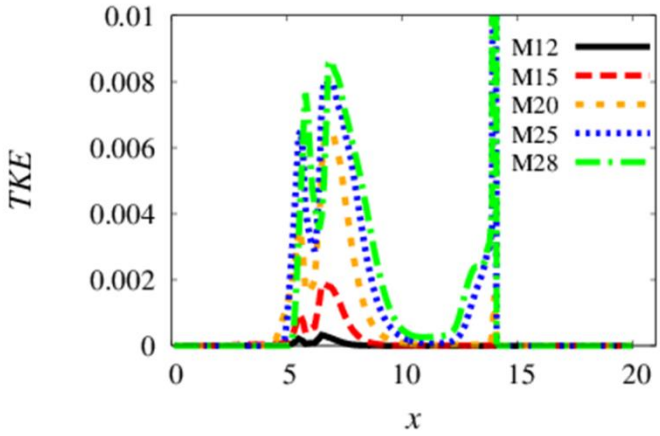


Figure 5 : Nondimensional turbulence kinetic energy

4 COMPUTATIONAL RESULT OF FLOW AROUND PARTICLES PASSED A SHOCK

4.1 Computational Condition

Figure 6 shows a computational domain. Three particles (objects number : o1, o2, o3) pass a shock from the low pressure side to the high pressure side. The Reynolds number, based on the particle diameter and freestream values (including viscosity), is fixed at 300, while the shock Mach number (freestream, initial particle velocity) are varied as Mach 2.5. Mesh size is 0.1*D*The calculation domain is 40*D* at *x*-direction and 10D at *y*- and *z*- direction. Other conditions are the same as the a particle at M2.5.

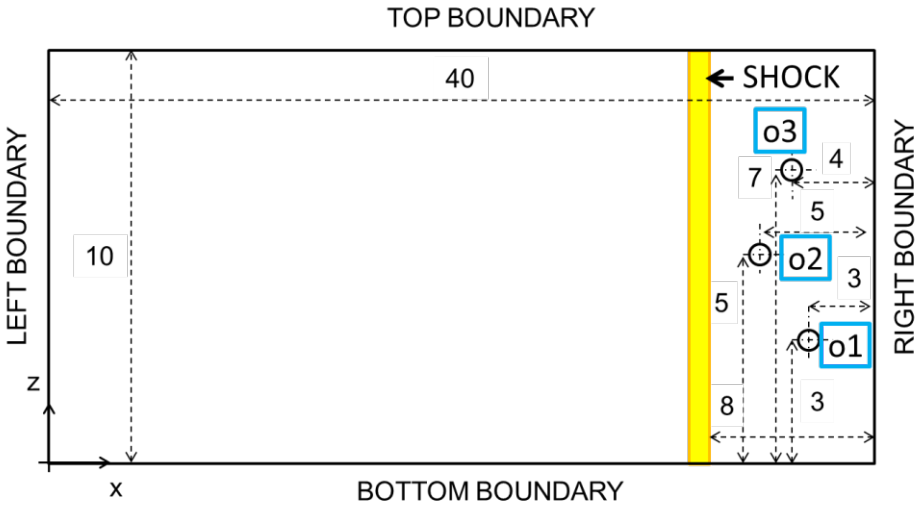
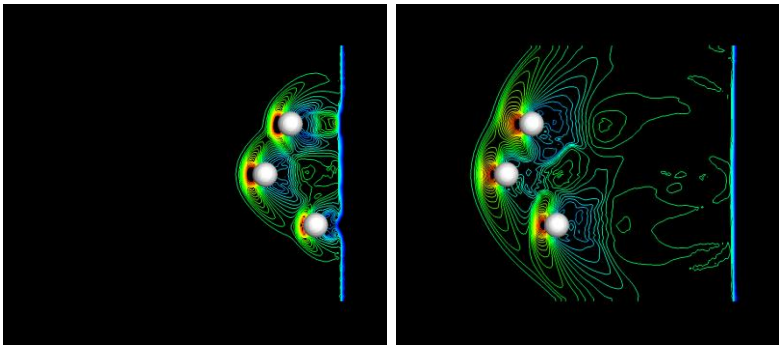


Figure 6 : Computational domain

4.1 Computational result

Figure 7 plots the pressure distribution at nondimensional time $t = 6.2$ and $t = 12$. Those particles are passing thorough the shock wave at $t = 6.0$. Those detached shock waves interfere with each other at $t = 12$. Figure 8 shows drag coefficient. *o1* has little influence from other particles. As a result, the tendency of the drag coefficient is comformed to the result of a single particle. *o2* has the influence from neighboring particles and the drag coefficient reduced. The value of *o3* increases by the influence of *o1*. Figure 9 plots those particles velocity at each point. The transition of the particle velocity is different with each particle because of interferes. Figure 10 displays the turbulence kinetic energy distribution that is nondimensionalized by the sound speed behind the shock wave. The particles are $x = 5.0$, 6.0 and 7.0, the turbulence kinetic energy increases the slipstream of particles and expands over a wide range. Moreover the maximum value becomes almost two times.



(a) t = 6.0 (b) t = 12

Figure 7 : Pressure coefficient distributions

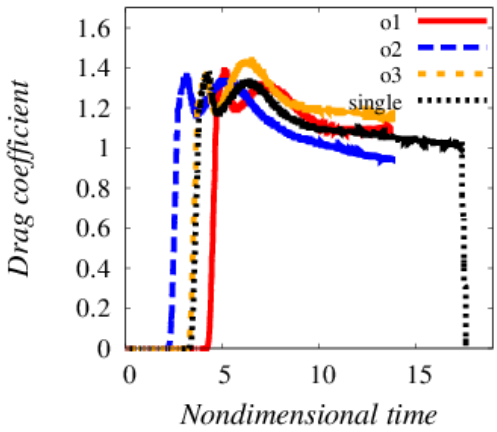


Figure 8 : Drag coefficient

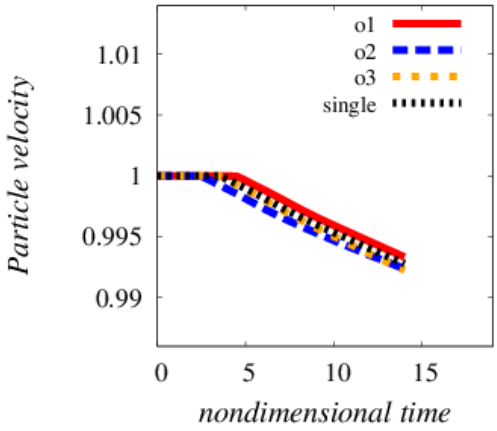


Figure 9 : Particles velocity

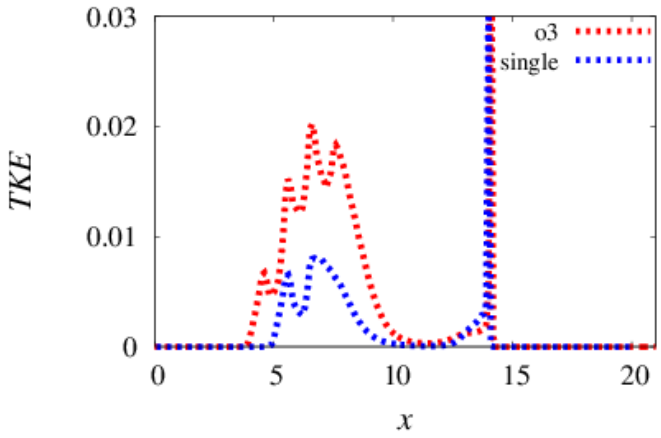


Figure 10 : Nondimensional turbulence kinetic energy

5 CONCLUSIONS

Direct numerical simulation on particles passing a shock wave was carried out. The flow condition is the shock Mach number at 1.2, 1.5, 2.0, 2.5, and 2.8 at Reynolds numeber 300. As a result, the drag coefficient has the error about 10 %. However, this error is due to the drag coefficient does not have the temperature ratio effect of the flowfield and the particle. After the particle passes the shock wave, the particle velocity decereases. The low turbulence kinetic energy hasthe wake of the particle. In the flow including multiple particles, the fluction of the drag coefficient is influenced by the interference between the particle. In near future, we will conduct large-scale gas-particle flow simulation and focus on the turbulence energy generation or the interference between the particle.

REFERENCES

[1] K. M. Eldred, "Acoustic Loads Generated by Propulsion System", NASA SP-8072, 1971

[2] T. Nonomura, K. Fujii, and S. Tsutsumi, "Toward accurate simulation and analysis of strong acoustic wave phenomenon – A review from the experience of our study on rocket problems", International Journal for Numerical Methods in Fluids, pp. 1412-1432, 2010

[3] K. Fukuda, S. Tsutsumi, K. Fuji, K. Ui, T. Ishii, H. Oinuma, J. Kazawa and K. Minesugi, "Acoustic Measurement and Prediction of Solid Rockets in Static Firing Tests", The 15th AIAA//CEAS Aeroacoustics Conference, AIAA paper pp. 2009-3368, 2011.

[4] S. Tsutsumi, T. Ishii, K. Ui, S. Tokudome and K. Wada, "Study on Acoustic Prediction and Reduction of Epsilon Launch Vehicle at Liftoff", Journal of Spacecraft and Rockets, 2014

[5] K. Fukuda, S. Tsutsumi, T. Shimizu, R. Takaki, and K. Ui, "Examination of Sound Suppression by Water Injection at Lift-off of Launch Vehicles", The 17th AIAA/CEAS Aeroacoustics Conference, AIAA paper pp. 2011-2814, 2011.

[6] T. Saito, M. Marumoto, K. Takayama, "Numerical investigations of shock waves in gas-particle mixtures", Shock Waves, pp. 299-322, 2003

[7] H. Tanno, K. Itoh, T. Saito, A. Abe, K. Takayama, "Interaction of a shock with a sphere suspended in a vertical shock tube", Shock Waves, pp. 191-200, 2003

[8] M. Sun, T. Saito, K. Takayama, H. Tanno, "Unsteady drag on a sphere by shock wave loading", Shock Waves, 2004

[9] B. van Leer, "Towards the Ultimate Conservative Difference Scheme. IV. A New Approach to Numerical Convection.", Journal of Computational Physics, Vol. 43, No. 2, pp. 357-372, 1981

[10] K. Georgesa, G. Winckelmans and P. Geuzainea, "Improving shock-free compressible RANS solver for LES on unstructured meshes" Journal of Computational and Applied Mathematics, Vol. 215, pp. 419-428, 2008

[11] S. Takahashi, T. Nonomura and K. Fukuda, "A Numerical Scheme Based on an Immersed Boundary Method for compressible Turbulent Flows with Shock: Application Two-dimensional Flows around Cylinders", Journal of Applied Mathematics, Vol. 2014

[12] C. B. Henderson, "Drag Coefficients of Spheres in Continuum and Rarefied Flows", AIAA Journal, Vol. 14, No. 6, 1976

[13] D. J. Carson and R. F. Hoglund, “Particle Drag and Heat Transfer in Rocket Nozzles”, AIAA Journal, Vol. 2, No. 11, 1980

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis (Eds)

DO TRANSPORT REGIMES ACROSS THE WATER-SEDIMENT INTERFACE UNDER LOW-ENERGETIC OSCILLATORY FLOWS

E. SÁNCHEZ-BADORREY[*] AND F. J. GARCÍA-ANGUITA[†]

[*] Dpt. M. E. e Ingeniería Hidráulica – Instituto del Agua
Universidad de Granada
c/Ramón y Cajal, 4. Granada, 18071, Spain
e-mail: elenasb@ugr.es

[†] Dpt. M. E. e Ingeniería Hidráulica – Instituto del Agua
Universidad de Granada
c/Ramón y Cajal, 4. Granada, 18071, Spain
e-mail: jgarcianguita@ugr.es

Key words: Water Sediment Interface, Transport Regimes, Dissolved Oxygen, Anoxia, Diffusive Boundary Layer, Multiple Scale Analysis.

Abstract. Dissolved oxygen (DO) transfer between bottom sediments and the overlaying water (Water-Sediment Interface, WSI) constitutes an essential coupling for benthic and pelagic ecosystems and influences a wide range of important geochemical processes. Very close to the bed, where turbulent diffusivity becomes negligible, DO fluxes are governed by molecular diffusion and the concentration gradients across the diffusive boundary layer (DBL) and diffusive sediment layer (DSL). The influence of the DBL dynamics induced by low-energetic oscillatory flows on the coupled transport of DO across the WSI is analyzed here by means of a 1D unsteady diffusive-reactive mass transport model in an evolving domain. A coordinate transformation is applied to map the evolving domain onto a fixed region and to solve the equivalent coupled transport problem. Using a multiple scale analysis, the DO transport regimes across de WSI are investigated as function of the DBL thickness variability. Model results show that the fluctuations of the DBL thickness induced by oscillatory flows lead to additional advective transport terms that enhance the DO transport across the WSI, playing an important role on the DO penetration depth and the occurrence of anoxia events in the sediment domain.

1 INTRODUCTION

Dissolved oxygen (DO) transfer between bottom sediments and the overlaying water (Water-Sediment Interface, WSI) constitutes an essential coupling for benthic and pelagic ecosystems and influences a wide range of important geochemical processes. For example, it regulates the organic matter degradation in the upper sediment.

Very close to the bed, where turbulent diffusivity becomes negligible, DO fluxes are governed by molecular diffusion and the concentration gradients across the diffusive boundary layer (DBL) and diffusive sediment layer (DSL). These regions, in which gradients

in physical, chemical and biological properties are sharp, control the vertical mass transfer at the benthic boundary, since molecular diffusion is generally orders of magnitude slower than turbulent and advective transports. In these conditions, the thickness of the DBL plays an important role, as it depends on the bulk flow conditions. Moreover, recent studies show that the thickness of the DBL response to the variability of the flow conditions even at time scales of the order of seconds or minutes [1, 2].

The thickness of the DBL and the DSL are time dependent. Here, they are denoted, respectively, as $\delta_D = \delta_{DBL}(t)$ and $\delta_S = \delta_{SBL}(t)$. δ_S is defined as the penetration depth in which the solute concentration is 5% of the solute concentration in the water column. Deep enough in the sediment, $h >> \delta_S$, the DO concentration can be considered negligible and constant. The time dependence of δ_D is given by the water column hydrodynamics and it is assumed known. It can be calculated as function of the friction velocity $u*$ or the turbulent dissipation rate, ε, depending of the bulk flow regime.

In this work, we analyze the influence of the temporal variability of the DBL thickness under low-energetic oscillatory flows on the coupled transport of DO across the WSI. Applying a multiple scale analysis to the 1D unsteady diffusive-reactive mass transport problem in an evolving domain, we analyze the possible DO transport regimes across the WSI as function of the characteristic time scales of: (1) the oscillatory flow, (2) the diffusive transport processes, and (3) the kinetics of the oxygen consumption processes involved.

2 THE DIFFUSIVE BOUNDARY LAYER MODEL

2.1 Governing equations

DO transport across the WSI located at $z = 0$ is described as a time dependent one-dimensional problem in an evolving domain. **Figure 1a** shows the sketch of the WSI problem. The location of the WSI is assumed to be constant in time. In the DBL region, $0 \le z \le \delta_D(t)$, the DO transport is diffusive. While in the SBL, $z < 0$, the DO transport is governed both by diffusive transport and consumption kinetics. The DO consumption has been modeled here assuming a first order Michaelis-Menten kinetics. Thus, the governing equation of the WSI problem can be written as,

$$\frac{\partial C_i}{\partial t} = \frac{\partial}{\partial z}\left(D_i \frac{\partial C_i}{\partial z} \right) - \alpha_i \nu \frac{C_i}{K + C_i} \quad i = 1,2 \tag{1}$$

where $i = 1$ denotes the water diffusive region (DBL) with $\alpha_1 = 0$, and $i = 2$ denotes the sediment diffusive layer (DSL) with $\alpha_2 = 1$. C_i represents the DO concentration in each region. The diffusivities (D_i) of the two regions are assumed constant within each region but discontinuous across the WSI. ν and K are constant kinetic parameters.

2.2. Coordinate transformation and problem mapping

The evolving domain of the WSI problem is transformed by defining the new variable

$$\xi = \frac{z}{\delta_D(t)} \tag{2}$$

In terms of ξ , the DBL problem given by Eq. (1) can be written as

$$\frac{\partial C'}{\partial t} = \frac{1}{T_w}\xi\frac{\partial C'}{\partial \xi} + \frac{1}{T_D}\frac{\partial^2 C'}{\partial \xi^2} \quad 0 \le \xi \le 1 \tag{3}$$

where T_w and T_D are, respectively, the characteristic time scales of the DBL thickness dynamics. They are be defined as,

$$\frac{1}{T_w} = \frac{1}{\delta_D}\frac{d\delta_D}{dt} \tag{4a}$$

$$\frac{1}{T_D} = \frac{D_w}{\delta_D^2} \tag{4b}$$

Notice that T_w and T_D depend on: (1) the characteristic diffusivity of the DBL in the water region ($D_1 = D_w$), and (2) on the bulk flow influence on the DBL thickness ($\delta_D = \delta_{DBL}(t)$).

Similarly, for the diffusive-reactive transport problem in the sediment region (DSL), the coordinate system can be transformed using

$$\varphi = \frac{z}{\delta_S(t)} \tag{5}$$

Being the characteristic time scales of the transport across this region,

$$\frac{1}{T_{w,s}} = \frac{1}{\delta_S}\frac{d\delta_S}{dt} \tag{6a}$$

$$\frac{1}{T_{D,s}} = \frac{D_s}{\delta_S^2} \tag{6b}$$

and $D_2 = D_s$. **Figure 1b** shows the sketch of the mapped domain after the coordinate transformation.

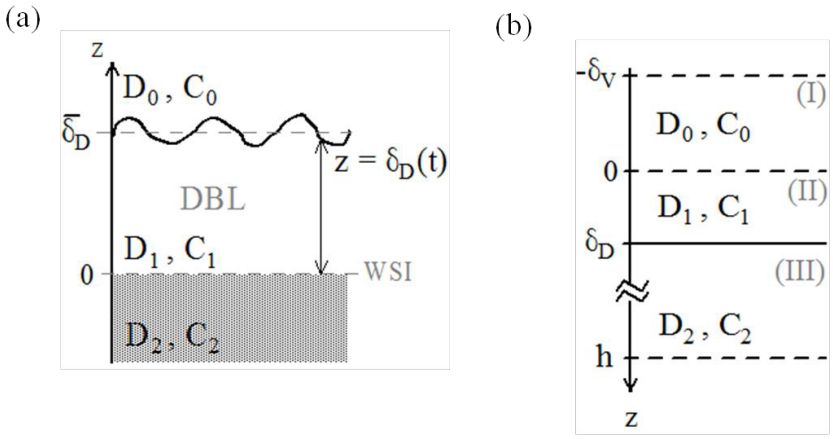


Figure 1. (a) Time dependent domain of the WSI under low-energetic oscillatory flows. (b) Mapped problem after coordinate transformation.

Let us define $t' = t / T_0$, where T_0 represents a selected characteristic time scale of the mapped problem. And, to express the time dependent variables as

$$\begin{aligned} T_w(t) &= \overline{T}_w + \Delta T_w(t) \\ T_D(t) &= \overline{T}_D + \Delta T_D(t) \\ \delta_D(t) &= \overline{\delta}_D + \Delta \delta_D(t) \end{aligned} \tag{7}$$

where the overbar denotes the mean (time-averaged) of the time dependent variable and Δ represents the instantaneous deviation from its mean value, e.g. $\Delta\delta_D = \delta_D(t) - \overline{\delta}_D$.

Eq. (3) can be now rewritten as

$$\frac{\partial C'}{\partial t'} = \frac{T_0}{\overline{T}_w}\left(\frac{1}{1+\frac{\Delta T_w}{\overline{T}_w}}\right)\xi\frac{\partial C'}{\partial \xi} + \frac{T_0}{\overline{T}_D}\left(\frac{1}{1+\frac{\Delta T_D}{\overline{T}_D}}\right)\frac{\partial^2 C'}{\partial \xi^2} \quad 0 \le \xi \le 1 \tag{8}$$

Eq. (8) makes apparent that the fluctuations of the DBL thickness, induced by the low-energetic oscillatory bulk flow, lead to an additional advective transport term in the mapped problem of the DBL.

Similarly, the characteristic time scales of the DSL can be expressed as

$$\begin{aligned} T_{w,s}(t) &= \overline{T}_{w,s} + \Delta T_{w,s}(t) \\ T_{D,s}(t) &= \overline{T}_{D,s} + \Delta T_{D,s}(t) \\ \tau(t) &= \overline{\tau} + \Delta\tau(t) \\ \delta_S(t) &= \overline{\delta}_S + \Delta\delta_S(t) \end{aligned} \tag{9}$$

where $\tau(t)$ is the characteristic time scale of the DO kinetics. The governing equation of diffusive-reactive transport in the DSL can be accordingly written as,

$$\frac{\partial C'}{\partial t'} = \frac{T_0}{\overline{T}_{w,s}}\left(\frac{1}{1+\frac{\Delta T_{w,s}}{\overline{T}_{w,s}}}\right)\varphi\frac{\partial C'}{\partial \varphi} + \frac{T_0}{\overline{T}_{D,s}}\left(\frac{1}{1+\frac{\Delta T_{D,s}}{\overline{T}_{D,s}}}\right)\frac{\partial^2 C'}{\partial \varphi^2} - \frac{T_0}{\overline{\tau}}\left(\frac{1}{1+\frac{\Delta \tau}{\overline{\tau}}}\right)\mu' F(C') \quad -1 \le \varphi \le 0 \tag{10}$$

In highly organic sediments, the DO consumption kinetics is typically faster than the diffusive transport, hence $\overline{\tau} << T_{D,s}$. Moreover, as $\delta_S(t) \approx \overline{\delta}_S = h$, the advective transport term in Eq.(9) can be considered negligible. Accordingly,

$$\frac{\partial C'}{\partial t'} = \frac{T_0}{\overline{T}_{D,s}}\frac{\partial^2 C'}{\partial \varphi^2} - \frac{T_0}{\overline{\tau}}\mu' F(C') \quad -1 \le \varphi \le 0 \tag{11}$$

It is concluded that the DSL transport problem in highly reactive sediments in the mapped domain can be described as a diffusive-reactive transport process.

3 MULTIPLE SCALE ANALYSIS AND WSI TRANSPORT REGIMES

The DO transport regimes induced by the variability of the DBL thickness under oscillatory bulk flows are analyzed next as function of the characteristic time scales of the physical processes involved.

First, the multiple scale analysis of the DBL problem is presented. For that, T_0 is defined as the characteristic time scale of the oscillatory bulk flow. Next, the transport regimes of the de DBL problem are established. Finally, the multiple scale analysis and transport regimes of the DSL are discussed. In this document we will discuss the particular case of highly organic sediments described by Eq. (11).

To simplify the notation, hereafter, we skip the overbar (mean value) and the comma (adimensional variable) of the formulation.

3.1. Multiple scale analysis of the DBL problem

Let us define the parameter $\varepsilon = \Delta\delta_D / \overline{\delta}_D$ as a measure of the relative magnitude of DBL thickness variability with respect to its mean value, and τ_D as the characteristic time scale of DBL thickness dynamics. Making use of Eq. (4a), it is found that

$$\frac{1}{T_w} \simeq \frac{1}{\tau_D}\left(\frac{\varepsilon}{1+\varepsilon}\right) \tag{12}$$

and

$$\frac{1}{\overline{\delta}_D}\frac{d\Delta\delta_D}{dt} \simeq O\left(\frac{\varepsilon}{\tau_D}\right) \tag{13}$$

Similarly,

$$\frac{1}{T_D} = \frac{1}{\overline{T}_D}\left(\frac{1}{1+\varepsilon}\right)^2 \tag{14}$$

where $\overline{T}_D = D_w / \overline{\delta}_D{}^2$ is the mean time scale of the diffusive transport across the WSI. Making use of this scale analysis, Eq. (8) can be written as

$$\frac{\partial C'}{\partial t'} = \frac{T_0}{\tau_D}\left(\frac{\varepsilon}{1+\varepsilon}\right)\xi\frac{\partial C'}{\partial \xi} + \frac{T_0}{\overline{T}_D}\left(\frac{1}{1+\varepsilon}\right)^2\frac{\partial^2 C'}{\partial \xi^2} \quad 0 \le \xi \le 1 \tag{15}$$

This equation shows that the fluctuations of the DBL thickness induced by low-energetic oscillatory flows lead to an additional advective transport term in the DBL mapped domain. The magnitude of the advective term varies linearly across the DBL and it is proportional to $\varepsilon/(1+\varepsilon)$, being $\varepsilon = \Delta\delta_D / \overline{\delta}_D$ the relative magnitude of DBL thickness variability with respect to its mean value. It is worthy to point out that the diffusive transport term in Eq. (15) has a constant diffusivity coefficient across the DBL and it is proportional to $1/(1+\varepsilon)^2$.

3.2. Transport regimes of the DBL problem

Three possible transport regimes across the DBL can be distinguished as function of the relative magnitude of the advective and diffusive transport terms in Eq. (15): (1) DDT: Diffusive Dominant Transport regime; (2) ADT: Advective Dominant Transport regime; and (3) MT: Mixed transport regime.

The *Diffusive Dominant Transport regime* (DDT) is defined here as those in which the characteristic time scales of the DBL problem fulfill that $\tau_D >> T_0 >> \overline{T}_D$. In this regime the DBL transport is governed by diffusion, with a decreasing diffusivity coefficient with the increasing amplitude of the DBL variability ($\varepsilon = \Delta\delta_D / \overline{\delta}_D$).

In the *Advective Dominant Transport regime* (ADT), the time scales of the problem satisfy the inequality: $\overline{T}_D >> T_0 >> \tau_D$. In this case, the advective contribution induced by the DBL thickness variability dominates the DO transport across the mapped DBL problem. The magnitude of the advective transport shows an asymptotic behavior as the parameter $\varepsilon = \Delta\delta_D / \overline{\delta}_D$ increases.

Finally, it is possible to define a *Mixed Transport regime* (MT) in which the advective and diffusive contributions have the same order of magnitude, i.e. $\tau_D \approx \overline{T}_D$. In this regime, both terms (diffusive and advective) in Eq. (15) contribute significantly to the DO transport across the DBL.

Figure 2 shows the relative amplitude of the diffusive and advective transport processes across the DBL as function of the parameter $\varepsilon = \left|\Delta\delta_D / \bar{\delta}_D\right|$. The mixed transport regime (MT) in which the transition from the diffusive dominant regime (DDT) to the advective dominant regime (ADT) was found equal to threshold value $\varepsilon \approx 0.618$, that is when $\left|\Delta\delta_D\right| \geq 0.618\bar{\delta}_D$. The amplitude of the DBL variability is found to be one of the key parameter to control the relative influence of the different transport mechanisms in the DBL mapped problem. More precisely, the mean value ($\bar{\delta}_D$), the amplitude ($\Delta\delta_D$) and the characteristic time scale (T_D) of the fluctuations of the DBL thickness are the key parameters to identify the dominant DO transport regime across the DBL under low energetic oscillatory flows.

These findings are consisting with previous numerical results by [2] for slow oscillatory motions of monochromatic type.

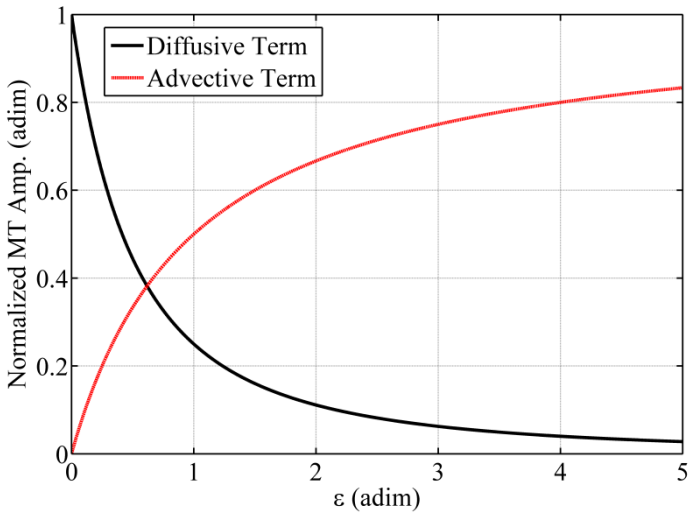


Figure 2. Dependence of diffusive and advective transport processes of the mapped DBL problem, Eq.(15), as function of $\varepsilon = \left|\Delta\delta_D / \bar{\delta}_D\right|$. Normalized amplitudes of the diffusive term [$1/(1+\varepsilon)^2$] and advective term [$\varepsilon/(1+\varepsilon)$] in the mixed transport regime (MT).

3.3. Multiple scale analysis and transport regimes of the SBL problem

The multiple scale analysis and the definition of the transport regimes of the DSL problem described by Eq. (10) was established following the same methodology applied to the DBL problem.

In the particular case of highly organic sediments, the DSL mapped problem consists on the simplified diffusive-reactive one dimensional transport equation given by Eq. (11). In these sediments, typically, $T_0 << \bar{\tau}$ and $\bar{T}_{D,s} << \bar{\tau}$. Consequently, the set of possible transport regimes reduces to one: the *Reactive Dominant Transport* regime (RDT).

The coupling between the DBL and the DSL in these highly reactive sediments is through the influence of the DBL transport regime on the DO concentration at the water side of the WSI. Accordingly, the DO penetration depth in the porous media increases when the DBL

dynamics is governed by the advective dominant transport (ADT) regime. On contrary, events of anoxia in the sediment media are mainly related to DBL dynamics governed by the diffusive dominant transport (DDT) regime.

4 CONCLUSIONS

- The influence of the temporal variability of the DBL thickness on the coupled transport of DO across the WSI under low-energetic oscillatory flows was investigated as function of the characteristic time scales of: (1) the oscillatory flow, (2) the diffusive transport processes, and (3) the kinetics of the oxygen consumption processes involved.
- Using a multiple scale analysis of the mapped DO transport problem induced by oscillatory flows it is showed that the variability of the DBL thickness is an additional degree of freedom in the WSI mass transport problem and enriches the possible mass transport regimes of the WSI. The analysis probes that the fluctuations of the DBL thickness lead to an additional advective transport term in the mapped transport problem of the DBL.
- The mean value ($\overline{\delta}_D$), the amplitude ($\Delta\delta_D$) and the characteristic time scale (T_D) of the fluctuations of the DBL thickness are found to be key parameters to identify the dominant DO transport regime across the WSI. Three different transport regimes were found in the DBL transport problem. Moreover, the threshold value of the ratio $\Delta\delta_D / \overline{\delta}_D$ for changing the transport regime across the DBL from diffusive-dominant to advective-dominant was found.
- The variability of the DO penetration depth and the occurrence of anoxia events in the sediment domain are found to be related with two of these transport regimes in highly organic sediments. These findings are consisting with previous numerical results by [2] for slow oscillatory motions of monochromatic type.

REFERENCES

[1] E. Sánchez-Badorrey, S. Bramato, C. Mans, M.A. Losada and E. Cowen, "Bottom boundary layer under standing-regular wave groups: Eulerian drift and BBL width", *Proc. 30th Int. Conf. Coastal Engineering*, ASCE. 15pp. Ed. Word Scientific, (2006).

[2] M. Chatelain and K.Guizien, "Modelling coupled turbulence – Dissolved oxygen dynamics near the sediment-water interface under wind waves and sea swell", *Water Research*, Vol. **44**, pp. 1361-1372, (2010).

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis(Eds)

NUMERICAL SIMULATION OF ACOUSTICS IN HETEROGENEOUS MEDIA

DOMINIK STAAB*, STEFANIE NOWAK*, DÖRTE C. STERNEL*, AND MICHAEL SCHÄFER*

*Institute of Numerical Methods in Mechanical Engineering
Technische Universität Darmstadt
Dolivostraße 15, 64293 Darmstadt, Germany
e-mail: staab@fnb.tu-darmstadt.de, web page: http://www.fnb.tu-darmstadt.de

Key words: acoustics, heterogeneous media, sound, reflection, transmission, impedance

Abstract. Noise influenced by multiphase flows becomes an increasingly important issue. Numerical simulations can help for a better understanding of sound behavior within heterogeneous media. In this work, we present a coupled approach for computing incompressible heterogeneous fluids with sound wave propagation. The linearized Eulerian equations are used for the transport of the sound waves. We investigate the propagation of the waves in the presence of a fluid/fluid interface and we verify the introduced approach. The method is also tested with an inclined interface, where different refraction angles are investigated. The refraction as well as the reflection of sound waves can be predicted with errors of less than one percent.

1 INTRODUCTION

In the modern world noise is a serious problem which often generates health problems [1]. Therefore, it is important to understand the sources of noise and its propagation mechanisms. A considerable part of noise is produced by multiphase flows, such as waterfalls, waves hitting ships, wave-breakers, or sloshing in car tanks. A further field of application is, for example, the meteorology, where rain on the sea is measured based on its sound generation [2].
Generally, the prediction of sound parameters can be simulated by the use of the compressible Navier-Stokes equations (direct noise computation, DNC [3]). Due to highly different spatial and temporal scales in fluid flow and acoustics, DNC is not efficient for low Mach numbers. To avoid the multiscale problem, the method of acoustic/viscous splitting [4] is employed, where the acoustic effects are considered by a superposition of the incompressible flow field and the fluctuating perturbations (expansion about incompressible flow, EIF). In this approach the incompressible Navier-Stokes equations are

solved to obtain the acoustic source terms. Shen and Sørensen [5, 6] modified the acoustic/viscous splitting method of Hardin and Pope [4] and later Seo and Moon [7] used the linearized compressible perturbation equations.
Based on the latter method, the sound field is computed by the linearized Euler equations in the context of a finite-volume method. We introduce an extension of the method for heterogeneous media based on [8]. In the following sections the basic theory is introduced, the numerical approach is verified and results for reflection and refraction test cases are shown.

2 GOVERNING EQUATIONS

In this section the governing equations of fluid dynamics and acoustics are given. Throughout the paper, the Einstein summation convention is used. Sound waves are a compressible phenomenon caused by varying flow variables. The general equations that describe the flow situation are the compressible Navier-Stokes equations [9]:

$$\frac{\partial \rho}{\partial t} + \frac{\partial (\rho u_i)}{\partial x_i} = 0, \tag{1}$$

$$\frac{\partial (\rho u_i)}{\partial t} + \frac{\partial (\rho u_i u_j)}{\partial x_j} = \frac{\partial}{x_j}\left[\mu\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} - \frac{2}{3}\frac{\partial u_k}{\partial x_k}\delta_{ij}\right)\right] - \frac{\partial p}{\partial x_i} + \rho f_i, \tag{2}$$

with the velocity vector u_i, the pressure p, the density ρ, the external forces f_i and the Cartesian coordinates x_i.
In the splitting approach it is assumed that the density, the velocity and the pressure are composed of an incompressible background flow (superscript *inc*) governed by the incompressible Navier-Stokes equations with a superimposed acoustic perturbation (superscript *ac*):

$$\rho = \rho^{inc} + \rho^{ac}, \tag{3}$$

$$u = u^{inc} + u^{ac}, \tag{4}$$

$$p = p^{inc} + p^{ac}. \tag{5}$$

This approach is part of the *Expansion about Incompressible Flow* [5, 6]. The inhomogeneous linearized Eulerian equations for the acoustic quantities read:

$$\frac{\partial \rho^{ac}}{\partial t} + \rho^{inc}\frac{\partial u_i^{ac}}{\partial x_i} + u_i^{inc}\frac{\partial \rho^{ac}}{\partial x_i} = 0, \tag{6}$$

$$\rho^{inc}\frac{\partial u_i^{ac}}{\partial t} + \rho^{inc}u_j^{inc}\frac{\partial u_i^{ac}}{\partial x_j} + \frac{\partial p^{ac}}{\partial x_i} = 0, \tag{7}$$

$$\frac{\partial p^{ac}}{\partial t} + c^2\rho^{inc}\frac{\partial u_i^{ac}}{\partial x_i} + c^2 u_i^{inc}\frac{\partial \rho^{ac}}{\partial x_i} = -\frac{\partial p^{inc}}{\partial t}. \tag{8}$$

A detailed derivation is given in [10].

3 NUMERICAL METHODS

The numerical treatment of linear acoustics for a uniform medium where the density ρ^{inc} and the speed of sound c are the same at every point, has been introduced and investigated in [10, 11]. In the following we show the differences in the treatment of heterogeneous media in comparison to uniform media. Due to the heterogeneous media the density ρ^{inc}, the viscosity μ^{inc}, and the speed of sound c depend on its location x, y, z:

$$\rho^{inc} = \rho^{inc}(x, y, z) = \rho^{inc}(\mathbf{x}), \tag{9}$$

$$\mu^{inc} = \mu^{inc}(x, y, z) = \mu^{inc}(\mathbf{x}), \tag{10}$$

$$c = c(x, y, z) = c(\mathbf{x}). \tag{11}$$

The acoustic equations can be written in flux formulation:

$$\frac{\partial U}{\partial t} + \frac{\partial F}{\partial x} + \frac{\partial G}{\partial y} + \frac{\partial H}{\partial z} = Q, \tag{12}$$

with the variable vector

$$U = \left[\rho^{ac}, u^{ac}, v^{ac}, w^{ac}, p^{ac}\right]^T. \tag{13}$$

The fluxes F, G, and H read in flux formulation:

$$F = \begin{bmatrix} u^{inc}\rho^{ac} + \rho^{inc}(\mathbf{x})u^{ac} \\ u^{inc}u^{ac} + \frac{p^{ac}}{\rho^{inc}(\mathbf{x})} \\ u^{inc}v^{ac} \\ u^{inc}w^{ac} \\ u^{inc}p^{ac} + \rho^{inc}(\mathbf{x})c^2u^{ac} \end{bmatrix}, \tag{14}$$

$$G = \begin{bmatrix} v^{inc}\rho^{ac} + \rho^{inc}(\mathbf{x})v^{ac} \\ v^{inc}u^{ac} \\ v^{inc}v^{ac} + \frac{p^{ac}}{\rho^{inc}(\mathbf{x})} \\ v^{inc}w^{ac} \\ v^{inc}p^{ac} + \rho^{inc}(\mathbf{x})c(\mathbf{x})^2v^{ac} \end{bmatrix}, \tag{15}$$

$$H = \begin{bmatrix} w^{inc}\rho^{ac} + \rho^{inc}(\mathbf{x})w^{ac} \\ w^{inc}u^{ac} \\ w^{inc}v^{ac} \\ w^{inc}w^{ac} + \frac{p^{ac}}{\rho^{inc}(\mathbf{x})} \\ w^{inc}p^{ac} + \rho^{inc}(\mathbf{x})c^2(\mathbf{x})w^{ac} \end{bmatrix}, \tag{16}$$

and the source term

$$Q = \left[0, 0, 0, 0, -\frac{\partial p^{inc}}{\partial t}\right]^T. \tag{17}$$

By transformation of the LEE into local coordinate system of a cell face with coordinate ξ normal to the cell face it can be reduced to a one-dimensional problem:

$$\frac{\partial U_\xi}{\partial t} + A(\xi)\frac{\partial U_\xi}{\partial \xi} = 0, \tag{18}$$

with the Jacobi matrix

$$A(\xi) = \begin{bmatrix} u(\xi)^{inc}_\xi & \rho(\xi)^{inc} & 0 \\ 0 & u(\xi)^{inc}_\xi & 1/\rho(\xi)^{inc} \\ 0 & \rho(\xi)^{inc}c(\xi)^2 & u(\xi)^{inc}_\xi \end{bmatrix}. \tag{19}$$

The resulting Riemann problem is assumed to have a piecewise constant impedance $Z = \rho^{inc}c$ (layered media), then the resulting problem is reduced to solve a general Riemann problem [8]. We use a high-resolution scheme consisting of an exact Riemann solver and the Lax-Wendroff method [10, 11].

4 TEST CASES AND RESULTS

In order to verify the implemented method the numerical order is computed. The computation domain and the initial location of the sound wave are presented in Fig. 3. The initial values of the acoustic density and of the sound particle velocity in y-direction are $\rho^{ac} = 0$ and $v^{ac} = 0$, respectively. The sound wave is defined by:

$$p^{ac} = 2^{-1200(x-0.25)^2} Pa, \tag{20}$$

$$u^{ac} = 0.001 \cdot 2^{-1200(x-0.25)^2} m/s, \tag{21}$$

where p^{ac} and u^{ac} are the sound pressure and the sound particle velocity in x-direction. The wave runs from fluid 1 to fluid 2. The speed of sound and the impedance are c_i and Z_i, where i indicates the fluid. The computations are carried out on four grid levels with 16384, 8192, 4096 and 2046 control volumes in x-direction. The Courant number for the finest grid is chosen to be 0.8. The resulting order is shown in Fig. 2. The order of the method tends to the analytical order of two. Table 1 shows the comparison of the numerically computed transmission coefficients C_T^{num} and reflection coefficients C_R^{num} as well as the analytical coefficients C_T^{ana} and C_R^{ana} with the 2064 cells grid.

Table 1: Transmission and reflection coefficients for the grid with 2064 cells

	$c_1 < c_2$	$c_1 > c_2$
C_T^{ana}	0.4	1.6
C_T^{num}	0.3998	1.5998
C_R^{ana}	-0.6	0.6
C_R^{num}	-0.59998	0.5998

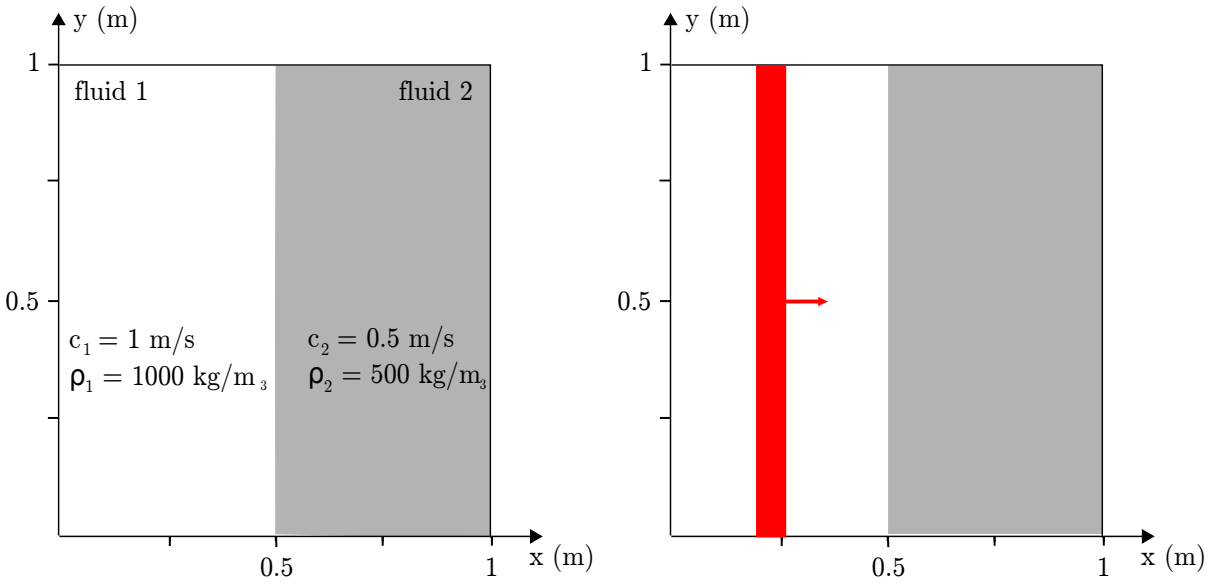


Figure 1: Test case setup with initial values and material parameters for each fluid (left side); the sound wave is indicated with a red line and runs from fluid 1 to fluid 2 (right side)

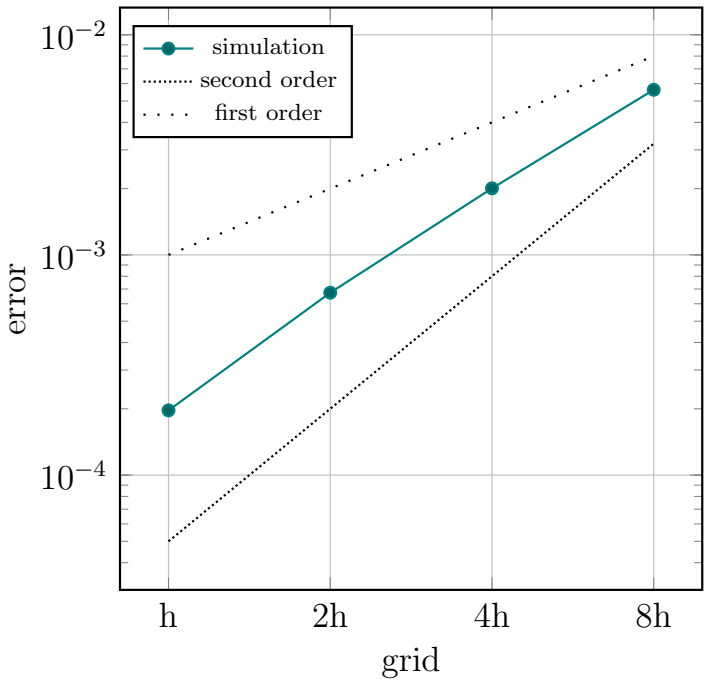


Figure 2: Order of accuracy of the quasi-one-dimensional acoustic test case; the order tends to the analytical order of two

Inclined interface

Next we investigate the sound wave behavior with an inclined interface. The setup is similar to the previous test case, but with an inclined interface, which is shown in Fig. 3. The number of cells in each coordinate direction is 512. The boundary conditions are set to outlets everywhere. We concentrate on the orientation of the transmitted wave as well as the reflected wave. We discuss the physical phenomena briefly. More details are given in [12, 13, 14], for example.

If the direction of an incoming plane wave is not perpendicular to the interface of the phases, it will be deflected. That means, the wave does not move over a straight line from a point in medium 1 to a point in medium 2, but takes the way which uses the least time. This phenomenon is called *Fermat's Principle* [15]. From this principle the law of refraction can be derived:

$$n = \frac{\sin(\alpha_1)}{\sin(\alpha_2)} = \frac{c_1}{c_2}, \tag{22}$$

where α_1 is the angle between the moving direction of the incoming wave and the normal vector to the interface. α_2 is the angle between the transmitted wave direction and the normal to the interface (Fig. 3). c_1 and c_2 are the speeds of sound connected to each medium and n is the refraction index. The angle of reflection should be the same as the incoming wave angle

$$\alpha_1 = \alpha_{1r}. \tag{23}$$

The transmission and the reflection angle is determined by the sound particle velocity. The order of accuracy for the particle velocity shown in Fig. 3 is given in Table 2 for different grid resolutions for $\alpha_1 = 10\,^\circ$.

Table 2: Order of accuracy of the two-dimensional acoustic computation

Grid resolution	L_2-norm (particle velocity)	order
124	$2.180 \cdot 10^{-4}$	
256	$2.624 \cdot 10^{-4}$	
512	$2.597 \cdot 10^{-4}$	1.55
1024	$2.624 \cdot 10^{-4}$	1.98

In the following the number of grid cells is 512 in both coordinate directions. In Table 3 results for four different incoming angles are shown. The reflected as well as the transmitted angles are in good agreement with the analytical values. The error E is computed by

$$E = \frac{|\alpha_{numeric} - \alpha_{analytic}|}{\alpha_{analytic}}. \tag{24}$$

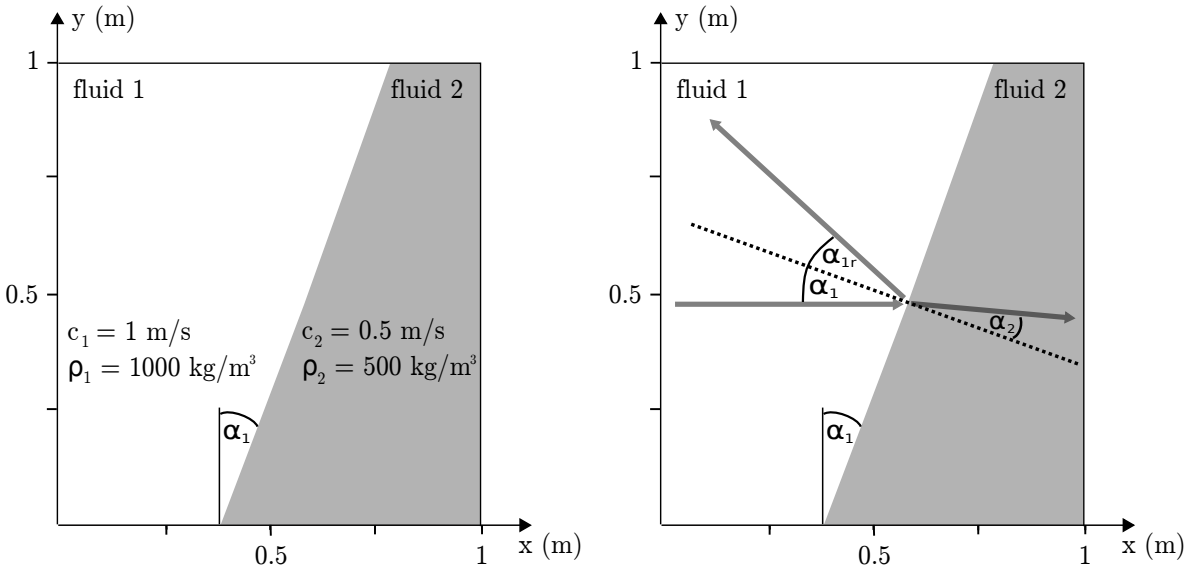


Figure 3: Test case setup with inclined interface; left: material properties; right: the wave refracts with angle α_w and reflects with angle α_{1r}

Table 3: Investigation of different reflection and transmission angles for $c_1 > c_2$ for 512 x 512 grid cells

α_1 (°)	α_{1r} (°)	E(α_1) (%)	α_{2ana} (°)	α_{2num} (°)	E(α_2) (%)
10	10.007	0.07	4.98	4.986	0.12
30	30.003	0.01	15	15.05	0.33
45	44.997	0.006	20.7	20.71	0.04
60	59.993	0.011	25.6	25.69	0.35

All errors are smaller than 1%. For an interchanged speed of sound ($c_2 > c_1$), total reflection is possible and the limitation angle can be computed by

$$\alpha_T = \arcsin\left(\frac{c_1}{c_2}\right) = \arcsin\left(\frac{0.5}{1}\right) = 30°. \tag{25}$$

That means, for $\alpha_1 \geq 30\,°$, the wave is totally reflected and not transmitted. Results for a lower angle are shown in Tab. 4. The computation error again is lower than 1%.

Table 4: Investigation of the reflection and transmission angles for $c_2 > c_1$ and $\alpha_1 = 10°$ and for 512 x 512 grid cells

α_1 (°)	α_{1r} (°)	E(α_1) (%)	α_{2ana} (°)	α_{2num} (°)	E(α_2) (%)
10	10.035	0.35	20.32	20.29	0.15

5 CONCLUSION AND OUTLOOK

In this paper we presented a coupled method for sound wave propagation in heterogeneous media, which is derived by the splitting approach. We introduced numerical test case setups to verify the method. The order of accuracy of the transmitted wave is in good agreement with the theoretical order of two. Furthermore, the method is applied to an inclined interface. The errors for the refraction and transmission of an incoming wave are less than one percent. The strength of this approach is the decoupling of the incompressible flow and the acoustic quantities, which allows a more flexible computation. The multiscale problem, consisting of a high speed of sound and a low flow velocity can be treated efficiently. In this paper we presented a first step of this method, where there is no background flow. As a next step, this method will be applied to more realistic problems. A background flow will be added, non-flat interface will be considered and surface tension should be taken into account.

REFERENCES

[1] Kazi A, Oommen A. *Chronic noise stress-induced alterations of glutamate and gamma-aminobutyric acid and their metabolism in the rat brain.* Noise and Health 16(73). pages 343-349. 2014.

[2] Nystuen J, McPhaden M, Freitag H. *Surface measurements of precipitation from an ocean mooring: The underwater acoustic log from the south china sea.* Journal of Applied Meteorology 39. pages 2182-2197. 2000.

[3] Bailly C, Bogey C, Marsden O. *Progress in direct noise computation.* International Journal of Aeroacoustics. Volume 9. Number 1. pages 123-143. 2010

[4] Hardin J, Pope D. *An acoustic/viscous splitting technique for computational aeroacoustics*. Theoretical and Computational Fluid Dynamics 6. pages 323-340. 1994

[5] Shen W, Sørensen J. *Aeroacoustic Modelling of Low-Speed Flows*. Theoretical and Computational Fluid Dynamics 13. pages 271-289. 1999

[6] Shen W, Sørensen J. *Comment on the aeroacoustic formulation of Hardin and Pope*. AIAA Journal 37(1). pages 141-143. 1999

[7] Seo JH, Moon YJ. *The perturbed Compressible Equations for Aeroacoustic Noise Prediction at Low Mach Numbers*. AIAA Journal 43(8). pages 10-23. 2005.

[8] Leveque RJ. *Finite-Volume Methods for Hyperbolic Problems*. Cambridge University Press. 2002.

[9] Schäfer M. 2006. *Computational Engineering - Introduction to Numerical Methods*. Springer.

[10] Kornhaas M. *Effiziente numerische Methoden fr die Simulation aeroakustischer Probleme mit kleinen Machzahlen*. PhD. TU Darmstadt. 2012.

[11] Kornhaas M, Sternel DC, Schäfer M. *Efficient numerical simulation of aeroacoustics for low Mach number flows interacting with structures*. Computational Mechanics. To appear.

[12] Lerch R, Sessler G, Wolf D. 2009. *Technische Akustik - Grundlagen und Anwendungen*. (4th edn). Springer.

[13] Filippi P, Habault D, Lefebvre JP, Bergassoli A. 1998. *Acoustics - Basic Physics, Theory and Methods*. (1th edn). Academic Press.

[14] Möser M. 2012. *Technische Akustik*. (9th edn). Springer.

[15] Perlick V. 2013. *Ray Optics, Fermat's Principle, and Applications to General Relativity*. Springer.

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis(Eds)

THE EFFICIENCY COMPARISON OF THE VORTEX ELEMENT METHOD AND THE IMMERSED BOUNDARY METHOD FOR NUMERICAL SIMULATION OF AIRFOIL'S HYDROELASTIC OSCILLATIONS

ILIA K. MARCHEVSKY, VICTORIYA S. MOREVA AND VALERIA V. PUZIKOVA

Bauman Moscow State Technical University (BMSTU)
Applied Mathematics department
Russia, 105005 Moscow, 2nd Baumanskaya, 5
e-mail: iliamarchevsky@mail.ru, morevavs@rambler.ru, valeria.puzikova@gmail.com

Key words: Vortex Element Method, Immersed Boundary Methods, The LS-STAG Method, Coupled Hydroelastic Problems, Wind Resonance

Abstract. In the present research the well-known test problem of wind resonance phenomenon simulation is considered. The Vortex Element Method and the LS-STAG method are used for its solving and their comparison is carried out. The obtained results can be useful for scientists and engineers who develop and operate the constructions which structural elements oscillate under hydrodynamic forces.

1 INTRODUCTION

Lagrangian meshless Vortex Element Methods [1, 2] are well-known numerical methods which efficiency can be very high when solving coupled aerohydroelastic problems. They allow to simulate both viscous and inviscid incompressible flows in bounded and unbounded domains. Vortex Element Methods for 2D flows are well developed and there are number of approaches for viscosity accounting (e.g., Viscous Vortex Domains method [3]) and for boundary conditions satisfaction. The main advantage of Vortex Element Method is that there is no necessity of mesh constructing and reconstructing when the airfoil moves and the airfoil can be of arbitrary shape. It also provides small numerical viscosity and requires sufficiently small memory and time of computations.

For flow simulation around airfoils with complicated shape or when the Reynolds number is about tens of thousands the number of vortex elements should be very large to provide the necessary accuracy. There are some approaches for accuracy improvement [4], mainly based on the modified mathematical models, an also number of approaches for

computations speedup, which presuppose parallel computational algorithms and fast approximate multipole methods usage. These approaches allow to simulate unsteady flows and to solve directly coupled hydroelastic problems even on personal computers.

Another effective method for coupled hydroelastic problems solving which also doesn't require mesh reconstruction is Immersed Boundary Method [5]. Its LS-STAG [6] modification is one of the most accurate algorithms because it provides correct approximation of the governing equations both on rectangular fluid cells and cut-cells. Because of rectangular mesh usage the uniform 5-point stencil inside the flow region and 4 or 3-point stencils at the boundaries are used, so it is possible to use high-efficiency numerical methods (e.g. Krylov subspaces, multigrid preconditioners etc.) for linear systems solving. RANS-based turbulence models have been recently implemented to LS-STAG method, so it can be used in coupled problems when the Reynolds number is about tens of thousands.

In the present research the well-known test problem of wind resonance phenomenon simulation is considered. The Vortex Element Method and the LS-STAG method are used for its solving and their comparison is carried out.

2 GOVERNING EQUATIONS

The problem is considered in 2D unsteady case when the flow around an airfoil is viscous and incompressible. The continuity and momentum equations are the following:

$$\nabla \cdot \vec{V} = 0, \qquad \frac{\partial \vec{V}}{\partial t} + (\vec{V} \cdot \nabla)\vec{V} = \nabla p + \frac{1}{\mathrm{Re}}\Delta \vec{V}. \tag{1}$$

Here $\vec{V} = \vec{V}(x, y, t) = u \cdot \vec{e}_x + v \cdot \vec{e}_y$ is the dimensionless velocity, $p = p(x, y, t)$ is the dimensionless pressure. The boundary conditions are the following:

$$\vec{V}\big|_{\text{inlet}} = \vec{V}_\infty, \qquad \frac{\partial \vec{V}}{\partial \vec{n}}\Big|_{\text{outlet}} = 0, \qquad \frac{\partial p}{\partial \vec{n}}\Big|_{\text{inlet\&outlet}} = 0, \tag{2}$$

$$\vec{V}\big|_{\text{airfoil}} = \vec{V}^{\,\text{ib}}, \qquad \frac{\partial p}{\partial \vec{n}}\Big|_{\text{airfoil}} = 0. \tag{3}$$

Here $\vec{V}^{\text{ib}}$ is the velocity of the immersed boundary. The airfoil is assumed to be rigid.

To simulate wind resonance phenomenon we consider the motion of the circular airfoil with diameter D across the stream (with one degree of freedom). Airfoil's constrain assumed to be linear viscoelastic Kelvin — Voigt-type (fig. 1) and its motion is described by the following ordinary differential equation:

$$m\ddot{y}_* + b\dot{y}_* + ky_* = F_y. \tag{4}$$

Here m is the airfoil mass, b is small damping factor, k is the constraint's elasticity coefficient, F_y is lift force, y_* is the deviation from the equilibrium. The natural frequency of the system $\omega \approx \sqrt{k/m}$ can be changed by varying of the coefficient k.

The deviation from the equilibrium on the n-th step of computation is $y_*^n = Y_C^n - Y_C^0$. Here Y_C^0 is the ordinate of the airfoil center at the initial time and Y_C^n is the ordinate of the airfoil center at the n-th step of computation.

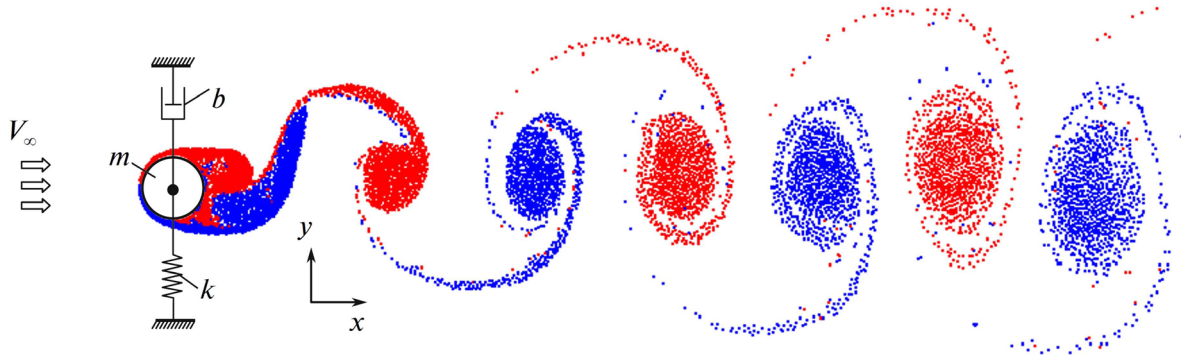


Figure 1: The circular airfoil with Kelvin — Voigt viscoelastic constraint and vortex wake behind it

In all numerical simulations with the following values of dimensionless parameters have been chosen: $\mathrm{Re} = 1000$, $V_\infty = 3.0$, $m = 39.15$, $b = 0.731$. The dimensionless natural frequency of the system is in the following range:

$$\mathrm{Sh}_\omega = \frac{\omega}{2\pi} \cdot \frac{D}{V_\infty} = 0.150 \ldots 0.280. \tag{5}$$

3 MAIN IDEAS OF THE VORTEX ELEMENT METHOD

Navier — Stokes equations (1) could be written down in Helmholtz form using vorticity vector $\vec{\Omega}(\vec{r},\, t) = \nabla \times \vec{V}(\vec{r},\, t)$:

$$\frac{\partial \vec{\Omega}}{\partial t} + \nabla \times (\vec{\Omega} \times \vec{U}) = 0. \tag{6}$$

Here $\vec{U}(\vec{r}, t) = \vec{V}(\vec{r}, t) + \vec{W}(\vec{r}, t)$, $\vec{W}(\vec{r}, t)$ is the so-called diffusive velocity [3], which is proportional to viscosity coefficient:

$$\vec{W}(\vec{r}, t) = \nu \frac{(\nabla \times \vec{\Omega}) \times \vec{\Omega}}{|\vec{\Omega}|^2}. \tag{7}$$

If vorticity distribution is known, flow velocity can be computed using Biot — Savart law:

$$\vec{V}(\vec{r}) = \vec{V}_\infty + \frac{1}{2\pi} \int_S \frac{\vec{\Omega}(\vec{\xi}, t) \times (\vec{r} - \vec{\xi})}{|\vec{r} - \vec{\xi}\,|^2} dS. \tag{8}$$

Equation (6) means that vorticity in the flow moves and its velocity is $\vec{U}$. 'New' vorticity is being generated only on airfoil surface ant this vortex layer intensity can be found from boundary condition of airfoil surface.

Vortex element method is meshless particle-type methods, so the vorticity field in the flow is discretized into separate vortex elements. Each vortex element is described by its position $\vec{r}_i$ and circulation Γ_i, $i = 1, \ldots, N$, where N is number of vortex elements in the flow. So the discretized Biot — Savar law has the following form:

$$\vec{V}(\vec{r}) = \vec{V}_\infty + \sum_{i=1}^{N} \frac{\Gamma_i}{2\pi} \frac{\vec{k} \times (\vec{r} - \vec{r}_i)}{|\vec{r} - \vec{r}_i|^2} + \oint_K \frac{\vec{k} \times (\vec{r} - \vec{\xi})}{2\pi |\vec{r} - \vec{\xi}|^2} \gamma(\vec{\xi})\, dl_\xi. \tag{9}$$

Here $\vec{k}$ is unit vector of the axis which is orthogonal to the plane of the flow.

Vortex elements movement according to (6) is simulated via solving the following ordinary differential equations system:

$$\frac{d\vec{r}_i}{dt} = \vec{V}(\vec{r}_i) + \vec{W}(\vec{r}_i), \quad i = 1, \ldots, N. \tag{10}$$

Number of vortex elements in the flow N changes at every time step because of the vorticity flux from the airfoil surface, which is simulated by vortex element generation near the airfoil. Their circulations are calculated from the vortex layer intensity $\gamma(\vec{\xi})$ on the airfoil surface. The circulations of all vortex elements in the flow remain constant and they can change only in special numerical procedure of vortex wake restructuring which allows to merge closely spaced vortex elements and lower their number in the flow.

Vortex layer intensity $\gamma(\vec{\xi})$ is unknown and can be found from the boundary condition

$$\vec{V}_-(\vec{r}) = \vec{V}^{\,\mathrm{ib}}(\vec{r}), \quad \vec{r} \in K. \tag{11}$$

Here $\vec{V}_-(\vec{r})$ is limit value of velocity from the airfoil side,

$$\vec{V}_-(\vec{r}) = \vec{V}_\infty + \sum_{i=1}^{N} \frac{\Gamma_i}{2\pi} \frac{\vec{k} \times (\vec{r} - \vec{r}_i)}{|\vec{r} - \vec{r}_i|^2} + \oint_K \frac{\vec{k} \times (\vec{r} - \vec{\xi})}{2\pi |\vec{r} - \vec{\xi}|^2} \gamma(\vec{\xi})\, dl_\xi - \frac{\gamma(\vec{r})}{2} \left(\vec{k} \times \vec{n}(\vec{r})\right) \tag{12}$$

$\vec{n}(\vec{r})$ is unit normal vector on the airfoil surface in point $\vec{r}$.

It can be shown [7] that we can solve one of scalar equation

$$\vec{V}_-(\vec{r}) \cdot \vec{n}(\vec{r}) = \vec{V}^{\,\mathrm{ib}}(\vec{r}) \cdot \vec{n}(\vec{r}) \qquad \text{or} \qquad \vec{V}_-(\vec{r}) \cdot \vec{\tau}(\vec{r}) = \vec{V}^{\,\mathrm{ib}}(\vec{r}) \cdot \vec{\tau}(\vec{r}) \tag{13}$$

instead of vector equation (11). Here $\vec{\tau}(\vec{r})$ is unit tangent vector on the airfoil surface. Mathematically there is no difference between solutions of these equations, but from computational point of view the corresponding numerical schemes are very different.

In 'classical' approach [1, 2, 3] unknown vortex layer intensity on the airfoil surface is assumed to be piecewise constant function and it satisfies equation $\vec{V}_- \cdot \vec{n} = \vec{V}^{\,\mathrm{ib}} \cdot \vec{n}$, which corresponds to equality of the flow and airfoil's velocity normal components on the airfoil surface and leads to singular integral equation of the 1-st kind

$$\oint_K \frac{\left[\vec{k} \times (\vec{r} - \vec{r}_0)\right] \cdot \vec{n}(\vec{r})}{2\pi |\vec{r} - \vec{r}_0|^2} \gamma(\vec{r}_0)\, dl_{r_0} = -\vec{n}(\vec{r}) \cdot \left(\vec{V}_\infty - \vec{V}^{\,\mathrm{ib}}(\vec{r}) + \sum_{i=1}^{N} \frac{\Gamma_i}{2\pi} \frac{\vec{k} \times (\vec{r} - \vec{r}_i)}{|\vec{r} - \vec{r}_i|^2} \right). \tag{14}$$

The solution of (14) certainly exists due to form of right side of this equation, but it is not unique. In order to select the unique solution an additional condition should be added:

$$\oint_K \gamma(\vec{r}) dl_r = G. \tag{15}$$

The kernel of equation (14) is unbounded and it has nonintegrable Hilbert-type singularity when $|\vec{r}-\vec{r}_0| \to 0$, so special numerical schemes are used for Cauchy principal value computation. They allow to obtain the solution of linear system approximating (14) with high accuracy when number of collocating points on the airfoil is large and its surface is smooth curve. It is proved [1] that in this case numerical solution converges to exact one in some integral (Hölder) norm. This approach lies in the basis of the 'classical' method; it can be called 'NVEM' — Vortex Element Method with normal components of velocity on airfoil surface.

The well-known numerical schemes [1], which are effective in vortex element method for inviscous fluids, can be generalized for viscous case, but the difference between numerical and exact solutions in uniform norm sometimes becomes significant. In order to take into account correctly the influence of the term with sum in the right side we need to discretize the airfoil extremely precisely, but it will lead to linear system with inadmissibly big dimension which coefficients and right side computation as well as solution have a very big computation cost.

At the same time if we simulate flow around the airfoil with angle points or sharp edges using NVEM, the error sometimes even increases proportionally to number of collocating points on the airfoil surface [4, 8]. So it is impossible to determine the vortex layer intensity with high accuracy and classical NVEM-schemes can't be applied for 2D Navier — Stockes equations solution for airfoils with angle points and sharp edges. The matrix of the linear system also can be ill-conditioned.

It also should be noted that linear algebraic system corresponding to (14) becomes ill-conditioned for airfoils with angle points or sharp edges.

In order to solve the mentioned problems, the alternative approach based on ideas [7] is developed by the authors [4, 8]. Vortex layer intensity is determined from solution of equation $\vec{V}_- \cdot \vec{\tau} = \vec{V}^{\,\mathrm{ib}} \cdot \vec{\tau}$, corresponding to equality of tangent component of flow velocity limit value and airfoil surface velocity. It leads to Fredholm-type integral equation of the 2-nd kind with bounded (for smooth airfoils) kernel:

$$\oint_K \frac{\left[\vec{k} \times (\vec{r}-\vec{r}_0)\right] \cdot \vec{\tau}(\vec{r})}{2\pi|\vec{r}-\vec{r}_0|^2} \gamma(\vec{r}_0) dl_{r_0} - \frac{\gamma(\vec{r})}{2} = -\vec{\tau}(\vec{r}) \cdot \left(\vec{V}_\infty - \vec{V}^{\,\mathrm{ib}}(\vec{r}) + \sum_{i=1}^{N} \frac{\Gamma_i}{2\pi} \frac{\vec{k} \times (\vec{r}-\vec{r}_i)}{|\vec{r}-\vec{r}_i|^2} \right). \tag{16}$$

Solution of equation (16) is also non-unique, so the same additional condition (15) as in classical method is used. This method is called 'TVEM' — Vortex Element Method with tangent components of velocity on airfoil surface.

Equation (16) also can be approximated with linear algebraic system which is well-conditioned both for smooth and non-smooth airfoils. Due to equation kernel boundness an arbitrary quadrature formula can be used for integral approximation in (16). In simplest case we also can consider vortex layer intensity to be piecewise constant function.

Results of numerical experiments show that errors are sufficiently big, but they could be significantly decreased if we consider some 'weak' formulation of (16): integral equa-

tion (16) in discrete numerical scheme will be satisfied not in separate collocation points $\vec{r}_j$, $j = 1, \ldots, N$ of airfoil surface, but on an average on airfoil surface parts (panels) K_p whose lengthes are L_p, $p = 1, \ldots, N$:

$$\frac{1}{L_p}\int_{K_p}\left[\oint_K \frac{\left[\vec{k}\times(\vec{r}-\vec{r}_0)\right]\cdot\vec{\tau}(\vec{r})}{2\pi|\vec{r}-\vec{r}_0|^2}\gamma(\vec{r}_0)dl_{r_0}\right]dl_r - \frac{1}{L_p}\int_{K_p}\frac{\gamma(\vec{r})}{2}dl_r =$$

$$= -\frac{1}{L_p}\int_{K_p}\vec{\tau}(\vec{r})\cdot\left(\vec{V}_\infty - \vec{V}^{\,\mathrm{ib}}(\vec{r})\right)dl_r - \frac{1}{L_p}\sum_{i=1}^{N}\frac{\Gamma_i}{2\pi}\int_{K_p}\frac{\vec{k}\times(\vec{r}-\vec{r}_i)}{|\vec{r}-\vec{r}_i|^2}dl_r, \quad p = 1, \ldots, N. \quad (17)$$

The other modification concerns uniform vorticity distribution on airfoil surface. In 'classical' NVEM method intensity of vortex layer assumed to be constant on every part (every panel) of the airfoil, but then all the vorticity from every panel concentrates in one point on the panel and integral in (14) transforms into a sum of influences of discrete (point) vortex elements. In the developed modified method TVEM vorticity assumed to be uniformly distributed over every panel. Every panel on the airfoil is straight-line segment, so internal integral in first term (17) transforms into a sum of influences of panels with uniformly distributed vorticity which intensity on the q-th panel is equal to γ_q:

$$\sum_{q=1}^{N}\frac{\gamma_q}{L_p}\int_{K_p}\left[\int_{K_q}\frac{\left[\vec{k}\times(\vec{r}-\vec{r}_0)\right]\cdot\vec{\tau}(\vec{r})}{2\pi|\vec{r}-\vec{r}_0|^2}dl_{r_0}\right]dl_r - \frac{\gamma_p}{2} =$$

$$= -\frac{1}{L_p}\int_{K_p}\vec{\tau}(\vec{r})\cdot\left(\vec{V}_\infty - \vec{V}^{\,\mathrm{ib}}(\vec{r})\right)dl_r - \frac{1}{L_p}\sum_{i=1}^{N}\frac{\Gamma_i}{2\pi}\int_{K_p}\frac{\vec{k}\times(\vec{r}-\vec{r}_i)}{|\vec{r}-\vec{r}_i|^2}dl_r, \quad p = 1, \ldots, N. \quad (18)$$

All the integrals in (18) can be calculated analytically, the corresponding formulae are derived in [4, 8]. Numerical results show that the developed TVEM-scheme and especially its 'weak' form allows to increase the accuracy significantly: sometimes the error becomes a tens or even hundreds times smaller [4, 8].

In order to compute the pressure distribution and hydrodynamic force which effect on the airflow, the analogues of Bernoulli and Cauchy — Lagrange integrals are used [9]. This approach is very effective in Vortex Element method and it can be equally applied for both NVEM and TVEM schemes.

It should be noted that computation cost of simulating fixed and movable rigid airfoils when using vortex element method remains nearly the same, so they are very suitable for coupled aerohydroelastic problems. However vortex element movement simulation is 'N-body'-type problem, and in practice number of vortex element can be on the order of tens and sometimes even hundreds of thousand, so special acceleration algorithms should be implemented. Well-known Barnes — Hut fast algorithm analogue [10] can be very effective, especially when using accurate analytical estimate of its computational cost [11] which allows to choose its parameters optimally. Parallel computation algorithms are also used in order to reduce time of computations [12].

4 MAIN IDEAS OF THE LS-STAG METHOD

The Cartesian mesh with cells $\Omega_{i,j} = (x_{i-1}, x_i) \times (y_{j-1}, y_j)$ is introduced in the rectangular computational domain. It is denoted that $\Gamma_{i,j}$ is the face of $\Omega_{i,j}$ and $\vec{x}^c_{i,j} = (x^c_i, y^c_j)$ is the center of this cell. Unknown components $u_{i,j}$ and $v_{i,j}$ of velocity vector $\vec{v}$ are computed in the middle of fluid parts of the cell faces. These points are the centers of control volumes $\Omega^u_{i,j} = (x^c_i, x^c_{i+1}) \times (y_{j-1}, y_j)$ and $\Omega^v_{i,j} = (x_{i-1}, x_i) \times (y^c_j, y^c_{j+1})$ with faces $\Gamma^u_{i,j}$ and $\Gamma^v_{i,j}$ and areas M^x_{ij} and $M^y_{i,j}$ respectively (fig. 2). If $i = \overline{1, N}$, $j = \overline{1, M}$, base mesh, x-mesh and y-mesh contain $E = N \cdot M$, $E_x = (N-1) \cdot M$ and $E_y = N \cdot (M-1)$ cells correspondingly.

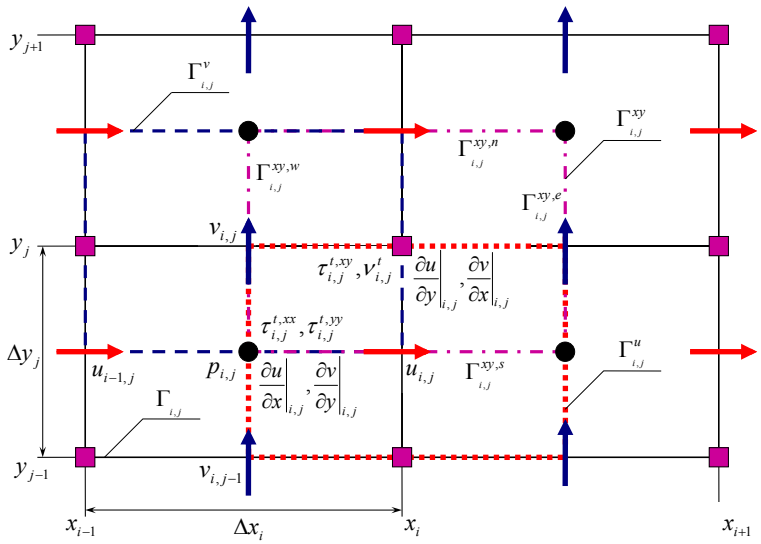


Figure 2: Staggered arrangement of the variables on the LS-STAG mesh

The level-set function $\varphi = \varphi(x, y)$ [13] is used for immersed boundary Γ^{ib} description [6]. Cells which the immersed boundary intersects are the so-called 'cut-cells' (fig. 3). These cells contain the solid part together with the liquid one. The boundary Γ^{ib} is represented by a line segment on the cut-cell $\Omega_{i,j}$. Location of this segment endpoints are defined by a linear interpolation of the variable $\varphi_{i,j} = \varphi(x_i, y_j)$.

The cell-face fraction ratios $\vartheta^u_{i,j}$ and $\vartheta^v_{i,j}$ are introduced in [6]. They take values in interval $[0, 1]$ and represent the fluid parts of the east and north faces of $\Gamma_{i,j}$ respectively. One-dimensional linear interpolation of $\varphi(x_i, y)$ on the segment $[y_{j-1}, y_j]$ and $\varphi(x, y_j)$ on the segment $[x_{i-1}, x_i]$ is used for the cell-face fraction ratios computing:

$$\vartheta^u_{i,j} = \frac{\min(\varphi_{i,j-1}, \varphi_{i,j})}{\min(\varphi_{i,j-1}, \varphi_{i,j}) - \max(\varphi_{i,j-1}, \varphi_{i,j})}, \quad \vartheta^v_{i,j} = \frac{\min(\varphi_{i-1,j}, \varphi_{i,j})}{\min(\varphi_{i-1,j}, \varphi_{i,j}) - \max(\varphi_{i-1,j}, \varphi_{i,j})}. \quad (19)$$

In 2D case, the cut-cells can be classified into trapezoidal, triangular and pentagonal cells. Examples of each type cut-cells are presented on fig. 3.

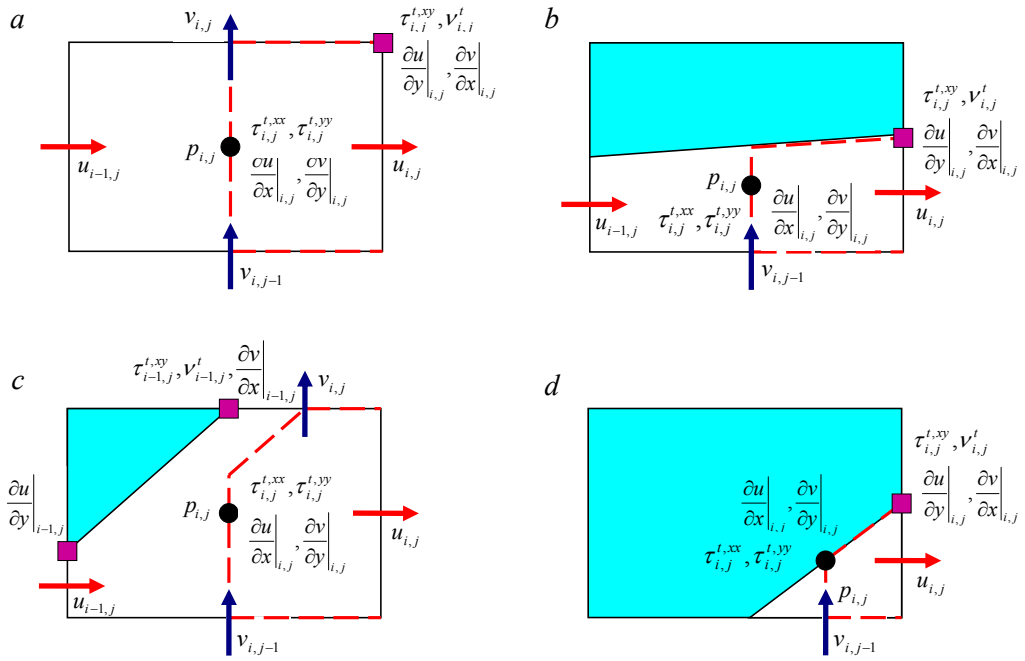


Figure 3: Location of the variables discretization points on the LS-STAG mesh: *a* – Cartesian Fluid Cell; *b* – North Trapezoidal Cell; *c* – Northwest Pentagonal Cell; *d* – Northwest Triangle Cell.

To preserve the five-point structure of the stencil of the MAC method we need to make distinction between the discretization of the normal and shear stresses (fig. 3).

Hydrodynamic force can be computed by the following formulae:

$$F_x = \sum_{\text{Cut-cells } \Omega^{ib}_{i,j}} \left[(\vartheta^u_{i-1,j} - \vartheta^u_{i,j}) \Delta y_j \left(p_{i,j} - \nu \frac{\partial u}{\partial x}\Big|_{i,j} \right) - \nu \text{Quad}^{ib}_{i,j} \left(\frac{\partial u}{\partial y} \vec{e}_y \cdot \vec{n} \right) \right], \tag{20}$$

$$F_y = \sum_{\text{Cut-cells } \Omega^{ib}_{i,j}} \left[-\nu \text{Quad}^{ib}_{i,j} \left(\frac{\partial v}{\partial x} \vec{e}_x \cdot \vec{n} \right) + (\vartheta^v_{i,j-1} - \vartheta^v_{i,j}) \Delta x_i \left(p_{i,j} - \nu \frac{\partial v}{\partial y}\Big|_{i,j} \right) \right].$$

The quadrature of the shear stresses $\text{Quad}^{ib}_{i,j}$ has to be adapted to each type of cut-cells.

According to the concept of the LS-STAG method equations (1) should be written in integral form for cell of base mesh, cell of x-mesh and cell of y-mesh respectively:

$$\int_{\Gamma_{i,j}} \vec{v} \cdot \vec{n} \, dS = 0, \tag{21}$$

$$\frac{d}{dt} \int_{\Omega^u_{i,j}} u \, dV + \int_{\Gamma^u_{i,j}} (\vec{v} \cdot \vec{n}) u \, dS + \int_{\Gamma^u_{i,j}} p \vec{e_x} \cdot \vec{n} \, dS - \int_{\Gamma^u_{i,j}} \nu \nabla u \cdot \vec{n} \, dS = 0,$$

$$\frac{d}{dt} \int_{\Omega^v_{i,j}} v \, dV + \int_{\Gamma^v_{i,j}} (\vec{v} \cdot \vec{n}) v \, dS + \int_{\Gamma^v_{i,j}} p \vec{e_y} \cdot \vec{n} \, dS - \int_{\Gamma^v_{i,j}} \nu \nabla v \cdot \vec{n} \, dS = 0.$$

The general form of the LS-STAG discretization for (21) is the following [6]:

$$D^x U_x + D^y U_y + \overline{U}^{ib} = 0, \tag{22}$$
$$\frac{d}{dt}(M^x U_x) + C^x U_x + G^x P - \nu K^x U_x + S_x^{ib,c} - \nu S_x^{ib,\nu} = 0,$$
$$\frac{d}{dt}(M^y U_y) + C^y U_y + G^y P - \nu K^y U_y + S_y^{ib,c} - \nu S_y^{ib,\nu} = 0.$$

Here $P \in \mathbb{R}^E$ is the discrete pressure, $U_x \in \mathbb{R}^{E_x}$ and $U_y \in \mathbb{R}^{E_y}$ are the discrete components of the velocity vector; $S_x^{ib,c} \in \mathbb{R}^{E_x}$, $S_x^{ib,\nu} \in \mathbb{R}^{E_x}$, $S_y^{ib,c} \in \mathbb{R}^{E_y}$, $S_y^{ib,\nu} \in \mathbb{R}^{E_y}$ are source terms; $\overline{U}^{ib} \in \mathbb{R}^E$ is the mass flux; $D^x \in M(\mathbb{R})_{E \times E_x}$, $D^y \in M(\mathbb{R})_{E \times E_y}$ are the divergence discrete analogues; $K^x \in M(\mathbb{R})_{E_x \times E_x}$ and $K^y \in M(\mathbb{R})_{E_y \times E_y}$ represent the discretization of the diffusive terms; $C^x \in M(\mathbb{R})_{E_x \times E_x}$ and $C^y \in M(\mathbb{R})_{E_y \times E_y}$ represent the discretization of the convective terms; $G^x = -D_x^T$ and $G^y = -D_y^T$ are the gradient discrete analogues.

The time integration of the differential algebraic system (22) is performed with a semi-implicit Euler scheme. Predictor step leads to discrete analogues of the Helmholtz equation for velocities prediction $\widetilde{U}_x$, $\widetilde{U}_y$ at the time $t_{n+1} = (n+1)\Delta t$:

$$\frac{M_x^{n+1}\widetilde{U}_x - M_x^n U_x^n}{\Delta t} + C_x^n U_x^n + S_x^{ib,c,n} - D_x^{T,n} P^n - \nu K_x^{n+1}\widetilde{U}_x - \nu S_x^{ib,\nu,n+1} = 0, \tag{23}$$
$$\frac{M_y^{n+1}\widetilde{U}_y - M_y^n U_y^n}{\Delta t} + C_y^n U_y^n + S_y^{ib,c,n} - D_y^{T,n} P^n - \nu K_y^{n+1}\widetilde{U}_y - \nu S_y^{ib,\nu,n+1} = 0.$$

Here Δt is the constant time discretization step. Corrector step leads to the following discrete analogue of Poisson equation for $\Phi = \Delta t(P^{n+1} - P^n)$:

$$A^{n+1}\Phi = D_x^{n+1}\widetilde{U}_x + D_y^{n+1}\widetilde{U}_y + \overline{U}^{ib,n+1}, \tag{24}$$

$A = -D^x(M^x)^{-1}(D^x)^T - D^y(M^y)^{-1}(M^y)^T$, $A \in M(\mathbb{R})_{E \times E}$. Then flow variables at the time point t_{n+1} are computed by the following formulae:

$$U_x^{n+1} = \widetilde{U}_x + (M_x^{n+1})^{-1} D_x^{T,n+1}\Phi,\ U_y^{n+1} = \widetilde{U}_y + (M_y^{n+1})^{-1} D_y^{T,n+1}\Phi,\ P^{n+1} = \frac{\Phi}{\Delta t} + P^n. \tag{25}$$

Linear systems (23), (24) are solved using the BiCGStab method with the ILU- and multigrid [14] preconditioning. The optimal parameters of the multigrid preconditioner were chosen using the original algorithm for the solver cost-coefficient estimation [15].

5 NUMERICAL EXPERIMENTS

Vortex Element Method implemented in POLARA software package [12] allows to simulate flow-induced vibrations of the airfoil. The computational cost of the simulation process is sufficiently small: time of computations in sequential mode and in parallel mode for 200 second of physical time is shown in table 1 for different values of Sh_ω value. All the computations were performed on cluster with Intel Core i7 2,4 GHz processors.

Table 1: Computational time (in hours) for vortex element method

Sh_ω		1 CPU	2 CPU	4 CPU	8 CPU	16 CPU
0.17	(no resonance)	41.3	22.6	12.0	7.1	4.7
0.21	(max amplitude)	63.4	34.7	17.9	10.1	6.6
0.24	(close to resonance)	45.3	25.4	13.3	7.6	4.8

About 80 computations have been produced for different values of the dimensionless frequency and the unsteady process have been simulated. At the initial time there were still flow and the airfoil in equilibrium position. Time step Δt was equal to 0.01, number of panels which approximates the airfoil $N_p = 200$. Then the velocity of the incident flow became greater; at time moment $t = 1.0$ (after 100 time steps) it was equal to $V_\infty = 3.0$ and then remained constant. After the transient mode airfoil's oscillation in all cases became close to periodical, their amplitudes dependency on the natural frequency Sh_ω is shown on fig. 4, *a* (dots connected by line).

The fig. 4, *a* shows that there is a sharp increase in the amplitude of oscillations at $\mathrm{Sh}_\omega \approx 0.198$. It's well known that there is hysteresis-type phenomenon [16] and in order to simulate it the following computations were performed: from $t = 0$ to $t = 100$ (10 000 time steps) Sh_ω was equal to 0.21; at this time the oscillations become steady with amplitude $A/D \approx 0.47$, then the constraint's elasticity coefficient was changed abruptly to the values which correspond to Sh_ω from 0.178 to 0.198 with step 0.00025. In each case after the transient mode new steady oscillations were generated, and their amplitudes are shown on fig. 4, *b* (dots connected by solid line).

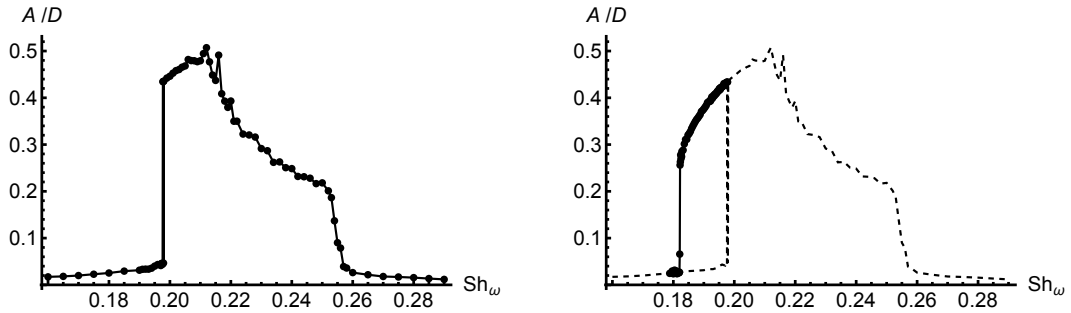


Figure 4: Maximum amplitude of the circular airfoil oscillations simulated using vortex element method: *a* – airfoil's initial state id equilibrium position, *b* — airfoil's initial state is close to resonance oscillations

The obtained results for maximum amplitude of oscillation, the resonance frequency and hysteresis properties are in good agreement with the results given in [16, 17].

When using the LS-STAG method number of computations have been performed on non-uniform grid 272×292 with time discretization step $\Delta t = 0.0001$. The maximum amplitudes of oscillations dependency on the Strouhal number Sh_ω is shown on fig. 5.

Computational results are in good qualitative agreement with the previous studies [17]. Maximum amplitude is about $0.4D$ and it occurs when the natural frequency of the system St_ω is close to the Strouhal number, calculated for a fixed airfoil $\mathrm{St} \approx 0.24$ [6].

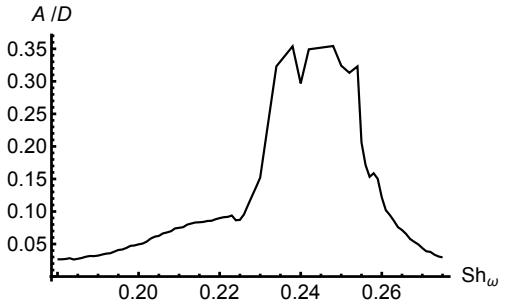


Figure 5: Maximum amplitude of the circular airfoil oscillations at Re = 1000 simulated using the LS-STAG method

The original program package developed by the authors requires approximately 250 hours to simulate 200 seconds of unsteady airfoil's oscillations at resonance mode and about 180 hours at non-resonance mode. In order to obtain more accurate results we need more detailed mesh, but the computational cost of the simulation for such mesh will be extremely high because the program package operates in sequential mode. Its parallelization and acceleration is very important but non-trivial problem.

6 CONCLUSIONS

Two different approaches to numerical simulation of flow-induced vibrations of the airfoil in incompressible flow are considered. The first approach is based on the meshless lagrangian Vortex Element Method while the second one corresponds to the LS-STAG immersed boundary method. Software packages are developed for the numerical simulation of the airfoils' motion in the flow by using the mentioned methods and their modifications. The model problem of wind resonance of the circular airfoil simulation is considered. Both methods allows to obtain satisfactory results; Vortex Element Method seems to be more accurate and it need approximately a fourth of time which the LS-STAG method need for unsteady oscillations simulation. Nevertheless, LS-STAG method seems to be useful in practice for high-Reynolds flows simulation because it allows to implement arbitrary turbulence model while there are no modifications for Vortex Element Method which allows to solve Navier — Stokes equations using RANS, LES or DES approaches.

7 ACKNOWLEGEMENTS

The work was partially supported by Russian Federation President Grants for young scientists [proj. MK-3705.2014.8, MK-5357.2015.8].

REFERENCES

[1] Lifanov, I.K., Belotserkovskii, S.M. *Methods of Discrete Vortices.* CRC Press, 1993.

[2] Cottet, G.-H., Koumoutsakos, P.D. *Vortex Methods: Theory and Practice.* CUP, 2008.

[3] Dynnikova, G.Ya. Lagrange method for Navier — Stokes equations solving. *Doklady Akademii Nauk.* (2004) **399**: 42–46.

[4] Marchevsky, I.K. and Moreva, V.S. Vortex Element Method for 2D Flow Simulation with Tangent Velocity Components on Airfoil Surface. *ECCOMAS 2012 — European Congr. on Comp. Meth. in Appl. Sc. and Eng., e-Book.* (2012) 5952–5965.

[5] Mittal, R. and Iaccarino, G. Immersed boundary methods. *Annu. Rev. Fluid Mech.* (2005) **37**: 239–261.

[6] Cheny, Y. and Botella, O. The LS-STAG method: A new immersed boundary/level-set method for the computation of incompressible viscous flows in complex moving geometries with good conservation properties. *J. Comp. Phys.* (2010) **229**:1043–1076.

[7] Kempka, S.N., Glass, M.W., Peery, J.S. and Strickland, J.H. Accuracy Considerations for Implementing Velocity Boundary Conditions in Vorticity Formulations. *SANDIA Report SAND96-0583* (1996)

[8] Kuzmina, K.S. and Marchevsky, I.K. On Numerical Schemes in 2D Vortex Element Method for Flow Simulation Around Moving and Deformable Airfoils. *Advanced Prob. in Mech. (APM 2014): Proc. of the XLII Summer School-Conf.* (2014) 335–344.

[9] Andronov, P.R., Guvernyuk, S.V. and Dynnikova, G.Ya. *Vortex methods for unsteady hydrodynamical forces calculation.* Moscow University Press, 2006. [in Russian]

[10] Dynnikova, G.Ya. Fast technique for solving the N-body problem in flow simulation by vortex methods. *Comp. Math. and Math. Phys.* (2009) **49**: 1389–1396.

[11] Kuzmina, K.S. and Marchevsky, I.K. Estimation of computational complexity of the fast numerical algorithm for calculating vortex influence in the vortex element method. *Science and Education.* (2013) **10**: 399–414. [in Russian]

[12] Marchevsky, I.K. and Moreva, V.S. High-Efficiency POLARA Program for Airfoil Aerodynamic Characteristics Calculation Using Vortex Elements Method. *ICVFM2010 – Int. conf. on vortex flow and vortex methods.* (2010) 10–15.

[13] Osher, S. and Fedkiw, R.P. *Level set methods and dynamic implicit surfaces.* Springer, (2003).

[14] Wesseling, P. *An introduction to multigrid methods.* John Willey & Sons Ltd., (1991).

[15] Marchevsky, I.K. and Puzikova, V.V. OpenFOAM iterative methods efficiency analysis for linear systems solving. *Proceedings of the Institute for System Programming of RAS.* (2013) **24**: 71–86. [in Russian]

[16] Klamo, J.T., Leonard, A. and Roshko, A. The effects of damping on the amplitude and frequency response of a freely vibrating cylinder in cross-flow. *J. of Fluids and Struct.* (2006) **22**: 845–856.

[17] Klamo, J.T., Leonard, A. and Roshko, A. On the maximum amplitude for a freely vibrating cylinder in cross flow. *J. of Fluids and Struct.* (2005) **21**: 429–434.

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis(Eds)

THE NAVIER-STOKES EQUATIONS IN TWO DIFFERENT FORMULATIONS WITH MODERATE AND HIGH REYNOLDS NUMBERS

B. Bermúdez*, and W. Fermín Guerrero S.†

* Facultad de Ciencias de la Computación, Benemérita Universidad Autónoma de Puebla.
14 Sur y San Claudio
Ciudad Universitaria, Puebla
e-mail:bbj@cs.buap.mx
†Facultad de Ciencias Físico-Matemáticas, Benemérita Universidad Autónoma de Puebla.
18 Sur y San Claudio
Ciudad Universitaria, Puebla
e-mail:willi@fcfm.buap.mx

Key words: Navier-Stokes equations, Stream Function-vorticity Formulation, Velocity-vorticity formulation, moderate and high Reynolds numbers

Abstract. The goal of this work is to present results for 2D viscous incompressible flows governed by the Navier-Stokes equations. Two different formulations will be used: The Stream Function-vorticity and the Velocity-vorticity formulation. To show that the schemes we are using are working for moderate and high Reynolds numbers, we are going to report results for the very well known un-regularized driven cavity problem, with Reynolds numbers in the range of $3200 \leq Re \leq 50000$.

1 INTRODUCTION

We are going to work with the Navier-Stokes in two different formulations: The Stream Function-vorticity and the Velocity-vorticity formulation. The problem we are going to solve is the well known un-regularized driven cavity problem, with Reynolds numbers in the range of $3200 \leq Re \leq 50000$. Results, in both formulations, are obtained using a simple numerical scheme based on a fixed point iterative process (see [1]), applied to a nonlinear elliptic system resulting after time discretization. The scheme has shown to be robust enough to handle such Reynolds numbers, from moderate to high, which is not an easy task to deal with (see [2] and [3]). As the Reynolds number increases the mesh has to be refined and a smaller time step has to be used, numerically, by stability matters and physically, to capture the fast dynamics of the flow, as pointed out in [4], although, with the Velocity-vorticity formulation (see [5] and [6]), a finer mesh has to be used. So, because of this, computing time is in general very large with this numerical scheme and

for both formulations, so we seek to reduce this time by, instead of working only with the matrix A, resulting from the discretization of the Laplacian term, using both matrixes A and B, the second one resulting from the discretization of the advective term. For the Stream Function-vorticity formulation and moderate and high Reynolds numbers, this second scheme has been faster than the fixed point iterative method (see [7], [8]). For the Velocity-vorticity formulation we are just showing results using the fixed point iterative method, and we are still looking forward to modify this scheme by using here too both matrixes A and B and in this way being able to reduce computing time when using this formulation.

2 Mathematical Models

Let $\Omega \subset R^N$ $(N = 2, 3)$ the region of a nonsteady, viscous, incompressible flow, and Γ its boundary.

$$\begin{cases} \frac{\partial u}{\partial t} - \nabla^2 \mathbf{u} + \nabla p + (\mathbf{u} \cdot \nabla)\mathbf{u} & = f, \qquad (a) \\ \nabla \cdot \mathbf{u} & = 0 \qquad (b) \end{cases} \tag{1}$$

These are the Navier-Stokes equations in primitive variables and this system has to be supplemented with appropiate boundary and initial conditions.

2.1 The Stream function-vorticity Formulation

We will restrict ourselves to a bidimensional region Ω. Taking the curl in both sides of the equation $(1a)$ and taking into account that

$$\left\{ \; u_1 = \frac{\partial \psi}{\partial y} \,, \quad u_2 = -\frac{\partial \psi}{\partial x} \,, \right. \tag{2}$$

which follows from $(1b)$, with ψ the stream function and u_1, u_2, the two components of the velocity. So we get:

$$\begin{cases} \nabla^2 \psi & = -\omega & (a) \\ \frac{\partial \omega}{\partial t} - \frac{1}{Re} \nabla^2 \omega + \mathbf{u} \cdot \nabla \omega = f_\omega & & (b) \end{cases} \tag{3}$$

where ω is the vorticity ($\omega = \frac{\partial u_2}{\partial x} - \frac{\partial u_1}{\partial y}$). These are the Navier-Stokes equations in the Stream function-vorticity formulation. The inconmpressibility condition $(1b)$, by (2) is automatically satisfied, and the pressure does not appear any more, which is a significant advantage with respect to the primitive variable formulation.

2.2 Velocity-Vorticity Formulation

Taking the curl in

$$\omega = -\nabla \times \boldsymbol{u} \tag{4}$$

and using the identity $\nabla \times \nabla \times \mathbf{a} = -\nabla^2 \mathbf{a} + \nabla(\nabla \cdot \mathbf{a})$ and $(1b)$, a velocity Poisson equation results:

$$\nabla^2 \mathbf{u} = -\nabla \times \boldsymbol{\omega}. \tag{5}$$

Then, two Poisson equations for the velocity components are obtained, which together with the equation for the vorticity gives us

$$\begin{cases} \frac{\partial u_1}{\partial t} + \nabla^2 u_1 = -\frac{\partial \omega}{\partial y} & (a) \\ \frac{\partial u_2}{\partial t} + \nabla^2 u_2 = \frac{\partial \omega}{\partial x}\,. & (b) \\ \frac{\partial \omega}{\partial t} - \frac{1}{Re}\nabla^2 \omega + \mathbf{u} \cdot \nabla \omega = f_\omega & (c) \end{cases} \tag{6}$$

These are de Navier-Stokes equations in the Velocity-vorticity Formulation.

3 The Numerical Schemes.

For the time derivative appearing in the vorticity equation in both schemes, the following second order approximation is used:

$$\frac{\partial f}{\partial t}(\mathbf{x}, (n+1)\Delta t) = \frac{3f^{n+1} - 4f^n + f^{n-1}}{2\Delta t} \tag{7}$$

where $\mathbf{x}\,\epsilon\,\Omega$, $n \geq 1$, Δt denotes the time step, and $f^r \approx f(\mathbf{x}, r\Delta t)$, assuming f is smooth enough.

3.1 The Stream function-Vorticity formulation

The following nonlinear elliptic system has to be solved at each time level:

$$\begin{cases} \nabla^2 \psi = -\omega, \quad \psi|_\Gamma = \psi_{bc}; & (a) \\ \alpha\omega - \nu\nabla^2\omega + \mathbf{u} \cdot \nabla\omega = f_\omega, \quad \omega|_\Gamma = \omega_{bc}\,, & (b) \end{cases} \tag{8}$$

where $\alpha = \frac{3}{2\Delta t}$, $\nu = \frac{1}{Re}$ and $f_\omega = \frac{4\omega^n - \omega^{n-1}}{2\Delta t}$. To obtain (ψ^1, ω^1), the first subinterval is divided into M subintervals, and a first order scheme, such as Euler, is applied to each of the M subintervals.

Let R_ω be defined by:

$$R_\omega(\omega,\psi) \equiv \alpha\omega - \nu\nabla^2\omega + \mathbf{u}\cdot\nabla\omega - f_\omega \ . \tag{9}$$

So, system (8) is equivalent to:

$$\left\{ \begin{array}{lll} \nabla^2\psi & = -\omega \ \ in\ \Omega\,, & \psi = \psi_{bc} \quad on\ \ \Gamma \\ R_\omega(\omega,\psi) & = \ \ 0 \ \ in\ \Omega & \omega|_\Gamma = \omega_{bc} \end{array} \right. \tag{10}$$

This system is solved, at time level (n+1), by the following fixed point iterative process:

Given $\omega^{n,0} = \omega^n$, $\psi^{n,0} = \psi^n$ *solve until convergence in* ω *and* ψ

$$\left\{ \begin{array}{l} \nabla^2\psi^{n,m+1} = -\omega^{n,m} \quad in\ \Omega, \\ \psi^{n,m+1} = \psi_{bc}^{n,m+1}\ on\ \Gamma \\ (\alpha I - \nu\nabla^2)\omega^{n,m+1} = (\alpha I - \nu\nabla^2)\omega^{n,m} - \rho_\omega R_\omega(\omega^m, \psi^{n,m+1})\ in\ \Omega, \\ \omega^{n,m+1} = \omega_{bc}^{n,m+1}\ on\ \Gamma,\ \rho_\omega > 0. \end{array} \right. \tag{11}$$

and then, take $(\omega^{n+1}, \psi^{n+1}) = (\omega^{n,m+1}, \psi^{n,m+1})$.

In order to reduce computing time, we worked on solving system by the following method at each time step:

$$\left\{ \begin{array}{ll} \nabla^2\psi^{n+1} = -\omega^n, \quad \psi^{n+1}|_\Gamma = \psi_{bc}^{n+1}; & (a) \\ (\alpha I - \nu A)\omega^{n+1} + B\omega^{n+1} = f_\omega, \quad \omega^{n+1}|_\Gamma = \omega_{bc}^{n+1}\,, & (b) \end{array} \right. \tag{12}$$

Here, A and B are the matrixes asociated with the discretization of the difussive (Laplacian) and the advective term, respectively. The linear system of equations resulting, is solved using Gauss-Seidel.

3.2 The Velocity-Vorticity Formulation

For the time derivatives appearing in the vorticity equation (7) is used, and the following fully implicit time discretization system is obtained, in Ω,

$$\left\{ \begin{array}{ll} \frac{\partial u_1}{\partial t} + \nabla^2 u_1 & = -\frac{\partial \omega}{\partial y} \\ \frac{\partial u_2}{\partial t} + \nabla^2 u_2 & = \frac{\partial \omega}{\partial x}, \quad \mathbf{u}^{n+1}|_\Gamma = \mathbf{u}_{bc} \\ R_\omega(\omega, \mathbf{u}) = \mathbf{0}, & \omega|_{\mathbf{\Gamma}} = \omega_{\mathbf{bc}} \end{array} \right. \tag{13}$$

where

$$R_\omega(\omega, \mathbf{u}) \equiv \alpha\omega - \nu\nabla^2\omega + \mathbf{u}\cdot\nabla\omega - \mathbf{f}_\omega \ , \tag{14}$$

and using again the fixed point iterative method, we have:

Given $\omega^{n,0} = \omega^n$, $u_1^{n,0} = u_1^n$, $u_2^{n,0} = u_2^n$ *solve until convergence on* ω, u_1 and u_2

$$\begin{cases} \frac{\partial u_1^{n,m+1}}{\partial t} + \nabla^2 u_1^{n,m+1} = -\frac{\partial \omega^{n,m}}{\partial y} \\ \\ \frac{\partial u_2^{n,m+1}}{\partial t} + \nabla^2 u_2^{n,m+1} = \frac{\partial \omega^{n,m}}{\partial x}, \qquad \mathbf{u}^{n,m+1}|_\Gamma = \mathbf{u}_{bc}^{n,m+1} \\ \\ (\alpha I - \nu\nabla^2)\omega^{n,m+1} \ = (\alpha I - \Delta)\omega^m - \rho_\omega R_\omega(\omega^{n,m}, \mathbf{u}^{n,m+1}), \\ \quad \rho_\omega > 0, \ \omega^{n,m+1}|_\Gamma = \omega_{bc}^{n,m}. \end{cases} \tag{15}$$

and then, take $(\omega^{n+1}, u_1^{n+1}, u_2^{n+1}) = (\omega^{n,m+1}, u_1^{n,m+1}, u_2^{n,m+1})$.

4 Numerical experiments

The numerical experiments take place in rectangular domains $\Omega = (0, a)\times(0, b)$, $a, b > 0$, in connection with the lid-driven cavity problem (in our case, $a = 1$ and $b = 1$). The boundary condition of $\mathbf{u}$ is given by $\mathbf{u} = (1, 0)$ at the moving boundary $y = b$ and $\mathbf{u} = (0, 0)$ elsewhere.

4.1 Stream function-vorticity formulation

A translation of the boundary condition in terms of the velocity primitive variable $\mathbf{u}$ to the $\psi - \omega$ variables has to be performed. Following [9], $\psi = 0$ is chosen on Γ, and by Taylor expansion of $(8a)$ on the boundary, with h_x and h_y the space steps, one obtains:

$$\begin{cases} \omega(0, y, t) = -\frac{1}{2h_x^2}[8\psi(h_x, y, t) - \psi(2h_x, y, t)] + O(h_x^2) \\ \omega(a, y, t) = -\frac{1}{2h_x^2}[8\psi(a - h_x, y, t) - \psi(a - 2h_x, y, t)] + O(h_x^2) \\ \omega(x, 0, t) = -\frac{1}{2h_y^2}[8\psi(x, h_y, t) - \psi(x, 2h_y, t)] + O(h_y^2) \\ \omega(x, b, t) = -\frac{1}{2h_y^2}[8\psi(x, b - h_y, t) - \psi(x, b - 2h_y, t)] - \frac{3}{h_y} + O(h_y^2). \end{cases} \tag{16}$$

With both schemes mentioned above (using just matrix A with the fixed point iterative method, and using A and B) results agree very well, so we present the graphs obtained with the second scheme. The advantage of the second scheme is that the computing time requiered is almost half of the time needed for the fixed point iterative process.

In Figure 1 we show the streamlines and isovorticity conturs for $Re = 25000$ with $h = h_x = h_y = 1/512$. In Figure 2 we show results for $Re = 31000$ and the same value of h. In figure 3, we show results for $Re = 50000$, with $h = h_x = h_y = 1/1024$. For these Reynolds numbers, since there is no steady state, results are shown for $T_final = 5$.

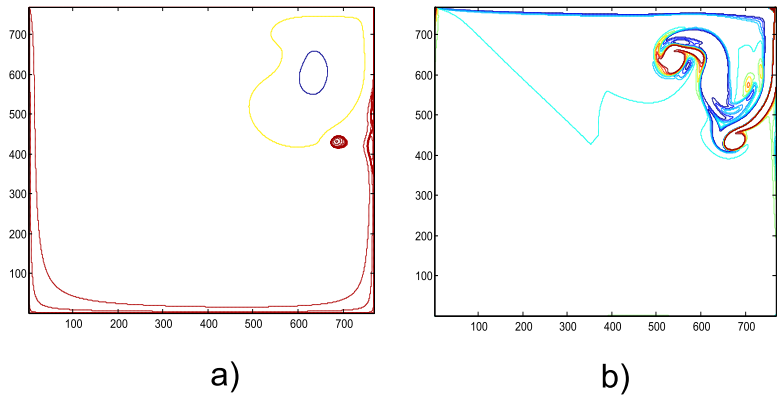


Figure 1: Streamlines (left) and isovorticity contours(right)for $Re = 25000, h = h_x = h_y = 1/728, dt = .00025, T_{final} = 5$

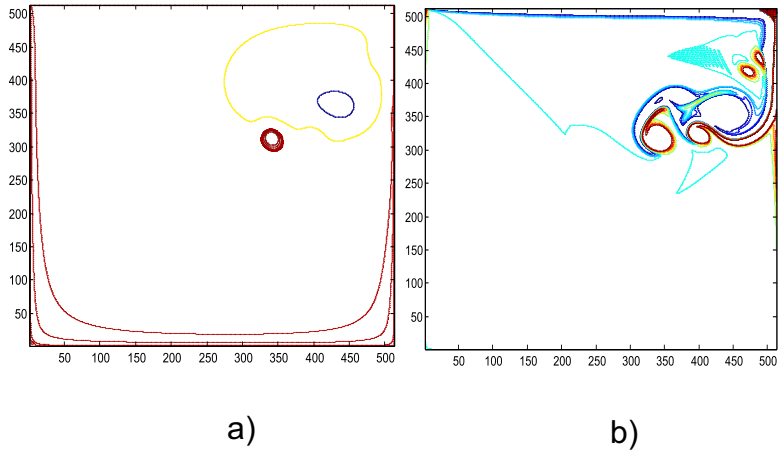


Figure 2: Streamlines (left) and isovorticity contours (right) for $Re = 31000, h = h_x = h_y = 1/512, dt = .00025, T_{final} = 5$

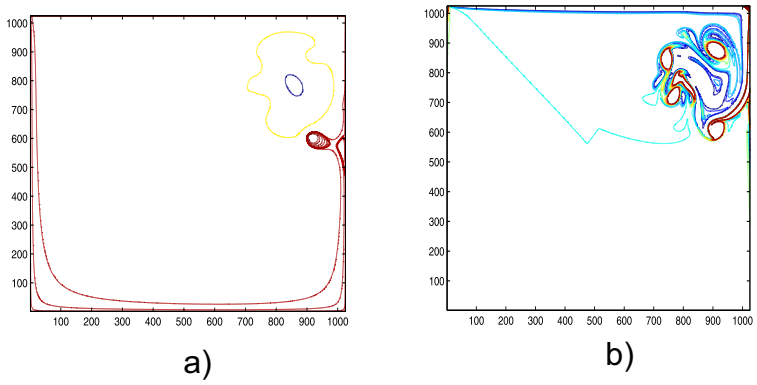


Figure 3: Streamlines (left) and isovorticity contours (right) for $Re = 50000, h = h_x = h_y = 1/1024, dt = .00025, T_{final} = 5$

4.2 Velocity-Vorticity formulation

The boundary conditions are given by:

$$\begin{cases} \mathrm{u}_1 = 0,\ u_2 = 0,\ \omega = \frac{\partial u_2}{\partial x} \text{ on } \Gamma_x = 0 \\ \mathrm{u}_1 = 0,\ u_2 = 0,\ \omega = \frac{\partial u_2}{\partial x} \text{ on } \Gamma_x = 1 \\ \mathrm{u}_1 = 0,\ u_2 = 0,\ \omega = -\frac{\partial u_1}{\partial y} \text{ on } \Gamma_y = 0 \\ \mathrm{u}_1 = 1,\ u_2 = 0,\ \omega = \frac{\partial u_1}{\partial x} \text{ on } \Gamma_y = 1 \end{cases} \tag{17}$$

The initial conditions are given by $\mathbf{u}(\mathbf{x}, \mathbf{0}) = (\mathbf{0}, \mathbf{0})$ and $\omega(\mathbf{x}, \mathbf{0}) = \mathbf{0}$

As remarked in [5], not all the results could be obtained with second order discretization. A fourth order discretization was required . This is accomplished by using the fourth order option of Fishpack ([10]) which was used in this work, for solving the elliptic problems arising.

In Figure 4 we show the streamlines and isovorticity contours for $Re = 3200, h = h_x = h_y = 1/512, T_{final} = 50$.

Finally, in Figure 5, we show just the vorticity contours For $Re = 20000$ with a) $h = h_x = h_y = 1/1512, T_{final} = 5$, obtained using the Velocity-vorticity formulation, and b) with the Stream Function-vorticity Formulation with $h = h_x = h_y = 1/768, T_{final} = 5$.

As can be seen, with the Stream Function-vorticity formulation we are using an h half of the one used for the Stream Function-vorticity formulation and we think that the results obtained with this formulation are much reliable; even more, the computing time used to obtain the results with the Velocity-vorticity formulation was much more greater than the one for the Stream Function-vorticity formulation. We think that there are still

some numerical problems with this formulation for high Reynolds numbers.

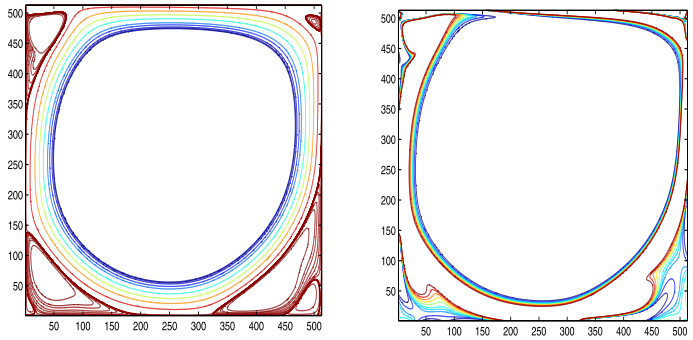


Figure 4: Streamlines (left) and isovorticity contours (right) for $Re = 3200, h = h_x = h_y = 1/512, dt = .0001, T_{final} = 50$

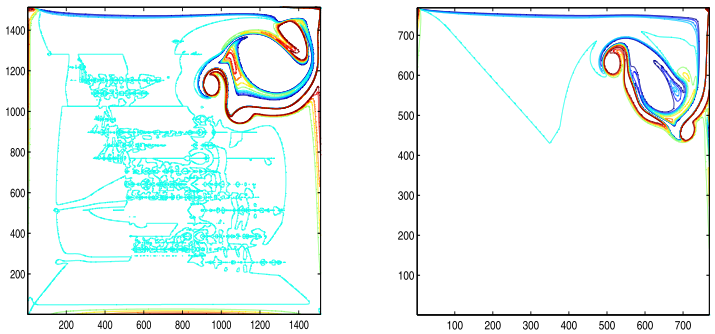


Figure 5: Isovorticity contours for $Re = 20000$, a) Velocity-vorticity formulation with $h = h_x = h_y = 1/1512, dt = .0001, T_{final} = 5$, b) Stream Function-vorticity formulation with $h = h_x = h_y = 1/768, dt = .0001, T_{final} = 5$.

5 Conclusions

We are presenting efficient numerical methods for solving the Navier-Stokes equations in the Stream funtion-vorticity and the Velocity-vorticity formulations.

Results agree very well with those reported in the bibliography (see [2], [3], [4]), but with the Stream function-vorticity formulation and working with matrixes A and B resulting from the discretization of the Laplacian and the advective term respectively, we were able to reduce computing time to almost one half with respect to the Fixed point iterative method, especially for high Reynolds numbers, for which the computing time increases a lot. In this case, as we have said, smaller values of h have to be used; numerically for stability and physically to capture the fast dynamics of the flow. We are still looking forward to reduce computing time. Previous results have already been obtained for $Re = 75000$ using the Stream function-vorticity formulation; they will be reported in a future work.

For the Velocity-vorticity formulation, as mentioned, we are just showing results using the fixed point iterative method, and we are still looking forward to modify this scheme by using here too both matrixes A and B and in this way being able to reduce computing time when using this formulation. This is the reason we are just showing results for Re up to 20000 and we did not go further.

In conclusion, the numerical procedure applied to the Stream Function-vorticity formulation is not as good for the Velocity-vorticity formulation, however, the way it behaves, through the discretization parameters, and the order of discretization, gives us another point of view of the behavior of flows under different numerical methods and different formulations, as pointed out in [5]. The difficulty of the Velocity-vorticity formulation is reinforced through some works such as [11], who, with very different methods reported driven cavity flows for moderate Reynolds numbers, lower than ours.

REFERENCES

[1] Nicolás, A. *A Finite element approach to the Kuramoto-Sivashinski equation.* Advances in Numerical Methods, SIAM (1991).

[2] Bermúdez B., Nicolás A., Sánchez F. J., Buendía E., Operator Splitting and upwinding for the Navier-Stokes equations, *Computational Mechanics* (1997) **20** 474–77

[3] Nicolás A., Bermúdez B., 2D incompressible viscous flows at moderate and high Reynolds numbers, *CMES* (2004) **6(5)** 441–451

[4] Nicolás-Carrizosa A., Bermúdez-Jurez B., Onset of two dimensional turbulence with High Reynolds numbers in the Navier-Stokes equations, *Proceedings IV International Conference on Computational methods for Coupled Problems,* (2011).

[5] Nicolás A., Bermúdez B., Viscous incompressible flows by the Velocity-Vorticity Navier-Stokes equations, *CMES* (2007) **20**No. 2, 73–83.

[6] Bermúdez B., Nicolás A., The Taylor Vortex and the Driven Cavity Problems by the Velocity-Vorticity Formulation, *Proceedings 7th. International Conference on Heat Transfer and Fluid Mechanics and Thermodynamics* (2010)

[7] Bermúdez B., Juárez L., Numerical Solution of an advection-diffusion equation, *Información Tecnológica* (2014) **25(1)** 151–160

[8] Bermúdez B., Posadas R., The Taylor Vortex and the Driven Cavity Problems in the Stream Function-Vorticity Formulation, Proceedings of the 11th. World Congress on Comp. Mechanics (WCCM XI), 5th European Congress on Computational Mechanics (ECCM V) 6th European Congress on Computational Fluid Dynamics (ECDF VI), *A publication of CIMNE* (2014) 3746–3755

[9] Goyon, O., High Reynolds numbers solutions of Navier-Stokes Equations using incremental unknownks, *Comput. Methods Appl. Math. Engrg.* (1996) 319–335

[10] Adams, J., Swarstrauber P., and Sweet R., FISHPACK: A Package of Fortran subroutines for the solution of separable Elliptic PDE's, The National Center for Atmospheric Research Noulder, Colorado, U.S.A. (1980)

[11] Grimaldi A., Pascazio G., Napolitano M., A Parallel Multi-block Method for unsteady Vorticity-velocity Equations, texitCMES, (2006) **14(1)** 45–56

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis (Eds)

A NUMERICAL COMPONENT TO EVALUATE THE IODINE CORROSION IN A NUCLEAR FUEL ACCOUNTING FOR CHEMISORPTION MECHANISMS ZIRCONIA LAYER IMPACT ON SCC.

N. BARNEL* AND D. PALERMO†

* Electricité de France, research and development (EDF R&D)
Avenue des Renardières
77 818 Morêt-sur-Loing, France
e-mail: nathalie.barnel@edf.fr, web page: http://www.edf.com

† YANTRA TECHNOLOGIES
Le Saouzé, Villa Nelumbo - 84 120 PERTUIS, France
email: david.palermo@yantra-technologies.com - Web page: http://www.yantra-technologies.com

Key words: chemisorptions, cladding, iodine, numerical component, SCC, zirconia

Abstract. In PWR plants, nuclear fuel pellets (cylinders 1.2 cm length, 0.84 cm diameter) are piled in nuclear fuel rods placed in assemblies, each containing 265 fuel rods. the fuel rod is a tube manufactured in a zirconium alloyed material. Fuel cylinders are piled to form a fuel stack shorter than the overall tube length to maintain a remaining upper volume called plenum, able to allow fuel stack elongation and accommodation of the gaseous and volatile fission products released during operation. The fuel stack (3.6-4 m depending on the design) is typically maintained during handling operation by a spring placed in the plenum. The fuel rod is closed at both ends by sealed end plugs after filling dead volumes with a neutral gas (helium) with a pressure ranging from 1 to 34.5 bars. The cladding insures the first confinement barrier against radionuclides dispersion. During severe power transients some fission products, like iodine, are expected to assist cracks initiation and propagation at the inner surface of the cladding. Our major concern here aims improving understanding the involved mechanisms, accounting for the zirconia layer able to develop as soon as a closed contact is locally established between the pellet fragments and the cladding inner surface.

In this work, we extend a specific numerical component, developed by EDF, coupling two numerical software codes [1]. The first one aims simulating the evolution of the fuel isotopic composition and the second one aims evaluating the fuel dioxide chemistry. The growth of the zirconia layer has been modelled and integrated in the numerical tool [2]. The scenario of Stress Corrosion Cracking (SCC) previously defined has been improved to consider the full impact of the zirconia layer on SCC. The role of the zirconia layer can be mechanical, thermal or even chemical. Assumptions and coupling of the phenomenon are discussed.

To make demonstration, a sensitivity analysis is performed. This study considers several

zirconia layer thickness for two cases: a high burn-up fuel rod operated in normal PWR conditions, and a fuel rod segment pre-irradiated in a power plant and then submitted to an incidental power transient simulation in a MTR. The results demonstrated a protective effect of the zirconia layer. As a consequence, the corrosive attack of the clad is delayed. The logarithm function of the iodine chemisorption kinetic appears to be proportional to the zirconia thickness. The simulation analysis validates the extension of the SCC scenario. The results are in good agreement with experimental observations.

1 INTRODUCTION

The numerical component NUAGES – Numerical Unit ensuring ANGE coupling with EDMOND and STRAP – is an embedded software structure "like russian dolls ", in which the "father" component call and plug three “son” components [Fig 1]-[1,2].

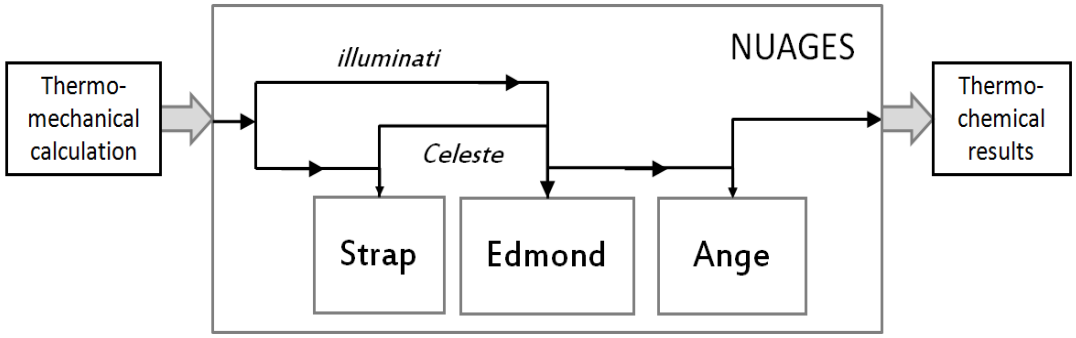


Figure 1 - Scheme of a calculation performed with NUAGES [3,4]

Version 1.1 of the component includes a scenario for modelling thermochemistry processes of the cladding Stress Constraint Corrosion (SCC). In parallel, it was developed to enable an estimation of the local zirconia layer growth at pellet cladding interface, and of the ballistic drop of fission products ejected from the pellet toward the cladding as well [5,6].

Bibliographical studies based on the analysis of a large number of destructive post ramp examinations, enlighten that the zirconia could have a protective effect against the SCC cracks propagations. The numerical component NUAGES, allows in a relatively simplified way modeling both phenomena and the way they could be coupled. We are going to present more in detailed the numerical tool developed and then show how results are quite promising regarding the prediction of the observations reported in literature [7,8].

2 PROTECTIVE EFFECT OF ZIRCONIA LAYER ON SCC

A very large number of post ramp examinations were conducted in the past and are still pursued to better understand and simulate the PCMI risk of failure during severe power transients. At this stage, most of observations lead to some main conclusions on which engineering studies agree:

1. The presence of oxygen in small quantity in the neighborhood accelerates SCC processes. It was evidenced by three tests: under air atmosphere, under inert gas atmosphere to lower the Oxygen potential and in presence of warmed shavings of zirconium (the oxidation of which lowers the partial pressure of oxygen with values close to those evaluated in the fuel rod) [9].

2. Presence of oxygen in high quantity in the test environment inhibits the reaction with the iodine. This was shown from local analyses realized on a surface of zirconium free of constraints, submitted to a molecular flow of iodine. The study concerns the impact of a flow of oxygen managed in the same place as the flow of iodine. It shows that there is a competition between the iodine and the zirconium chemical reactions on one hand, and the iodine and the oxygen chemical reactions on the other hand [10, 11].

3. Presence of a pre-oxidation layer important (upper to 0.2μm thickness) on the inner surface of the metal is beneficial; nevertheless this layer can break easily with the differential strain, letting the iodine reach locally the surface. Tests on pre-oxidized claddings either in an oven or before fabrication, show that a too thin layer, like the one naturally formed in ambient atmosphere (10-15 nm), has no protective effect on the metal [9].

Eric Walle proposed some additional experiments to specify the role of the oxygen during SCC [12]. Samples used are leaky rings (10 mm in diameter) of recrystallized Zircaloy-4. A strain control loading path is applied in non–irradiated conditions. As there is no additional mechanical stress on the sample, the loading is only representative of the crack initiation mechanism (and not for propagation). The tests differ by the oxygen pressure, which is controlled just before the test, when the sample is sealed in a glass tube. They also differ either from the presence (or not) of a pre-oxidized layer on the cladding inner surface, either from the test duration. The results seem to indicate:

4. That crack initiation is due to mechanics. On pre-oxidized samples, perpendicular cracks in the interface are developing in front of the oxidation forehead. The authors consider thus reasonable to think that the combined effect of the formation of oxide and the mechanical loading are at the origin of a purely mechanical initiation of cracks in the metal.

5. That the presence of oxygen localizes the phenomenon of SCC. Without oxygen we observe a generalized corrosion of the samples, while the presence of oxygen allows the localization of the SCC [Fig. 2]. The protection of the surface by a pre-oxidized layer of 1 or 10μm turns out to be a sufficient (but not necessary) condition for a localization of the chemical attack relative to SCC.

6. That the oxygen can lead to a passivation of the metal if the quantity is important enough so that the formation of an oxide layer on the metal surface is faster than the attack by the iodine [Fig 2].

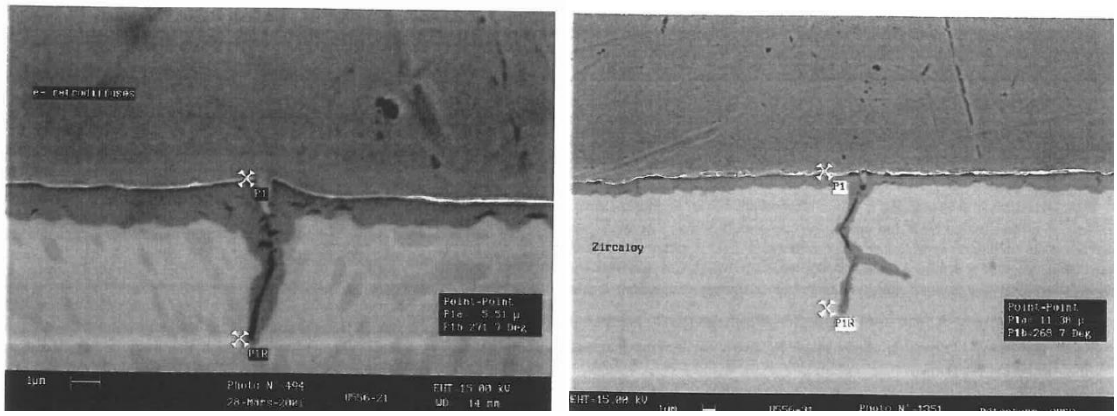


Figure 2 : Cross section of two samples, *in* [12]

The sample on the left is a reference that has been preoxydized at 1 μm, without iodine, and submitted to a 500 hour testing. The sample on the right has been preoxydized at 1 μm, with a partial pressure of oxygen of 0.01 bars at ambient temperature, the a diiodine partial pressure of 0.01 bars and finally submitted to a 150 hour testing.

3 EXTENDED MODELING SCENARIO OF SCC

Here are the main steps and assumptions for the modeling scenario retained for SCC without considering the impact of the zirconia layer [7,8]:

Step 1. Inventory of the neutronic balance sheet

Step 2. Ballistic migration and drop of fissions products through the pellet-cladding interface

Step 3. Thermochemistry: evaluation of condensed compounds and gaseous species (like I, I_2, CsI O_2) and also their chemical potential.

Step 4. Evaluation of corrosive source term

First assumptions. Fission gas released will be consider either from a conservative point of view (100% of fission gas is instantaneously released), either by means of a best estimate evaluation (the fission gas released rate is predicted by a thermo-mechanical code).

Step 5. Corrosive attack of the cladding

Second assumption: Cracking initiation is due to the mechanical loading (the corrosion only assists the failure mechanism)

Third assumptions: The corrosive attack is due to gaseous iodine. The monoatomic specie has been preferred to the diatomic one.

Fourth assumption: Because of lack of evidences, an hypothetical radiolysis of gaseous species is not considered. As a consequence gaseous CsI does not dissociate.

The main result of this scenario consists in an evaluation of the iodine and oxygen deposit rate on a zirconium crystal due to chemisorption. Indeed chemisorption is the first step for metallic corrosion.

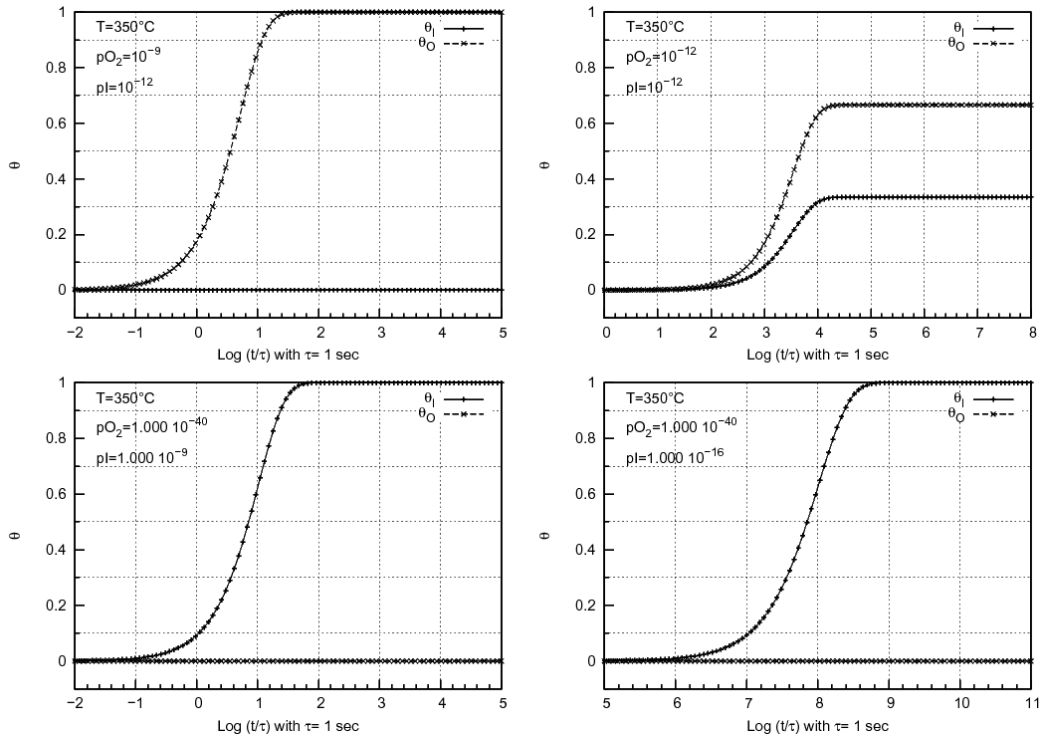


Figure 3. Chemisorption examples curves [13]

Figure 3 shows examples of chemisorption curves which can be obtained at the end of SCC modeling. Those of the top correspond to two virtual cases where the oxygen is either in excess compared with iodine, or in equivalent quantity. Both bottom graphs correspond to two cases presented in Eric Walle's pHD thesis [12], the one on the left would be representative of a rod at the end of its life and the one on the right of a rod submitted to a power transient.

Let us now consider the role of the zirconia layer:

From a mechanical point of view, the zirconia is known for its very low mechanical resistance, even if the literature does not mention very homogeneous quantitative data on the subject [14,15]. As far as the previous study led on the SCC assume mechanical cracking initiation process, the wisest choice is to adopt a conservative approach:

Fifth assumption: No mechanical resistance of the zirconia layer.
In case of a power transient, we consider that the zirconia layer is systematically broken.

So, according to this assumption, zirconia will have no mechanical protective role towards SCC. The latter could be only thermal or chemical.

From a thermal point of view, zirconia has a lower thermal conductivity than zicaloy. As a result, the thermal gradient in the cladding is modified. Considering that the thickness of the zirconia layer is low compared to the cladding thickness (ratio of the order of 1,5 %), it is likely that the modification of the thermal gradient only affects the vicinity of the pellet-cladding interface. Nevertheless, it cannot be neglect because it is exactly at this location that the SCC (known as a localized phenomenon), happens.

This impact was taken into account in a new version of NUAGES [1.2]. Modelling considers a thermal conductivity of 15 W/m/K in the cladding, and 1.5 W/m/K in the zirconia layer (so a ratio of 10 between the two materials).

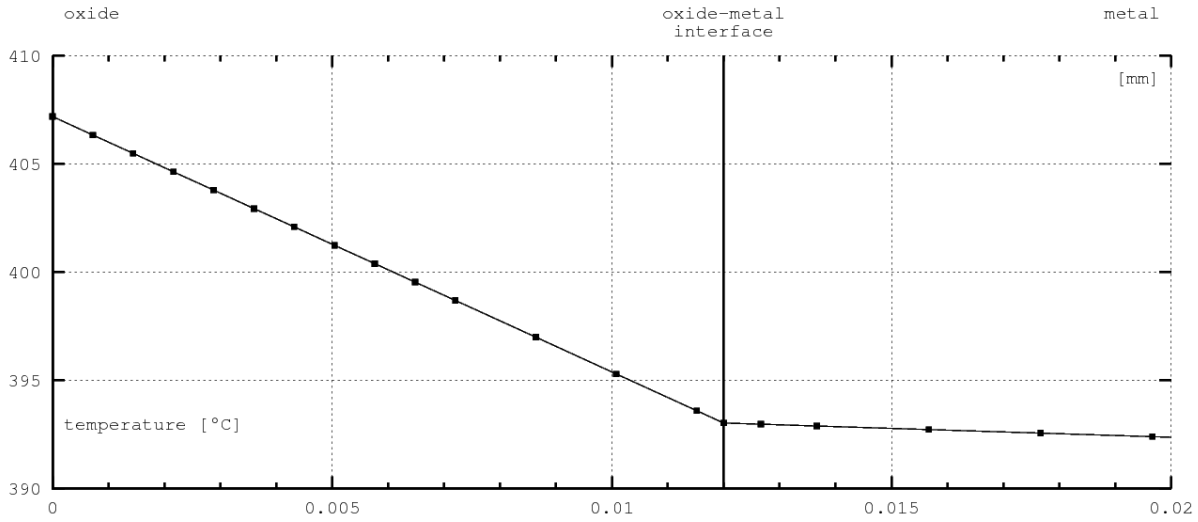


Figure 4 – Example of a thermal gradient, in a rod submitted to a power transient, considering a 12 μm thickness of zirconia at the inner surface of the cladding [16].

Figure 4 shows an example of thermal gradient considered in NUAGES v1.2, for a rod and a zirconia layer thickness that favor a huge thermal impact. The temperature drop through the zirconia layer is in this case about 14°C. We assume in the same time a systematic presence of crack in the zirconia layer. The temperature to be considered for chemisorption is then the temperature at the oxide-metal interface. This temperature is slightly lower than the temperature at the zirconia layer surface. The temperature drop in the zirconia layer is nevertheless low to explain a protective effect in absence of crack initiation.

From a chemical point of view, the presence of zirconia results from an oxygen flow from the fuel oxide toward the cladding. The oxygen is taken in the periphery of the pellet. Logically, the oxygen potential decreases. As a consequence, the chemical balances are modified. In other words, for a given element (iodine for example) the molar quantity is not modified, but chemical forms proportions between I(g) I_2 (g) or CsI (g) are modified.

Such effect is directly taken into account in NUAGES, in the sequence between EDMOND and ANGE [Fig 1]. The only constrain is to start calculations with a positive zirconia thickness and to provide an adequate time increment in the dataset file for the zirconia growth modeling [5].

4 CALCULATIONS DEALING WITH THE PROTECTIVE ROLE OF ZIRCONIA ON SCC

Figure 5 presents the zirconia layer thickness that has been measured in fuel rods:

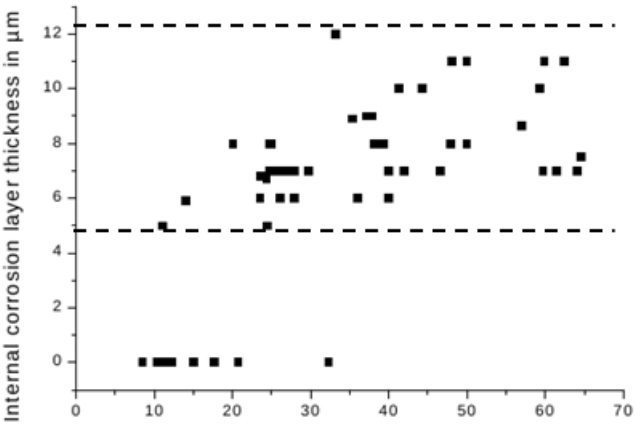


Figure 5 – Zirconia layer thickness, in [17,18]

As far as we know, there are no observations below 5 µm. The growth starts with localized zirconia islands, between which a continuous layer progressively develops. Studies let expect that this initiation phase is very quick, as soon as very punctual hard contacts happens between the pellet and the cladding. It is also observed that whatever the burn-up is, the zirconia layer growth is limited to about twelve micrometers. If we analyze more in details, for high burn-up the inner surface of the cladding penetrates HBS specific structure in the pellet rim, showing a wrinkled aspect. This phenomena strongly increases the specific surface between the cladding and the ceramic, making it difficult to measure the zirconia layer thickness [5].

For calculations, we choose to model the evolution of the observed thicknesses from 5 µm to 12 µm, with a 1 µm step. For demonstration, we used again the two reference calculations R_1 and R_2 rod, previously studied when validating SCC modeling scenario without taking into account any zirconia layer [8].

R_1 is a UO_2-Zy4 fuel rod irradiated during 5 cycles, to an average burn-up of 60 GWj/t (base irradiation)

R_2 is a power ramped segment taken from a fuel rod pre-irradiated during 3 cycles to an average burn-up of 45.3 GWd/t. This R_2 segment was conditioned at 195W/cm during 16h40. The conditioning rise was of 10 W/cm/mn. The power of the high level was of 466 W/cm. The average power was 470 W/cm. The power increase rate was of 100W/cm/mn. These conditions correspond to a severe power transient. Thus, this rod failed after 4mn04s. The holding time was of 1h34mn.

Each simulation was performed twice:

- Assuming 100 % fission gas release (conservative approach of SCC scenario),
- Providing a best-estimate evaluation of the fission gas release history (CYRANO3 calculations).

Finally, considering different zirconia layer thicknesses, we performed 36 simulations for

this sensitivity analyses.

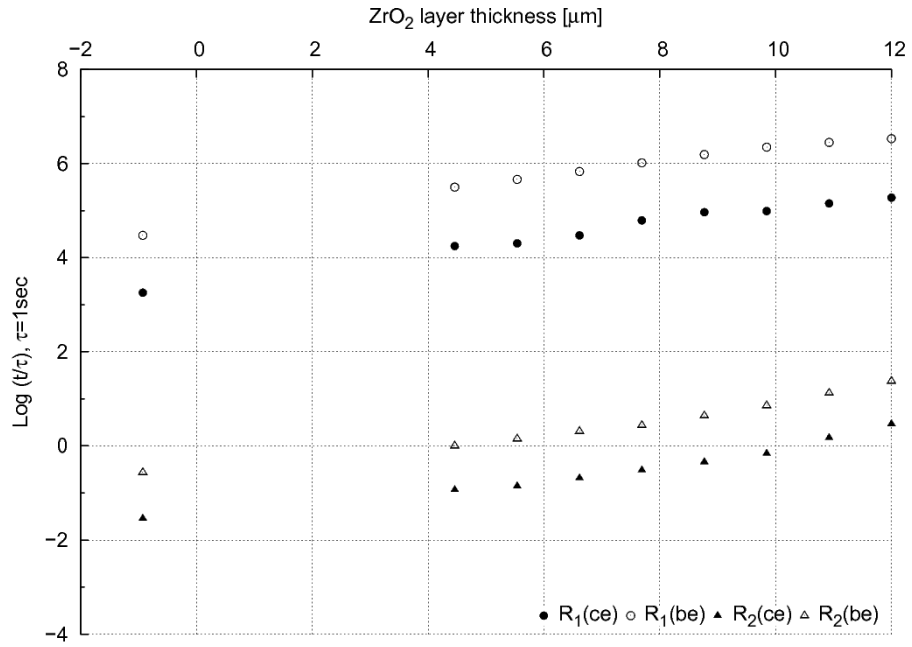


Figure 6 – Zirconia impact on chemisorption kinetics, NUAGES v1.2

The covering rate of a zirconium monocrystal by iodine (to 1 more or less 10^{-6}) is delayed in presence of a zirconia layer. The delay is proportional to the layer thickness [Fig 6].

There is a linear relationship between zirconia layer thickness and the logarithm of delayed impact. Comparing the four cases here above, the slope of the linear relationship seems to be the same. Then, we manage to draw the results considering: $Log(t/\tau) - Log(t/\tau)_0$, $Log(t/\tau)_0$ standing for $Log(t/\tau)$ in case of a zero layer thickness [Fig. 7].

The linear relationship that quantifies the delayed impact of the zirconia layer on chemisorption can be then written:

$$Log(t/\tau) = Log(t/\tau)_0 + 0.167 e_{ZrO_2} - 0.044$$

with e_{ZrO_2} standing for zirconia thickness [µm] and τ=1sec.

The regression coefficient gives 0.897. It is quite satisfying. All the simulations points are quite uniformly distributed on each side of the regression line, at a maximum distance of twice the standard deviation. From a quantitative point of view, for a zirconia layer of 6 µm, the chemical reaction step is multiplied by ten. So far, this modelling confirmed the protective role of the zircon layer against SCC [Fig. 7].

We showed previously that the temperature drop in the zirconia layer was not enough to explain its protective effect. The modelling presented here suggests that the protective effect has a chemical explanation. Indeed, the decrease of the surrounding oxygen potential induces a decrease of the Iodine monoatomic gaseous potential, which is the corrosive specie retained

in SCC scenarios. We have now to check that the SCC modelling scenario remains valid.

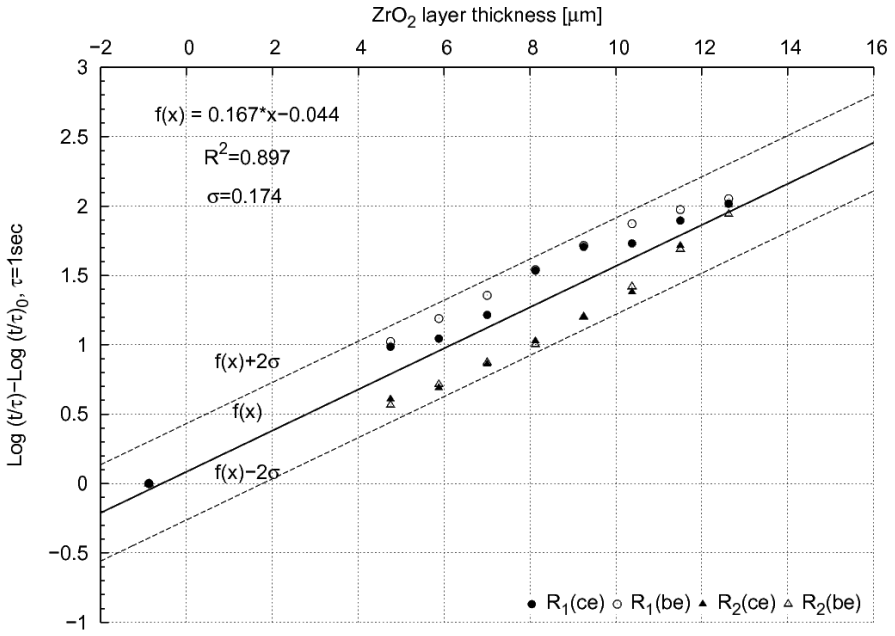


Figure 7 – Quantification of zirconia impact on chemisorption kinetics, NUAGES v1.2

5 VALIDATION OF SCC SCENARIO ASSUMPTIONS

The first assumption considers an upper bound and a lower bound to take into account, with very fast computing times, the fission gas release. We can observe on figure 6 that this two bounds do not prevent the differentiation of the high burn-up rod (R_1) and the rod submitted to a power transient (R_2). Whatever the assumption on fission gas released is, the chemisorption is always faster in the case of the power transient, with three orders of magnitude. As a consequence, the extension of SCC scenario accounting for the the zirconia layer does not modify assumptions relative to fission gas release. They are adequate to answer the industrial specification in term of simulation.

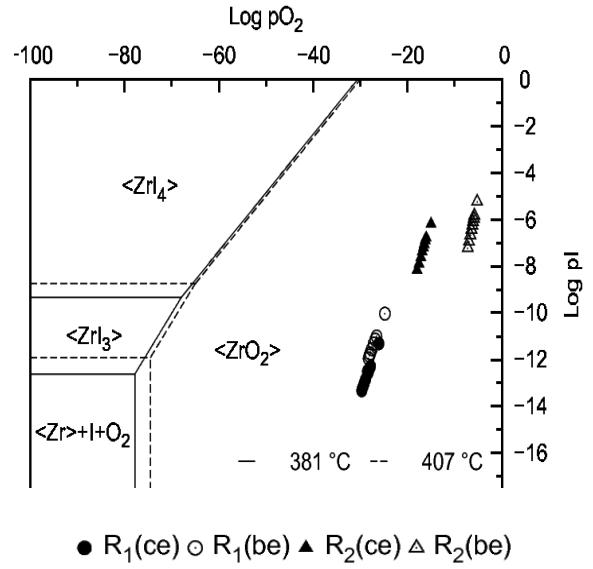


Figure 8 Zr-I-O binary diagram, NUAGES v1.2 built with equations available in [12]

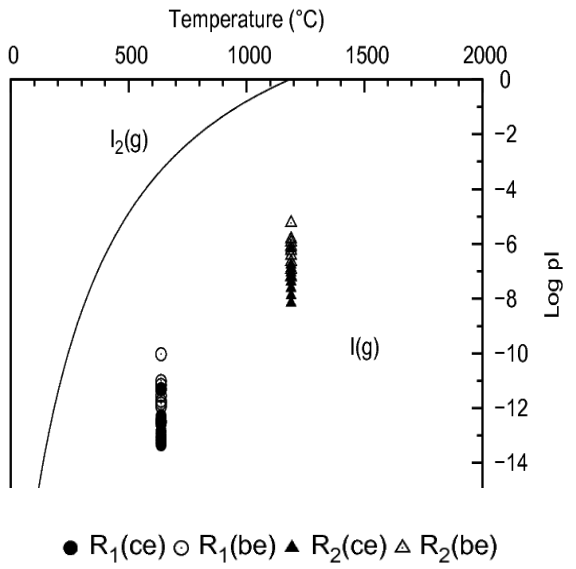


Figure 9 Vapor pressure of gaseous iodine, NUAGES v1.2, built with equations available in [19]

When equilibrium is established, the thermodynamic diagram predicts the development of zirconia at the inner surface of the cladding [Fig. 8]. This is in accordance with the in pile observations. On another hand the presence of a crack crossing the zirconia layer and reaching the zirconium interface, can lead to a steady-state thermodynamically compatible to initiate Iodine corrosion. In this case, even if the presence of the zirconia layer lowers the oxygen potential in the gap, it has no effect at the zirconia-zircaloy interface where ZrI_x species are formed. Therefore, the hypothesis of a crack mechanical initiation stays a necessary condition for the SCC modeling scenario.

The third assumption considers that diatomic gaseous iodine can be neglected compared with the monoatomic form. Figure 9 shows the proportions of monoatomic iodine vs diatomic iodine. Nevertheless gases are produced inside the pellet. Thus, we had to postpone on the diagram the partial pressures of iodine estimated by calculations, according to the average temperature of the pellet. Because it lowers the potential of gaseous monoatomic iodine for an equivalent temperature, the presence of zirconia does not call into question the predominance of this form against the gaseous diatomic iodine. On the contrary, it tends to strengthen it. The assumption considering the monoatomic Iodine form mainly responsible of the corrosive attack remains valid [Fig. 9].

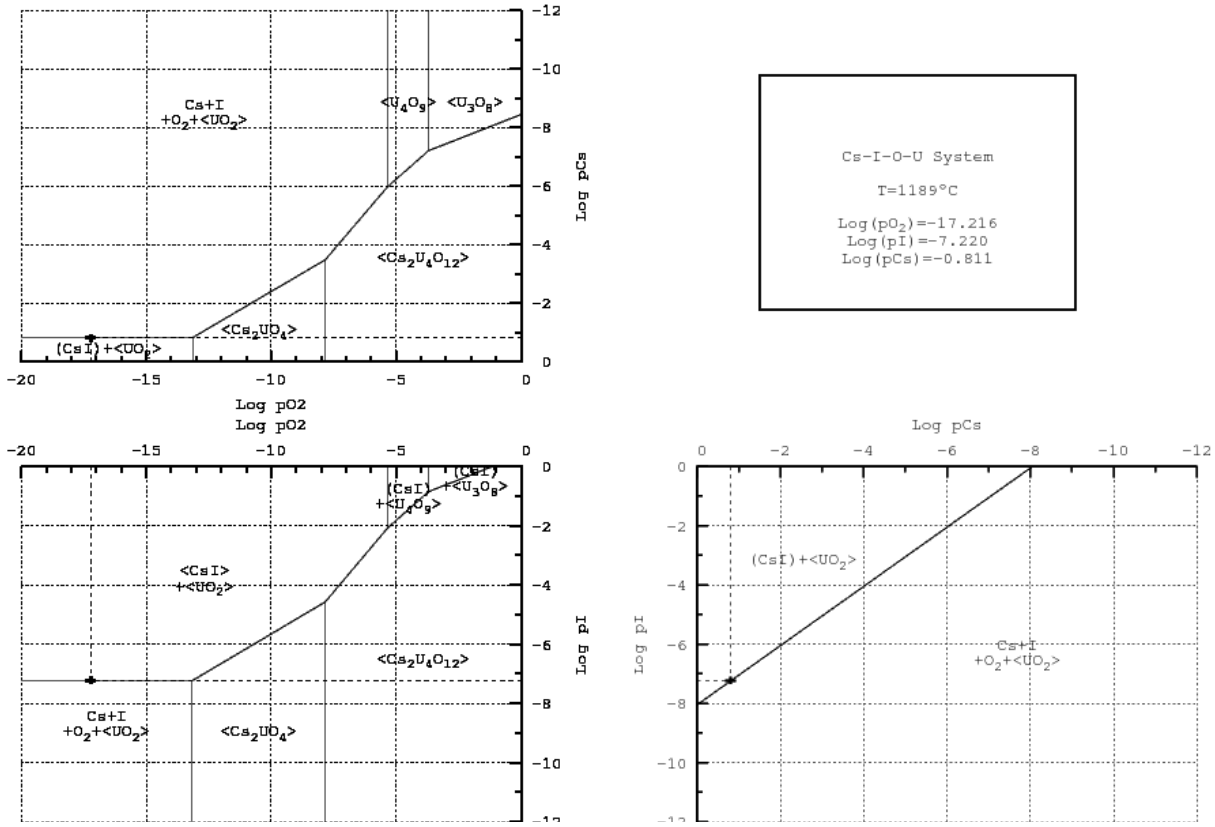


Figure 10 : Ternary Cs-I-O-U diagram, R2 calculation with a 12 µm zirconia layer thickness, NUAGES v1.2

Finally, the fourth assumption considers the radiolysis phenomenon in pile as marginal. CsI(g) is not a corrosive gas for the cladding. However in case of dissociation, it would provide monatomic or diatomic iodine. CsI(g) is produced in large quantities within the pellet. Moreover, according to the thermodynamics diagram balance, the potential of cesium is widely higher than that those of iodine and oxygen. We also notice that the calculations are always situated in the balance between I (g) and CsI (g) [Fig. 10].

At first glance, the zirconia layer does not carry new information about the hypothetical role of radiolysis. But, we also can consider another point of view. Let's consider that radiolysis plays a significant role in pile. It would mean a large scale dissociation of CsI (g). As a consequence the potential of gaseous iodine species would be the same whatever is the zirconia layer thickness for a given burn-up. Calculations would not be able anymore to predict the protective role of the zirconia on SCC observed experimentally.

In conclusions coupling a protective effect of the zirconia layer to the CsC model, remains consistent with all assumptions underlying the industrial CSC model. It even strengthened the original model.

6 ADEQUACY BETWEEN SIMULATIONS AND EXPERIMENTAL OBSERVATIONS

We suggest here to confront experimental observations which describe the protective role of the zirconia towards SCC with our scenario and the results of simulations.

1. Presence of oxygen in small quantity in the middle accelerates SCC The calculation consolidates this observation, the more the quantity of oxygen decreases the more the reaction mechanism of chemisorption of iodine is delayed [Fig. 6].

2. Presence of oxygen in high quantity in the test environment inhibits the reaction with iodine. In our calculations, the presence of oxygen is always unimportant compared with the quantity of iodine. However, considering a hypothetical case in which the quantity of oxygen would be equivalent to the initial quantity of iodine [Fig. 3], the oxygen would coat zirconium. inhibiting reaction with iodine. The chemisorption model thus accredits the experimental observations.

3. The presence of a pre-oxidation layer protects the metal if the thickness is self-important (upper than 200 nm). The one naturally formed in ambient atmosphere (10-15 nm), has no protective effect on the metal. Our simulations allowed to quantify the impact of zirconia towards SCC :

$$\text{if } e_{ZrO_2} = 0.2\ \mu\text{m, then } Log\left(t/\tau\right) - Log\left(t/\tau\right)_0 = -0.0116 \text{ and } Log\left(t/\tau\right) = 0.974\ Log\left(t/\tau\right)_0$$

So, a 0.2 μm zirconia layer would delay chemisorption reaction mechanism of less than 3%. Considering this result, we can easily deduce that native zirconia layer has no protective effect.

4. The origin of SCC cracking is not chemical but mechanical. This point has been previously discussed. The SCC scenario would be able to predict SCC chemical risk without this assumption.

5. The presence of oxygen localizes the phenomenon of SCC. As the numerical component is 1D axisymmetric, we cannot give any comment on this observation.

6. The oxygen can lead to a passivation of the metal if the quantity of oxygen is important enough to induce formation of an oxide layer faster than the iodine attack. The proportion of oxygen is always higher in the gap compared to iodine, favoring the zirconia formation kinetic [Fig. 3]. The final balance is also in favor of zirconia formation [Fig. 5]. The theory underlying the scenario of thermochemistry confirms this observation.

Finally, the current scenario confirms five experimental observations over six. The sixth is not contradicted today, but not accessible with our 1D axisymmetric calculations.

7 CONCLUSIONS AND PERSPECTIVES

Thanks to the numerical component NUAGES, a sensitivity analysis has been performed allowing to confirm the protective effect of the zirconia layer against SCC.

- This protective effect looks to be not mechanical or thermal but only chemical.
- It can be quantified as follow:

$$Log\left(t/\tau\right) = Log\left(t/\tau\right)_0 + 0.167 e_{ZrO_2} - 0.044$$

with $Log\left(t/\tau\right)_0$ standing for $Log\left(t/\tau\right)$ in the case of a zero thickness zirconia layer.

e_{ZrO_2} represents the zirconia layer thickness [μm] and τ=1sec.

After coupling the zirconia layer model presented here and the SCC standard model, all hypotheses underlying the industrial modeling scenario remains valid, even strengthened.

The current modeling scenario agrees with five experimental observations over six. The sixth is not contradicted, but simply inaccessible today in our 1D axisymmetric approach.

An interesting perspective would be to enrich the present study with other simulations, in order to:

- Confirm the generic character of an equation to represent the impact of the zirconia against the SCC risk
- Refine when necessary the parameters of the equation modelling the delayed effect of zirconia on chemisorption kinetic.

8 REFERENCES

[1] F. Laugier, C. Garzenne and E. Cabrol, "Very fast isotopic inventory and decay heat calculations in PWR spent nuclear fuel for industrial application", *Proc. Int. Top. Meet. on Mathematics and Computation, Supercomputing, Reactor Physics and Nuclear Biological Applications (M&C 2005),* Avignon, France, pp. 12–15, (2005).

[2] P. Garcia P., J-P. Piron and D. Baron, "A Model for the Oxygen Potential of Oxide Fuels at High Burn-up", *IAEA TCM on Water Reactor fuel Modelling at high Burn-up and its Experimental support*, Windermere (UK) 19-23 September (1994).

[3] Barnel N. "Spécification informatique du composant Nuages v1.0", Note EDF HT-25-2013-01899-FR, (2013).

[4] Palermo D., Grondein C., "Méthodologie de Migration Codes Fortran en Composants C++ - Yantra-Tachnologies", ISBN 2-9527-3440-2, (2006).

[5] Minne J.-B., "Contribution à la modélisation du couplage mécanique/chimique de l'évolution de l'interface pastille-gaine sous irradiation", Thèse de l'Université de

Bourgogne, (2013).

[6] Barnel N., “Mise en évidence du rôle protecteur de la zircone sur la CsC-I avec le composant numérique NUAGES v1.2”, Note EDF HT-25-2014-04496-FR, (2014).

[7] Barnel N., “Etude d’un cas concret de thermochimie avec Nuages : la CsC-I”, Note EDF HT-25-2013-00239-FR, (2013).

[8] Barnel N., Palermo D – “Evaluation of iodine corrosion potential, by means of an optimized numerical component coupling neutron physics and chemistry. Application to a nuclear fuel rod.”, *VI International Conference on Computational Methods for Coupled Problems in Science and Engineering , Coupled Problems*

[9] Yang T.T., Tsai C.H., “On the susceptibility to Stress-Corrosion-Cracking of Zircaloy in an iodine containing environment”, *J. Nucl. Mater*., **166**, 1989, pp. 252-264, 1989.

[10] Fregonese M. Lefevre F., Lemaignan C. Magnin T., “Influence of recoil-implanted and thermally released iodine on I-SCC of Zircaloy-4 in PCI conditions: chemical aspects”, *Journal Of Nuclear Materials*, **265**, pp 245-254, (1999).

[11] Balooch M., Olander D.R., “Kinetic study of zirconium-iodine reaction by modulated molecular beam mass spectroscopy”, *J. Electrochem. Soc.* **130**, n°1, pp 151-157, (1982).

[12] Walle E., “Evaluation du risque de corrosion sous contrainte par l’iode des gaines de combustible en conditions d’entreposage à sec”, thèse de l’université des Sciences et Technologie de Lille, (2002).

[13] Krishnan G., Wood B., Cubiciotti D., “Auger electron spectroscopy of the chemisorption of iodine on zirconium”, *J. Electrochem. Soc.*, **127**, pp 2738-2742, (1980).

[14] Cox B., “Pellet-clad interacton (PCI) failures of zirconium alloy fuel cladding”, a review, *J. Nucl. Mater*., **172**, pp 249-292, (1990).

[15] Bailly H., Ménessier D., Prunier C., “Le combustible nucléaire des réacteurs à eau sous pression et des réacteurs à neutrons rapides, conception et comportement”, Eyrolles, CEA Série Synthèses, (1996).

[16] Jannot Y., “Transferts thermiques”, Cours de l’Ecole des Mines de Nancy, (2012).

[17] Desgranges L., “Internal corrosion layer in PWR fuel”, Proceedings of the seminar on the Thermal Performance of High Burn-Up, OCDE IAEA, Aix en Provence, 187-196,

(1998).

[18] Lozano N., Desgranges L., Aymes D. Niepce J.C., "High magnification SEM observations for two types of granularity in a high Burn-up PWR fuel RIM", *Journal of Nuclear Materials*, **257**, 78-887, (1998).

[19] Weast R.C., "Handbook of chemistry and physics – 56th edition", *CRC press*, INC, (1976).

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis (Eds)

COMPLETE FORMULATION OF THE SUBLOADING SURFACE MODEL

KOICHI HASHIGUCHI*

* Joining and Welding Research Institute, Osaka University
Mihogaoka 11-1, Ibaragi-shi, Osaka, 567-0047, Japan

Key words: Cyclic plasticity, Multiplicative elastoplasticity, Subloading surface model

Abstract. The subloading surface model is endowed the noticeable ability to describe the wide classes of irreversible mechanical behavior [1]. However, the past formulations of the subloading surface model have contained several inexact equations, which have been modified repeatedly after the concept of the subloading surface was proposed in 1977 [2]. The exact formulation is presented first in this article for the hypoelastic-based plasticity, which enjoys the distinguished superiority in the both aspects of the description of material behavior in high accuracy and of the numerical calculation in high efficiency.

1 INTRODUCTION

The subloading surface is based on the quite natural physical insight that the plastic strain rate develops continually as the stress approaches the yield surface. The formulation of the subloading surface model for the elastoplastic deformation has been modified repeatedly and developed from the initial ones [2,3,4,5]. Further, it has been applied to the descriptions of wide classes of irreversible mechanical phenomena, i.e. the monotonic and cyclic loadings of metals and soils, the viscoplastic deformation behavior 0, the damage behavior, the phase transformation behavior, the friction behavior and the crystal plasticity [3]. On the other hand, the other unconventional elastoplasticity models, e.g. the multi surface model [6], the two surface model [7] and the superposed kinematic hardening model [8], which does not assume that the inside of yield surface is an elastic domain but assume the existence of the small purely-elastic domain, have been formulated only for the description of cyclic loading behavior of metals. The extensive applicability of the subloading surface model to wide classes of irreversible mechanical behavior is based on the noticeable advantages such that it does not require the yield judgment on whether or not the stress reaches the yield surface and it is furnished with the controlling function such that the stress is pulled-back automatically to the yield surface when it goes out from the yield surface in numerical calculation due to finite incremental steps. However, the translation rules of the anisotropic hardening variable (back-stress) and the elastic-core, i.e. similarity-center of the yield and the subloading surfaces and the accurate expression of the Masing rule have been modified repeatedly but they have been formulated in the inexact forms in the past [1,3,9]. Now, the exact formulation of the subloading surface model will be attained in this article, passing near a half century after the

concept of the subloading surface was proposed in 1977.

2 STRAIN RATE AND HYPOELASTIC RELATION

In the framework of the hypoelastic-based plasticity, incorporating velocity gradient $\boldsymbol{l}=\partial\mathbf{v}/\partial\mathbf{x}$, the strain rate $\mathbf{d}\equiv\mathrm{sym}[\boldsymbol{l}]$ and the continuum spin $\mathbf{w}\equiv\mathrm{ant}[\boldsymbol{l}]$ are used as the measures of the rates of deformation and rotation, respectively, where $\mathrm{sym}[\ ]=(\mathbf{t}+\mathbf{t}^T)/2$ and $\mathrm{ant}[\ ]=(\mathbf{t}-\mathbf{t}^T)/2$ stand for the symmetric and the anti-symmetric parts, respectively, for an arbitrary second-order tensor $\mathbf{t}$ and $(\ )^T$ for the transpose. Let the strain rate $\mathbf{d}$ be decomposed additively into elastic strain rate $\mathbf{d}^e$ and the plastic strain rate $\mathbf{d}^p$ as

$$\mathbf{d}=\mathbf{d}^e+\mathbf{d}^p \tag{1}$$

First, assume that the elastic strain rate is given by the hypoelastic relation on the premise that the elastic deformation is small compared with the plastic deformation:

$$\mathbf{d}^e=\mathbf{E}^{-1}:\mathring{\boldsymbol{\sigma}} \tag{2}$$

where $\boldsymbol{\sigma}$ is the Cauchy stress. $\mathbf{E}$ is the elastic modulus tensor which is given by

$$E_{ijkl}=K\delta_{ij}\delta_{kl}+2G\left\{\frac{1}{2}(\delta_{ik}\delta_{jl}+\delta_{il}\delta_{jk})-\frac{1}{3}\delta_{ij}\delta_{kl}\right\} \tag{3}$$

for the Hooke's law, where K and G are the bulk modulus and the shear modulus, respectively. $(^\circ)$ stands for the proper objective corotational rate, i.e.

$$\mathring{\mathbf{t}}\equiv\dot{\mathbf{t}}-\boldsymbol{\omega}\mathbf{t}+\mathbf{t}\boldsymbol{\omega} \tag{4}$$

for an arbitrary second order tensor $\mathbf{t}$, where $\boldsymbol{\omega}$ is the spin of substructure of material. The continuum spin $\mathbf{w}$ may be used for $\boldsymbol{\omega}$ up to a moderate deformation. Needless to say, the corotational rate is also used for tensor-valued internal variables to describe anisotropy.

The continuum spin $\mathbf{w}$ is additively decomposed into the elastic spin $\mathbf{w}^e$ and the *plastic spin* $\mathbf{w}^p$, i.e.

$$\mathbf{w}=\mathbf{w}^e+\mathbf{w}^p \tag{5}$$

Here, we adopt the *isoclinic concept* insisting that the rigid-body spin is involved in the elastic spin $\mathbf{w}^e$ under the postulate that the spin of the substructure is induced by the rotations due to the rigid-body motion and the elastic distortion. Then, the spin of substructure $\boldsymbol{\omega}$ is given by $\mathbf{w}^e$, i.e.

$$\boldsymbol{\omega}=\mathbf{w}^e=\mathbf{w}-\mathbf{w}^p \tag{6}$$

3 REFINEMENT FOR FORMULATION OF PLASTIC STRAIN RATE

The plastic strain rate will be formulated based on the concept of subloading surface in the following.

3.1 Yield surface

The normal-yield surface with the isotropic and the kinematic hardenings is described as

$$f(\hat{\boldsymbol{\sigma}})=F(H) \tag{7}$$

where H is the isotropic hardening variable and

$$\hat{\boldsymbol{\sigma}} \equiv \boldsymbol{\sigma} - \boldsymbol{\alpha} \tag{8}$$

$\boldsymbol{\alpha}$ $(\mathrm{tr}\boldsymbol{\alpha} = 0)$ being the kinematic hardening variable, i.e. back stress. The rates of these internal variables can be described by

$$\left.\begin{aligned} \overset{\bullet}{H} &= f_{hd}(\boldsymbol{\sigma}, F; \mathbf{d}^p) = f_{hd}(\boldsymbol{\sigma}, F; \mathbf{d}^p / \|\mathbf{d}^p\|)\|\mathbf{d}^p\| \\ \overset{\circ}{\boldsymbol{\alpha}} &= \mathbf{f}_{kd}(\boldsymbol{\sigma}, F, \boldsymbol{\alpha}; \mathbf{d}^{p\prime}) = \mathbf{f}_{kd}(\boldsymbol{\sigma}, F, \boldsymbol{\alpha}; \mathbf{d}^{p\prime}/\|\mathbf{d}^p\|)\|\mathbf{d}^p\| \end{aligned}\right\} \tag{9}$$

since they are homogeneous functions of $\mathbf{d}^p$ in degree-one since they are induced only in the plastic loading process $\mathbf{d}^p \neq \mathbf{O}$ and the first-order time-differential quantities, where $\|\ \|$ designates the magnitude and $(\ )'$ the deviatoric part, . Here, assume that $f(\hat{\boldsymbol{\sigma}})$ is the homogeneous function of $\hat{\boldsymbol{\sigma}}$ in degree-one and thus it follows by the Euler's theorem that

$$\frac{\partial f(\hat{\boldsymbol{\sigma}})}{\partial \hat{\boldsymbol{\sigma}}} : \hat{\boldsymbol{\sigma}} = f(\hat{\boldsymbol{\sigma}}) \tag{10}$$

3.2 Concept of subloading surface model

The plastic strain rate is induced explicitly when the stress lies on the yield surface. Here, in facts, the plastic strain rate is induced not suddenly at the moment when the stress reaches the yield surface but it is induced gradually as the stress approaches the yield surface. The accurate description of the plastic strain rate induced by the rate of stress inside the yield surface is required in order to predict the cyclic loading behavior of materials, although it has been ignored in the conventional elastoplasticity. Then, let the following postulate be incorporated, which is the basic concept of the subloading surface model [1,2,3,4,5].

Fundamental postulate of elastoplasticity (Subloading surface concept): *The plastic strain rate is induced when the stress approaches the yield surface but only the elastic strain rate is induced when the stress moves towards of the inside the yield surface* as shown in **Fig. 1**, while the stress rate causes the elastic strain rate inevitably. In other words, **the stress approaches the yield surface when a plastic strain rate is induced but it moves towards the inside of the yield surface when only an elastic strain rate is induced.**

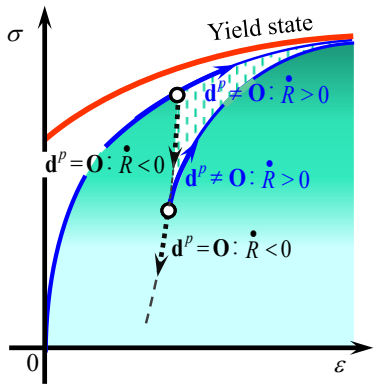


Fig. 1 Background of the subloading surface concept.

Then, it is first required to incorporate the measure which describes the approaching degree of the stress to the yield surface, renamed the *normal-yield surface*, in order to formulate the plastic strain rate based on the above-mentioned subloading surface concept.

Then, let the following ***subloading surface*** which passes through the current stress and keeps the similar shape and orientation to the normal-yield surface be introduced, which plays the general measure of approaching degree of the stress to the normal-yield surface (see **Fig. 2**).

$$f(\bar{\boldsymbol{\sigma}}) = RF(H) \tag{11}$$

where $R\,(0 \le R \le 1)$ is the ratio of the size of the subloading surface to that of the normal-yield surface and referred to as the *normal-yield ratio* which plays the role for the general measure to describe the approaching degree to the normal-yield surface.

$$\left.\begin{aligned} &\hat{\mathbf{c}} \equiv \mathbf{c} - \boldsymbol{\alpha} = (\mathbf{c} - \bar{\boldsymbol{\alpha}})/R \\ &\bar{\boldsymbol{\alpha}} \equiv \mathbf{c} - R\hat{\mathbf{c}} \\ &\bar{\boldsymbol{\sigma}} \equiv \boldsymbol{\sigma} - \bar{\boldsymbol{\alpha}}, \quad \tilde{\boldsymbol{\sigma}} \equiv \boldsymbol{\sigma} - \mathbf{c} \end{aligned}\right\} \tag{12}$$

leading to the expressions

$$\bar{\boldsymbol{\sigma}}\,(= \boldsymbol{\sigma} - \bar{\boldsymbol{\alpha}} = \boldsymbol{\sigma} - (\mathbf{c} - R\hat{\mathbf{c}})) = \tilde{\boldsymbol{\sigma}} + R\hat{\mathbf{c}} \tag{13}$$

$\mathbf{c}$ represents the center of similarity of the normal-yield and the subloading surfaces, i.e. the similarity-center, while let it be called the *elastic-core* since the most elastic deformation behavior is induced when the stress lies on it. $\bar{\boldsymbol{\alpha}}$ stands for the conjugate (similar) point in the subloading surface to the point $\boldsymbol{\alpha}$ in the normal-yield surface. All of the relations of variables in Eq. (12) hold by virtue of the similarity of the subloading surface to the normal-yield surface as known from Fig. 2.

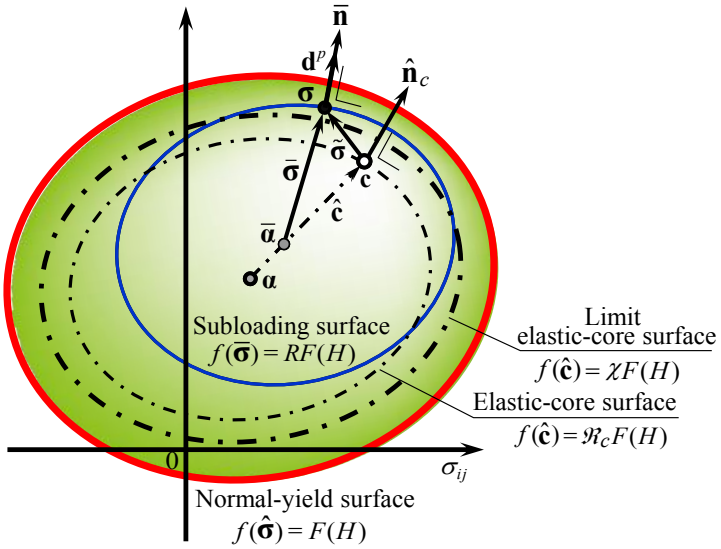


Fig. 2 Normal-yield, subloading and elastic-core surfaces.

3.3 Evolution rule of normal-yield ratio

Based on the afore-mentioned fundamental postulate of elastoplasticity, the rate of the

normal-yield ratio must satisfy the following conditions.

$$\dot{R}\begin{cases}\to\infty \text{ for } R=0\\ >0 \text{ for } R<1\\ =0 \text{ for } R=1\\ (<0 \text{ for } R>1)\end{cases} \quad \text{for } \mathbf{d}^p\neq\mathbf{O} \tag{14}$$

$$\dot{R}\begin{cases}=0 \text{ for } \mathbf{d}^e=\mathbf{O}\\ <0 \text{ for } \mathbf{d}^e\neq\mathbf{O}\end{cases} \quad \text{for } \mathbf{d}^p=\mathbf{O} \tag{15}$$

Taking account of the requirement in Eq. (14), let the evolution rule of the normal-yield ratio in the plastic deformation process be formulated as follows:

$$\dot{R}=U(R)\|\mathbf{d}^p\| \quad \text{for } \mathbf{d}^p\neq\mathbf{O} \tag{16}$$

where $U(R)$ is the monotonically-increasing function of R fulfilling the conditions (**Fig. 3**).

$$U(R)\begin{cases}\to+\infty \text{ for } R=0 \text{ (quasi-elastic state)}\\ >0 \text{ for } R<1 \text{ (subyield state)}\\ =0 \text{ for } R=1 \text{ (normal-yield state)}\\ <0 \text{ for } R>1 \text{ (over normal-yield state)}\end{cases} \tag{17}$$

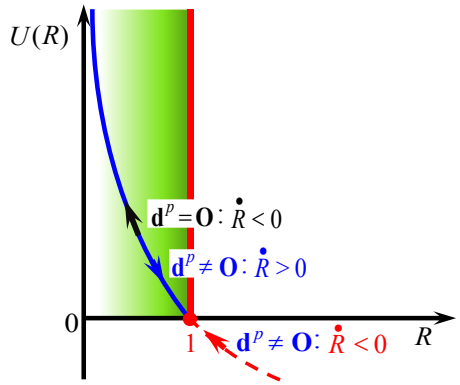


Fig. 3 Function $U(R)$ in the evolution rule of normal-yield ratio.

The function $U(R)$ is given by the following equation, where u is the material parameter.

$$U(R)=u\cot\left(\frac{\pi}{2}R\right) \tag{18}$$

Equation (17) should be extended to the following equation for metals in which plastic strain rate is not induced until the normal-yield ratio R reaches a certain value R_e (<1) which is the material parameter.

$$U(R)\begin{cases}\to+\infty \text{ for } 0\le R\le R_e \text{ (quasi-elastic state)}\\ >0 \text{ for } R_e<R<1 \text{ (subyield state)}\\ =0 \text{ for } R=1 \text{ (normal-yield state)}\\ <0 \text{ for } R>1 \text{ (over normal-yield state)}\end{cases} \tag{19}$$

Eqs. (18) is modified for Eq. (19) as follows:

$$U(R)=u\cot\left(\frac{\pi}{2}\frac{\langle R-R_e\rangle}{1-R_e}\right) \tag{20}$$

where $\langle\ \rangle$ is the Macauley's bracket. If u is fixed to be constant, Eq. (16) with Eq. (20) can be integrated analytically as

$$R=\frac{2}{\pi}(1-R_e)\cos^{-1}\left[\cos\left(\frac{\pi}{2}\frac{R_0-R_e}{1-R_e}\right)\exp\left(-u\frac{\pi}{2}\frac{\varepsilon^p-\varepsilon_0^p}{1-R_e}\right)\right]+R_e \quad \text{for } R_0\geq R_e \tag{21}$$

under the initial condition $\varepsilon^p=\varepsilon_0^p: R=R_0$, where $\varepsilon^p\equiv\int\|\mathbf{d}^p\|dt$ (t: time), whilst one must set $R_0=R_e$ for $R_0<R_e$. The use of the analytical integration in Eqs. (21) contributes to the enhancement of the numerical calculation in the return-mapping projection. However, it spoils the automatic controlling function to attract the stress to the normal-yield surface, which is inevitable in the numerical calculation in the forward-Euler method.

The subloading surface model possesses the following distinguished abilities.

1) Smooth transition from elastic to plastic state is described, fulfilling always the smoothness condition 0 as shown schematically in **Fig. 4**.

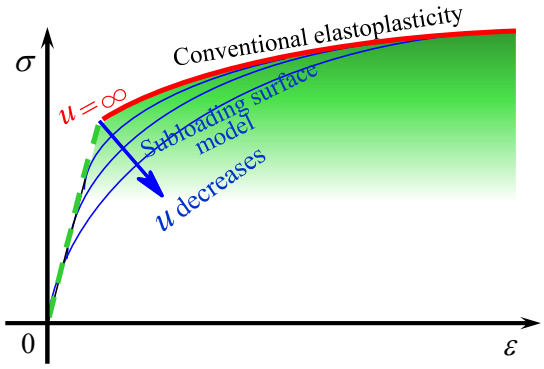


Fig. 4 Influence of material parameter u on curvature of stress vs. strain curve.

Then, we don't need suffer from the determination of an offset-value of strain for yielding. On the other hand, the determination of an offset value is required in the conventional elastoplasticity with an abrupt elastic-plastic transition, although it is accompanied with an arbitrariness. Smoother stress-strain curve is described for smaller value of the material parameter u.

2) Plastic strain rate can be described even for the loadings under a low stress level and a small stress amplitude since a purely-elastic domain is not assumed.

3) The yield-judgment whether or not the stress reaches the yield surface is unnecessary for the loading criterion, since the plastic strain rate develops continuously as the stress approaches the normal-yield surface.

4) The stress is automatically attracted to the normal-yield surface in the plastic loading process. Therefore, in the numerical calculation due to the forward-Euler method, it is pulled back automatically to the normal-yield surface when it goes out from the normal-yield surface in numerical calculation because of $\dot{R}<0$ for $R>1$ from Eq. (16) with Eq.

$(17)_4$ or $(19)_4$ as seen in **Fig. 5**.

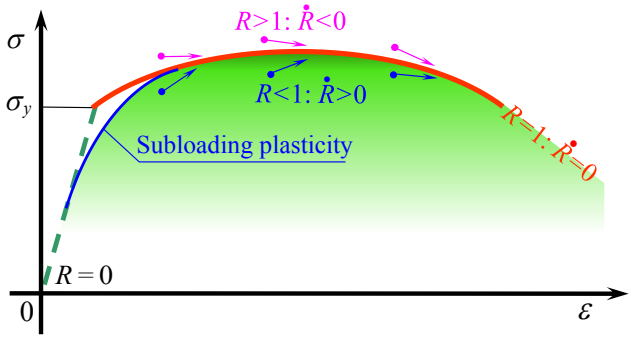


Fig. 5 Stress is automatically attracted to normal-yield surface in plastic loading process.

Consequently, the realistic description of mechanical behavior is attained by virtue of 1) and 2) and further the efficient numerical calculation is realized by virtue of 3) and 4). Thus, the subloading surface model possesses the noticeable advantages in both aspects of the physical description and the numerical calculation.

3.4 Associated flow rule for subloading surface

The associated flow rule for the subloading surface is adopted:

$$\mathbf{d}^p = \dot{\lambda}\bar{\mathbf{n}} \quad (\dot{\lambda}>0) \quad (\|\mathbf{d}^p\| = \dot{\lambda}) \tag{22}$$

where 10

$$\bar{\mathbf{n}} \equiv \frac{\partial f(\bar{\boldsymbol{\sigma}})}{\partial \boldsymbol{\sigma}} \Big/ \Big\| \frac{\partial f(\bar{\boldsymbol{\sigma}})}{\partial \boldsymbol{\sigma}} \Big\| = \frac{\partial f(\bar{\boldsymbol{\sigma}})}{\partial \bar{\boldsymbol{\sigma}}} \Big/ \Big\| \frac{\partial f(\bar{\boldsymbol{\sigma}})}{\partial \bar{\boldsymbol{\sigma}}} \Big\| \quad (\|\bar{\mathbf{n}}\|=1) \tag{23}$$

designating the magnitude and the direction of plastic strain rate by $\dot{\lambda}$ and $\bar{\mathbf{n}}$, respectively. Then, $\dot{H}$ and $\mathring{\boldsymbol{\alpha}}$ in Eq. (9) are expressed as follows:

$$\dot{H} = \dot{\lambda} f_{hn}(\boldsymbol{\sigma}, F; \bar{\mathbf{n}}), \quad \mathring{\boldsymbol{\alpha}} = \dot{\lambda}\mathbf{f}_{kn}(\boldsymbol{\sigma}, F, \boldsymbol{\alpha}; \bar{\mathbf{n}}) \tag{24}$$

where f_{hn} and $\mathbf{f}_{\alpha n}$ are the homogeneous functions of $\bar{\mathbf{n}}$ in degree-one.

3.5 Unified nonlinear kinematic hardening rule

The nonlinear kinematic-hardening rule of Armstrong and Frederick [10] is inapplicable to the general anisotropic hardening material with the plastically-compressibility causing the rotational-hardening as known from the fact that the anisotropic hardening is not induced by the isotropic plastic deformation but is induced by the deviatoric plastic deformation. Then, let the generalized evolution rule of anisotropic hardening variable $\boldsymbol{\alpha}$ $(\mathrm{tr}\boldsymbol{\alpha}=\mathrm{tr}\mathring{\boldsymbol{\alpha}}=0)$ be given as follows:

$$\mathring{\boldsymbol{\alpha}} = c_k\Big(\mathbf{d}^{p\prime} - \frac{1}{b_k}\|\mathbf{d}^{p\prime}\|\boldsymbol{\alpha}\Big) = \dot{\lambda}\mathbf{f}_{kn}, \quad \mathbf{f}_{kn} = c_k\Big(\bar{\mathbf{n}}' - \frac{1}{b_k}\|\bar{\mathbf{n}}'\|\boldsymbol{\alpha}\Big) \tag{25}$$

where c_k is the material constant but b_k is the material function in general. Equation (25) is applicable not only to the kinematic hardening in metals but also to the rotational hardening, i.e. the rotation of yield surface in soils [1].

3.6 Translation rule of elastic-core

The most elastic deformation behavior is induced in the state that the stress lies on the similarity-center, i.e. $\boldsymbol{\sigma}=\mathbf{c}$ leading to $R=0$. Then, the similarity-center $\mathbf{c}$ in the mathematical sense is interpreted physically as the most elastic stress state so that let it be called the *elastic-core* or *elastic-center*. Here, note that the elastic-core $\mathbf{c}$ approaches the normal-yield surface, following the stress $\boldsymbol{\sigma}$ in the plastic loading process. However, from the physical point of view the elastic-core should not approach the normal-yield surface without limitation as known from the fact that the abrupt transition from the elastic to the plastic state is predicted if the elastic-core lies on the normal-yield surface. On the other hand, the small yield surface enclosing a purely-elastic region is allowed to contact with the yield surface in the cyclic kinematic-hardening models [6,7] predicting the abrupt elastic-plastic transition. In addition, from the mathematical point of view the subloading surface is not determined uniquely if the stress coincides with the similarity-center lying just on the normal-yield surface.

Now, let the following *elastic-core surface* be introduced, which always passes through the elastic-core $\mathbf{c}$ and keeps the similar shape and orientation to the normal-yield surface with respect to the kinematic-hardening variable $\boldsymbol{\alpha}$.

$$f(\hat{\mathbf{c}})=\mathscr{R}_c F(H), \text{ i.e. } \mathscr{R}_c=f(\hat{\mathbf{c}})/F(H) \tag{26}$$

where $\mathscr{R}_c$ designates the ratio of the size of the elastic-core surface to the normal-yield surface (see Fig. 2) so that let it be called the *elastic-core yield ratio*. Here, the elastic-core should not reach (lie on) the normal-yield surface as described above so that the elastic-core does not go over the following *limit elastic-core surface.*

$$f(\hat{\mathbf{c}})=\chi F(H) \tag{27}$$

where $\chi(<1)$ is material parameter and the following inequality must be satisfied.

$$f(\hat{\mathbf{c}})\le\chi F(H), \text{ i.e. } \mathscr{R}_c\le\chi \tag{28}$$

Let the translation rule of elastic-core be formulated as

$$\overset{\circ}{\mathbf{c}}=c\Big(\mathbf{d}^p-\frac{\mathscr{R}_c}{\chi}\|\mathbf{d}^p\|\hat{\mathbf{n}}_c\Big)=\overset{\bullet}{\bar{\lambda}}\mathbf{f}_{cn},\quad \mathbf{f}_{cn}=c\Big(\bar{\mathbf{n}}-\frac{\mathscr{R}_c}{\chi}\hat{\mathbf{n}}_c\Big) \tag{29}$$

$$\left(=\begin{cases} c\mathbf{d}^p & \text{for } \mathscr{R}_c=0 \\ \mathbf{O} & \text{for } \bar{\mathbf{n}}=\hat{\mathbf{n}}_c \\ 2c\mathbf{d}^p & \text{for } \bar{\mathbf{n}}=-\hat{\mathbf{n}}_c \end{cases}\right\} \text{for } \mathscr{R}_c=\chi$$

where c is a material constant or material parameter and

$$\hat{\mathbf{n}}_c\equiv\frac{\partial f(\hat{\mathbf{c}})}{\partial\mathbf{c}}\Big/\Big\|\frac{\partial f(\hat{\mathbf{c}})}{\partial\mathbf{c}}\Big\|=\frac{\partial f(\hat{\mathbf{c}})}{\partial\hat{\mathbf{c}}}\Big/\Big\|\frac{\partial f(\hat{\mathbf{c}})}{\partial\hat{\mathbf{c}}}\Big\| \quad (\|\hat{\mathbf{n}}_c\|=1) \tag{30}$$

Here, it follows from Eq. (29) that

$$\hat{\mathbf{n}}_c:\overset{\circ}{\mathbf{c}}=c\overset{\bullet}{\bar{\lambda}}\Big(\hat{\mathbf{n}}_c:\bar{\mathbf{n}}-\frac{\mathscr{R}_c}{\chi}\Big)<0 \quad \text{for } \mathscr{R}_c>\chi \tag{31}$$

Therefore, the evolution rule of the elastic-core in Eq. (29) is furnished with the distinguished ability in numerical calculation that the elastic-core is automatically pulled-back to the limit elastic-core surface when it goes out from that surface by the input of finite numerical increment as shown in **Fig. 6**.

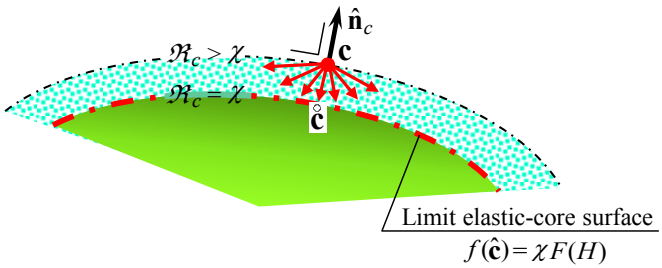


Fig. 6 Elastic-core is automatically attracted to the limit elastic-core surface when it goes out from the surface.

3.7 Plastic strain rate

The material-time derivative of Eq. (11) leads to the consistency condition of the subloading surface in the corotational time-derivative:

$$\frac{\partial f(\bar{\boldsymbol{\sigma}})}{\partial \bar{\boldsymbol{\sigma}}} : \mathring{\boldsymbol{\sigma}} - \frac{\partial f(\bar{\boldsymbol{\sigma}})}{\partial \bar{\boldsymbol{\sigma}}} : \mathring{\bar{\boldsymbol{\alpha}}} - \dot{R} F - R \dot{F} = 0 \tag{32}$$

It holds for Eq. (11) that

$$\frac{\partial f(\bar{\boldsymbol{\sigma}})}{\partial \bar{\boldsymbol{\sigma}}} : \bar{\boldsymbol{\sigma}} = RF \tag{33}$$

instead of Eq. (10), and then it follows that

$$\bar{\mathbf{n}} : \bar{\boldsymbol{\sigma}} = \frac{\partial f(\bar{\boldsymbol{\sigma}})}{\partial \bar{\boldsymbol{\sigma}}} : \bar{\boldsymbol{\sigma}} \Big/ \left\| \frac{\partial f(\bar{\boldsymbol{\sigma}})}{\partial \bar{\boldsymbol{\sigma}}} \right\| = RF \Big/ \left\| \frac{\partial f(\bar{\boldsymbol{\sigma}})}{\partial \bar{\boldsymbol{\sigma}}} \right\| \tag{34}$$

leading to

$$1 \Big/ \left\| \frac{\partial f(\bar{\boldsymbol{\sigma}})}{\partial \bar{\boldsymbol{\sigma}}} \right\| = \frac{\bar{\mathbf{n}} : \bar{\boldsymbol{\sigma}}}{RF} \tag{35}$$

The substitution of Eq. (35) into Eq. (32) leads to

$$\bar{\mathbf{n}} : \mathring{\boldsymbol{\sigma}} - \bar{\mathbf{n}} : \left[\left(\frac{\dot{F}}{F} + \frac{\dot{R}}{R} \right) \bar{\boldsymbol{\sigma}} + \mathring{\bar{\boldsymbol{\alpha}}} \right] = 0 \tag{36}$$

Here, the rate variable $\mathring{\bar{\boldsymbol{\alpha}}}$ is described from Eq. $(12)_2$ as

$$\mathring{\bar{\boldsymbol{\alpha}}} = R \mathring{\boldsymbol{\alpha}} + (1 - R) \mathring{\hat{\mathbf{c}}} - \dot{R} \hat{\mathbf{c}} \tag{37}$$

The substitution of Eq. (37) into Eq. (36) leads to

$$\bar{\mathbf{n}} : \mathring{\boldsymbol{\sigma}} - \bar{\mathbf{n}} : \left[\frac{\dot{F}}{F} \bar{\boldsymbol{\sigma}} + \frac{\dot{R}}{R} (\bar{\boldsymbol{\sigma}} - R \hat{\mathbf{c}}) + R \mathring{\boldsymbol{\alpha}} + (1 - R) \mathring{\hat{\mathbf{c}}} \right] = 0 \tag{38}$$

Further, noting the relation

$$\bar{\boldsymbol{\sigma}}-R\hat{\mathbf{c}}=\boldsymbol{\sigma}-\bar{\boldsymbol{\alpha}}-(\mathbf{c}-\bar{\boldsymbol{\alpha}})=\tilde{\boldsymbol{\sigma}} \tag{39}$$

it follows from Eq. (38) that

$$\bar{\mathbf{n}}:\mathring{\boldsymbol{\sigma}}-\bar{\mathbf{n}}:\left[\frac{\dot{F}}{F}\bar{\boldsymbol{\sigma}}+R\mathring{\boldsymbol{\alpha}}+\frac{\dot{R}}{R}\tilde{\boldsymbol{\sigma}}+(1-R)\mathring{\mathbf{c}}\right]=0 \tag{40}$$

Substituting Eqs. (16), (25) and (29) into Eq. (40), one has

$$\bar{\mathbf{n}}:\mathring{\boldsymbol{\sigma}}-\bar{\mathbf{n}}:\left[\frac{F'}{F}\dot{\bar{\lambda}}f_{hn}\bar{\boldsymbol{\sigma}}+R\dot{\bar{\lambda}}\,\mathbf{f}_{kn}+\frac{U}{R}\dot{\bar{\lambda}}\tilde{\boldsymbol{\sigma}}+c(1-R)\dot{\bar{\lambda}}\mathbf{f}_{cn}\right]=0 \tag{41}$$

where $F'\equiv dF/dH$.

Further substituting the associated flow rule in Eq. (22) into Eq. (41), one has

$$\bar{\mathbf{n}}:\mathring{\boldsymbol{\sigma}}-\bar{M}^p\dot{\bar{\lambda}}=0 \tag{42}$$

where

$$\bar{M}^p=\bar{\mathbf{n}}:\left[\frac{F'}{F}f_{hn}\bar{\boldsymbol{\sigma}}+c_kR\left(\bar{\mathbf{n}}-\frac{1}{b_k}\boldsymbol{\alpha}\right)+\frac{U}{R}\tilde{\boldsymbol{\sigma}}+c(1-R)\left(\bar{\mathbf{n}}-\frac{\mathfrak{R}_c}{\chi}\hat{\mathbf{n}}_c\right)\right] \tag{43}$$

by Eqs. (25) and (29).

The plastic multiplier $\dot{\bar{\lambda}}$ and the plastic strain rate $\mathbf{d}^p$ are given from Eq. (42) and (22) as follows:

$$\dot{\bar{\lambda}}=\frac{\bar{\mathbf{n}}:\mathring{\boldsymbol{\sigma}}}{\bar{M}^p},\quad \mathbf{d}^p=\frac{\bar{\mathbf{n}}:\mathring{\boldsymbol{\sigma}}}{\bar{M}^p}\bar{\mathbf{n}} \tag{44}$$

3.8 Stress-strain relations

The strain rate is given by substituting Eqs. (2) and $(44)_2$ into Eq. (1) as follows:

$$\mathbf{d}=\mathbf{E}^{-1}:\mathring{\boldsymbol{\sigma}}+\frac{\bar{\mathbf{n}}:\mathring{\boldsymbol{\sigma}}}{\bar{M}^p}\bar{\mathbf{n}} \tag{45}$$

from which the proportionality factor described in terms of the strain rate, denoted by $\dot{\Lambda}$ instead of $\dot{\bar{\lambda}}$, in the flow rule (22) is given as follows:

$$\dot{\Lambda}=\frac{\bar{\mathbf{n}}:\mathbf{E}:\mathbf{d}}{\bar{M}^p+\bar{\mathbf{n}}:\mathbf{E}:\bar{\mathbf{n}}},\quad \mathbf{d}^p=\frac{\bar{\mathbf{n}}:\mathbf{E}:\mathbf{d}}{\bar{M}^p+\bar{\mathbf{n}}:\mathbf{E}:\bar{\mathbf{n}}}\bar{\mathbf{n}} \tag{46}$$

The stress rate is given from Eq. (45) by use of Eq. $(46)_1$ instead of Eq. $(44)_1$, as follows:

$$\mathring{\boldsymbol{\sigma}}=\mathbf{E}:\mathbf{d}-\frac{\bar{\mathbf{n}}:\mathbf{E}:\mathbf{d}}{\bar{M}^p+\bar{\mathbf{n}}:\mathbf{E}:\bar{\mathbf{n}}}\mathbf{E}:\bar{\mathbf{n}} \tag{47}$$

3.9 Loading criterion

The loading criterion is given as follows :

$$\left.\begin{aligned}&\mathbf{d}^p\neq\mathbf{O}\ \text{for}\ \dot{\Lambda}>0\\&\mathbf{d}^p=\mathbf{O}\ \text{for}\ \dot{\Lambda}\le 0\end{aligned}\right\}\quad\text{or}\quad\left.\begin{aligned}&\mathbf{d}^p\neq\mathbf{O}\ \text{for}\ \bar{\mathbf{n}}:\mathbf{E}:\mathbf{d}>0\\&\mathbf{d}^p=\mathbf{O}\ \text{for}\ \bar{\mathbf{n}}:\mathbf{E}:\mathbf{d}\le 0\end{aligned}\right\} \tag{48}$$

where the judgment whether or not the stress reaches the yield surface is not required since the plastic strain rate is induced continuously as the stress approaches the normal-yield surface. It should be noted that the loading judgment by the plastic multiplier $\dot{\bar{\lambda}}$ in terms of

the stress rate in Eq. $(44)_1$ cannot be used for the softening behavior and even for the hardening behavior since the stress is required to be pulled-back to the normal-yield surface leading to the contraction of the subloading surface after it goes over the normal-yield surface in the plastic loading process.

3.10 Calculation of normal-yield ratio

Substituting Eq. (13) into Eq. (11), the subloading surface is described as follows:

$$f(\tilde{\boldsymbol{\sigma}} + R\hat{\mathbf{c}}) = RF(H) \tag{49}$$

from which the normal-yield ratio R is calculated by substituting the updated values of $\boldsymbol{\sigma}$, $\boldsymbol{\alpha}$, $\mathbf{c}$, F.

The normal-yield ratio can be calculated by the following two methods:

1) We calculate it from Eq. (49) in both of the plastic (loading) and the elastic (unloading) processes after all the other variables are calculated.

2) We calculate it by Eq. (49) in the elastic (unloading) process and in the state $R \leq R_e$ but we calculate it by the time-integration of Eq. (16) in the plastic (loading) process. Here, the analytical time-integration in Eq. (21) is beneficial to the enhancement of numerical analysis in the return-mapping projection 0. However, its use spoils the controlling function to pull-back the stress to the normal-yield surface numerical analysis in the forward-Euler method.

The second method 2) would be superior to the first method 1), since the normal-yield ratio is calculated directly from the plastic strain rate.

3.11 Expression of Masing rule

Note the following facts:

1) The difference between the curvatures in the reloading and the reverse loading curves becomes larger as the plastic deformation proceeds, which is called the Masing rule.

2) The elastic core approaches the normal-yield surface, following the current stress, when the plastic deformation proceeds continuously, and the approaching degree of the elastic core to the normal-yield surface is expressed by the elastic-core yield ratio $\mathscr{R}_c$ in Eq. (26).

3) The transition from the elastic to plastic state is more abrupt, i.e. the curvature of stress–strain curve is greater for a larger value of the material parameter u in the function $U(R)$ in Eq. (16). Therefore, the increase in the curvature of stress-strain curve can be described by giving a larger value to the material parameter u.

4) By the facts 1)-3), the difference between the values u in the reloading process and the reverse loading process is greater for a larger value of $\mathscr{R}_c$.

5) The direction $\bar{\mathbf{n}}$ of plastic strain rate is near to the outward-normal $\hat{\mathbf{n}}_c$ of the elastic core surface (Fig. 2) in the reloading process but it is far from $\hat{\mathbf{n}}_c$ in the unloading process. Then, the degree how the process is near the reloading process can be expressed by the following scalar product of these unit tensors:

$$C_\sigma \equiv \hat{\mathbf{n}}_c : \bar{\mathbf{n}} \quad (-1 \leq C_\sigma \leq 1) \tag{50}$$

Eventually, introducing the variables $\mathscr{R}_c$ and C_σ, let the material parameter u in Eq. (18) or Eq. (20) be extended as follows:

$$u = \bar{u}\exp(u_c \mathscr{R}_c C_\sigma) \tag{51}$$

where $\bar{u}$ (average value of u) and u_c are the material constant. u is the continuous function of the variables $\mathscr{R}_c$ and C_σ. $C_\sigma = 1,\ 0$ and -1 designate the states that the current stress has the outward-normal, tangential and inward-normal directions, respectively, of the similarity-center surface. Then, u increases in the loading direction but inversely it decreases in the opposite direction.

3.12 Plastic spin

The plastic spin in Eq. (6) is given following Zbib and Aifantis [11] by

$$\mathbf{w}^p = \eta^p(\boldsymbol{\sigma}\mathbf{d}^p - \mathbf{d}^p\boldsymbol{\sigma}) = \eta^p \overset{\bullet}{\lambda}(\boldsymbol{\sigma}\bar{\mathbf{n}} - \bar{\mathbf{n}}\boldsymbol{\sigma}) \tag{52}$$

where η^p is the material parameter.

REFERENCES

[1] Hashiguchi, K. *Elastoplasticity Theory*, Lecture Note in Appl. Compt. Mech., 2nd Edition, Springer, (2013).

[2] Hashiguchi, K. and Ueno, M. Elastoplastic constitutive laws of granular materials. *Constitutive Equations of Soils* (*Proc. 9th Int. Conf. Soil Mech. Found. Eng., Spec. Ses. 9*, S. Murayama and A.N. Schofield (eds.), Tokyo, JSSMFE, pp. 73-82 (1977).

[3] Hashiguchi, K. General description of elastoplastic deformation/sliding phenomena of solids in high accuracy and numerical efficiency: Subloading Surface Concept. *Arch. Compt. Meth. Eng.* (2013) **20**: 361-417.

[4] Hashiguchi, K. Constitutive equations of elastoplastic materials with elastic-plastic transition. *J. Appl. Mech.* (ASME) (1980) **47**: 266-272.

[5] Hashiguchi, K. Subloading surface model in unconventional plasticity. *Int. J. Solids Structures* (1980) **25**: 917-945.

[6] Mroz, Z. On the description of anisotropic workhardening. *J. Mech. Phys. Solids*. (1967) **15**:163-175.

[7] Dafalias, Y.F. and Popov, E.P. A model of nonlinearly hardening materials for complex loading. *Acta Mech*. (1975) **23**: 173-192.

[8] Chaboche, J.L., Dang-Van, K. and Cordier, G. Modelization of the strain memory effect on the cyclic hardening of 316 stainless steel. *Trans. 5th Int. Conf. SMiRT*, Berlin, Division L. Paper No. L. 11/3, (1979).

[9] Hashiguchi, K., Ueno, M. and Ozaki, T. Elastoplastic model of metals with smooth elastic-plastic transition", *Acta Mech*. (2012) **223**: 985-1013.

[10] Armstrong, P.J. and Frederick, C.O. A mathematical representation of the multiaxial Bauschinger effect. *CEGB Report RD/B/N 731*, 1966 (or in: *Materials at High Temperature* (2007) **24**: 1-26).

[11] Zbib, H.M. and Aifantis, E.C. On the concept of relative and plastic spins and its implications to large deformation theories. Part I: Hypoelasticity and vertex-type plasticity. *Acta Mech*. (1988) **75**: 15-33.

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis (Eds)

FINITE ELEMENT ANALYSIS OF DAMAGE-HEALRING BEHAV-IOUR IN SELF-HEALING CERAMIC MATERIALS

SHINGO OZAKI[*], YUTARO HORIE[*], KOSUKE HINATA[*], TOSHIO OSADA[†] AND WATARU NAKAO[*]

[*] Faculty of Engineering, Yokohama National University
Yokohama National University
Tokiwadai 79-5, Hodogaya-ku, Yokohama 240-8501, Japan
e-mail: s-ozaki@ynu.ac.jp, wakao@ynu.ac.jp

[†]Research Center for Strategic Materials
, National Institute for Materials Science
Sengen 1-2-1, Tsukuba, Ibaraki 305-0047, Japan
email: OSADA.Tosiho@nims.go.jp

Key words: Self-healing Material, Constitutive Equation, Continuum Damage Mechanics, Finite Element Method, Ceramics Matrix Composite.

Abstract. In this study, we develop the constitutive model to analyse the self-healing ceramic materials within the framework of FEM. The self-hearing and isotropic damage constitutive model for ceramic materials can describe not only the damage process under a certain boundary condition, but also the self-healing process under a high-temperature condition. The damage process is formulated based on the fracture mechanics, and the self-healing process is formulated based on the kinetic model of self-healing time and velocity. Then, we apply the proposed model to analyses of homogeneous ceramic materials and unit cell model of fiber-reinforced ceramic material.

1 INTRODUCTION

The self-healing fiber-reinforced ceramic material (shFRC) developed by the authors is a new functional material [1-3]. The self-crack/damage-hearing function is one of most valuable phenomena to overcome the reliability decrease of blittle ceramics that are caused by non-acceptable cracking. When a micro-crack propagates in this material, the self-healing occurs under high-temperature. Then, the strength of the material recovers to its initial state because the crack is re-bonded. The reason why the self-crack/hearing automatically attains the complete recovery of damaged strength is that the passive oxidation of self-healing agent, for example SIC, is caused by a crack itself. However, in order to apply the self-hearing ceramic material to machines and constractions, it is important to develop the novel numerical simulation method.

In this study, we develop the constitutive model to analyse the self-healing ceramic materials within the framework of FEM. The damage process is formulated based on the fracture

mechanics [3], and the self-healing process is formulated based on the kinetics of self-healing time and velocity [1]. Therefore, the proposed self-hearing and isotropic damage constitutive model for ceramic materials can describe not only the damage process under a certain boundary condition, but also the self-healing process under a high-temperature condition. Then, we apply the proposed model to FE analyses of homogeneous ceramic material and unit cell model of ceramic matrix composites.

2 CONSTITUTIVE EQUATION

To formulate both the isotropic damage and self-healing behavior, we adopt the damage constitutive equation proposed by Kurumatani et al.[4]. In this model, the strain softening behavior due to damage is formulated based on the fracture mechanics of concrete material. Concretely, the critical energy release rate and the work due to the cohesive force and relative displacement relation are related, i.e. the damage behavior is described like a cohesive element model. In addition, the dependence on mesh size and mesh kind can be reduced because the size of finite element is incorporated as a characteristic length. In the following, first, we will explain the formulation under one-dimensional problem. Then, we will explain about the extension to three-dimensional problem using equivalent strain. Finally, we will explain about the evolution rule for self-healing behavior based on kinetics.

2.1 Formulation for one-dimensional problem

The one-dimensional damage model is described using scalar variable D as follows:

$$\sigma = (1-D)E\varepsilon \tag{1}$$

where σ, ε and E are stress, strain and Young's modulus, respectively. The damage variable D $(0 \le D \le 1)$ prescribes magnitude of damage, i.e. D=0 is a non-damage state and D=1 is a perfectly fracture state.

We assume that the relation of cohesive force and relative displacement (crack opening displacement) is expressed by an exponential function. Thus, the relation of cohesive force and crack opening displacement by tensile fracture in one-dimensional problem is given by the following equation.

$$f = Ae^{-Bw} \tag{2}$$

where f is the cohesive force on fracture surface, w is the crack opening displacement. And A and B are unknown parameters. If we assume f_t is one-dimensional tensile fracture strength, we can obtain the unknown parameter A as follows:

$$f_t = Ae^0 \rightarrow A = f_t \tag{3}$$

The fracture energy G_f is given as

$$G_f = \int_0^\infty f dw = A\int_0^\infty e^{-Bw} dw \tag{4}$$

Thus, the unknown parameter B is obtained as follows:

$$G_f = \frac{A}{B} \rightarrow B = \frac{A}{G_f} = \frac{f_t}{G_f} \tag{5}$$

Consequently, the relation of cohesive force and crack opening displacement in one-dimensional problem is obtained as follows:

$$f = f_t \exp\left(-\frac{f_t}{G_f} w \right) \tag{6}$$

In the following, Eq.(6) based on the fracture energy is applied to the damage constitutive model. An strain corresponding to the tensile strength is defined as the initial damage strain ε_0 and is given by

$$\varepsilon_0 = \frac{f_t}{E} \tag{7}$$

The crack opening displacement w is given by the relationship between displacement and strain as follows:

$$w = \varepsilon h_e - \varepsilon_0 h_e = (\varepsilon - \varepsilon_0) h_e \tag{8}$$

where h_e stands the length of finite element. In the damage model, the cohesive force (traction) f on the fracture surface is equivalent to stress σ, as the following equation.

$$f = \sigma \tag{9}$$

The relation of cohesive force and crack opening is obtained from Eqs. (6)-(9), as follows:

$$\begin{aligned} \sigma &= E\varepsilon_0 \exp\left(-\frac{E\varepsilon_0 h_e}{G_f} (\varepsilon - \varepsilon_0) \right) = \left[1 - \left\{ 1 - \frac{\varepsilon_0}{\varepsilon} \exp\left(-\frac{E\varepsilon_0 h_e}{G_f} (\varepsilon - \varepsilon_0) \right) \right\} \right] E\varepsilon \\ &= [1 - D(\varepsilon)] E\varepsilon \end{aligned} \tag{10}$$

Representing ε in the damage variable $D(\varepsilon)$ by the maximum strain $\bar{\varepsilon} \geq 0$ during deformation history, the damage variable $D(\bar{\varepsilon})$ is rewritten as follows:

$$D(\bar{\varepsilon}) = 1 - \frac{\varepsilon_0}{\bar{\varepsilon}} \exp\left(-\frac{E\varepsilon_0 h_e}{G_f} (\bar{\varepsilon} - \varepsilon_0) \right) \tag{11}$$

We can judge the loading criterion in accordance with the magnitude relationship between ε and $\bar{\varepsilon}$, i.e.

$$\begin{cases} \text{if } \ \bar{\varepsilon} \leq \varepsilon \rightarrow \bar{\varepsilon} = \varepsilon \ \text{(loading)} \\ \text{if } \ \bar{\varepsilon} > \varepsilon \rightarrow \bar{\varepsilon} = \bar{\varepsilon} \ \text{(unloading)} \end{cases} \tag{12}$$

In addition, the judgment of damage is given as follows:

$$\begin{cases} \text{if } \ \varepsilon < \varepsilon_0 \rightarrow D = 0 \ \text{(undamage)} \\ \text{if } \ \varepsilon \geq \varepsilon_0 \rightarrow D = D(\bar{\varepsilon}) \end{cases} \tag{13}$$

2 Formulation for three-dimensional problem

In general, the equivalent strain, which is scalar value, is used for the isotropic damage model in multi-dimensional problems. Kurumatani et al.[4] have adopted the following equivalent strain ε_{eq} based modified von-Mises type [5].

$$\varepsilon_{eq} = \frac{k-1}{2k(1-2\nu)} I_1 + \frac{1}{2k}\sqrt{\left(\frac{k-1}{1-2\nu} I_1\right)^2 + \frac{12k}{(1+\nu)^2} J_2} \tag{14}$$

where ν is Poisson's ratio, and k is a ratio of the compressive and tensile strengths. I_1 is the first invariant of strain tensor $\boldsymbol{\varepsilon}$ and is given by

$$I_1 = \mathrm{tr}\boldsymbol{\varepsilon} = \varepsilon_{kk} \tag{15}$$

J_2 is the second invariant of deviatoric strain tensor $\boldsymbol{\varepsilon}^*$ and is given by

$$J_2 = \frac{1}{2}\boldsymbol{\varepsilon}^* : \boldsymbol{\varepsilon}^* = \frac{1}{2}\varepsilon^*_{kl}\varepsilon^*_{kl} \tag{16}$$

The deviatoric strain is defined by

$$\boldsymbol{\varepsilon}^* = \boldsymbol{\varepsilon} - \frac{1}{3}\mathrm{tr}\boldsymbol{\varepsilon}\mathbf{I} \tag{17}$$

Here, $\mathbf{I}$ is unit tensor.

Next, we will describe the stress-strain relation in multi-dimensional problem. The isotropic damage model is described using damage variable D as follows:

$$\boldsymbol{\sigma} = (1-D)\mathbf{c} : \boldsymbol{\varepsilon} \tag{18}$$

where $\boldsymbol{\sigma}$ is Cauchy stress tensor, $\mathbf{c}$ is elastic coefficient tensor. In this study, we adopt the hypo-elastic coefficient tensor.

By using the maximum value of equivalent strain $\kappa \geq 0$ in deformation history, the damage variable $D(\kappa)$ is given as follows:

$$D(\kappa) = 1 - \frac{\kappa_0}{\kappa}\exp\left(-\frac{E\kappa_0 h_e}{G_f}(\kappa - \kappa_0)\right) = 1 - \frac{\kappa_0}{\kappa}e^{-\beta(\kappa-\kappa_0)} \tag{19}$$

where κ_0 is the initial damage strain. In what follows, κ is called the history variable or the maximum equivalent strain. The length of three-dimensional solid element is defined based on the volume V_e as follows:

$$h_e = V_e^{1/3} \tag{20}$$

The loading criterion is given by

$$\begin{cases} \text{if } \kappa \leq \varepsilon_{eq} \rightarrow \kappa = \varepsilon_{eq} \text{ (loading)} \\ \text{if } \kappa > \varepsilon_{eq} \rightarrow \kappa = \kappa \text{ (unloading)} \end{cases} \tag{21}$$

$$\begin{cases} \text{if } \kappa < \kappa_0 \rightarrow D = 0 \text{ (undamage)} \\ \text{if } \kappa \geq \kappa_0 \rightarrow D = D(\kappa) \end{cases} \tag{22}$$

2.3 Evolution rule for self-hearing

To incorporate the self-healing behavior to the damage constitutive model, we assume that a micro cracks (damage) are recovered by self-healing function. In other words, the damage variable is varied from the damage state $D \neq 0$ to the non-damage state $D = 0$ in accordance with the self-healing that is affected by the temperature and atmosphere conditions. Thus, we assume that the history variable κ evolves to be closer to κ_0 with the self-healing. That is, the damage history is gradually dissipated by the self-healing. Based on abobe-mentioned concept, we will formulate the evolution rule for self-healing.

First, we assumed that the history variable is additively decomposed into the equivalent strain part κ_ε and the self-healing part $-\kappa_h$ as follows:

$$\kappa = \kappa_\varepsilon - \kappa_h \tag{23}$$

where the contributed part by progress of equivalent strain ε_{eq} is given as

$$\kappa_\varepsilon = \int_t \dot{\varepsilon}_{eq} dt \quad \text{if} \quad \varepsilon_{eq} \geq \bar{\kappa} \tag{24}$$

$$\kappa_\varepsilon = \kappa_\varepsilon + < \Delta\varepsilon_{eq} > \tag{25}$$

On the other hand, the self-healing part κ_h is assumed a monotonic increasing function to approach the state $\kappa \to \kappa_0$. we then assume the following simple function for the evolution rule of κ_h.

$$\dot{\kappa}_h = \xi_1 v_h (\kappa - \kappa_0) = \xi_1 v_h (\kappa_\varepsilon - \kappa_h - \kappa_0) \tag{26}$$

$$\kappa_h = \int_t \dot{\kappa}_h dt \approx \kappa_h + \dot{\kappa}_h dt \tag{27}$$

where ξ_1 is the parameter that influences the rate of self-healing. v_h [s^{-1}] is the average velocity of self-healing based on kinetics. In this study, we adopt the following equation proposed by Osada et al.[1].

$$v_h = \frac{1}{t_h^{\min}} = A_h \cdot \exp\left(\frac{-Q_h}{RT_h}\right) a_{O_2}{}^{3n/2} \tag{28}$$

where A_h is the frequency factor, Q_h is the activation energy for self-healing, R is the gas constant, and T_h is the temperature. a_{O_2} and n are the activity of Oxygen and the reaction order of Oxygen, respectively. Thus, we can consider $h_e \dot{\kappa}_h$ as a filling velocity of mcro-crack.

Furthermore, to describe that a restart of damage depends on a degree of the self-healing, the maximum equivalent strain $\bar{\kappa}$ has to be re-defined. In the loading state, as well as the normal damage model, $\bar{\kappa}$ is given by the following manners.

$$\bar{\kappa} = \max\{\kappa_\varepsilon\} \geq \kappa_0 \quad \text{for loading with damage} \tag{29}$$

On the other hand, in the self-healing state, we assume that the maximum equivalent strain $\bar{\kappa}$ gradually approaches to κ_0. We then assume the evolution rule of the maximum equivalent strain as follows:

$$\dot{\bar{\kappa}} = -\xi_2 v_h \left(\bar{\kappa} - \kappa_0 \right) \tag{30}$$

$$\bar{\kappa} = \int_t \dot{\bar{\kappa}} dt \approx \bar{\kappa} + \dot{\bar{\kappa}} dt \tag{31}$$

That is, in the reloading state, the damage does not occur until $\min(\varepsilon_{eq},\ \kappa) \geq \bar{\kappa}$. Here, ξ_2 is the parameter that influences the rate of self-healing. Although the initial damage strain in the reloading state is $\bar{\kappa}$, strain necessary to reach it is adjusted by the magnitude of ξ_1 and ξ_2. In the case of $\xi_1 \geq \xi_2$, the super healing phenomenon, in which the fracture strength of healed material becomes higher than that of virgin material, is naturally described. Note that if we set the average velocity of self-healing to zero, i.e. $v_h = 0$, the condition $\kappa = \kappa_\varepsilon = \bar{\kappa}$ is satisfied, and thus the model becomes the normal damage model as described in Section 2.2.

By considering of the evolution rules for above-mentioned history variables, the loading criterion is summarized as follows:

$$\begin{cases} \text{if } \left\langle \Delta\varepsilon_{eq} \right\rangle > 0 \rightarrow \kappa_\varepsilon = \kappa_\varepsilon + \Delta\varepsilon_{eq} \text{ (loading)} \\ \text{if } \left\langle \Delta\varepsilon_{eq} \right\rangle = 0 \rightarrow \kappa_\varepsilon = \kappa_\varepsilon \text{ (unloading)} \end{cases} \tag{32}$$

$$\begin{cases} \text{if } \min(\varepsilon_{eq},\ \kappa) < \bar{\kappa} \rightarrow D = 0 \quad \text{for } \kappa < \kappa_0 \\ \text{if } \min(\varepsilon_{eq},\ \kappa) < \bar{\kappa} \rightarrow D = D(\kappa) \quad \text{for } \kappa \geq \kappa_0 \\ \text{if } \min(\varepsilon_{eq},\ \kappa) \geq \bar{\kappa} \rightarrow D = D(\kappa) \quad \text{for } \bar{\kappa} = \kappa \end{cases} \tag{33}$$

It should be noted that the following equations based on the logistic curve would be suitable for the evolution rule of self-healing, instead of Eqs.(26) and (30).

$$\dot{\kappa}_h = \xi_1 v_H \kappa_h (\kappa_\varepsilon - \kappa_0 - \kappa_h) \tag{34}$$

$$\dot{\bar{\kappa}} = -\xi_2 v_h \bar{\kappa} \left(\kappa_0 - \bar{\kappa} \right) \tag{35}$$

3 FINITE ELEMENT ANALYSIS

In this section, we demonstrate the analysis of damage and healing phenomena within the framework of continuum theory using the FEM implemented with the proposed model.

Table 1: Material parameters.

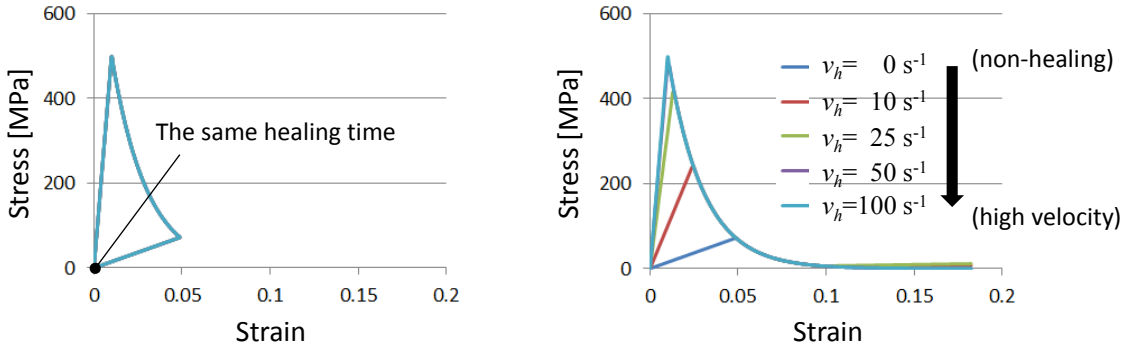
E [MPa]	50000	k	10
G [MPa]	25000	h_e [mm]	10
ν	0.3	v_h [s^{-1}]	0-100
G_f [J]	100	ξ_1	1.0
κ_0	0.01	ξ_2	1.0

3.1 Mechanical response of proposed model

In this study, the self-healing and isotropic damage model is implemented in the commercial FEM software package LS-DYNA Ver.971 [6]. To demonstrate the capability of

the proposed model, we created a simplest FE model. The model is discretized by one eight-node solid element. The element is $10\times10\times10$ mm in size. Prescribed forced cycle displacement into uniaxial direction is imposed at upper surface. Material parameters used for the analysis is listed in Table 1. Note that, to reasonably verify the mechanical response of proposed model, we adopt quite high values for the average velocity of self-healing, v_h.

Figure 1(a) and 1(b) show the effect of v_h on the stress-strain relation under cyclic loading. Figure 1(a) and 1(b) present result of "loading-unloading-rest" and "reloading" states, respectively, where the healing time is the same for each condition. It can be confirmed from Figure 1 that the self-healing occurs during stop of loading and the fracture strength recovers in accordance with the average velocity of self-healing. Figure 2 shows the variation of damage variable D during cyclic loading as shown in Figure 1. By introducing of the evolution rule of Eqs.(26) and (30), the damage variable gradually approaches to zero. It is also confirmed that the variation rate of damage variable depends on v_h.



(a) Loading-unloading- rest (b) Reloading

Figure 1: Effect of v_h on stress-strain relation under cyclic loading.

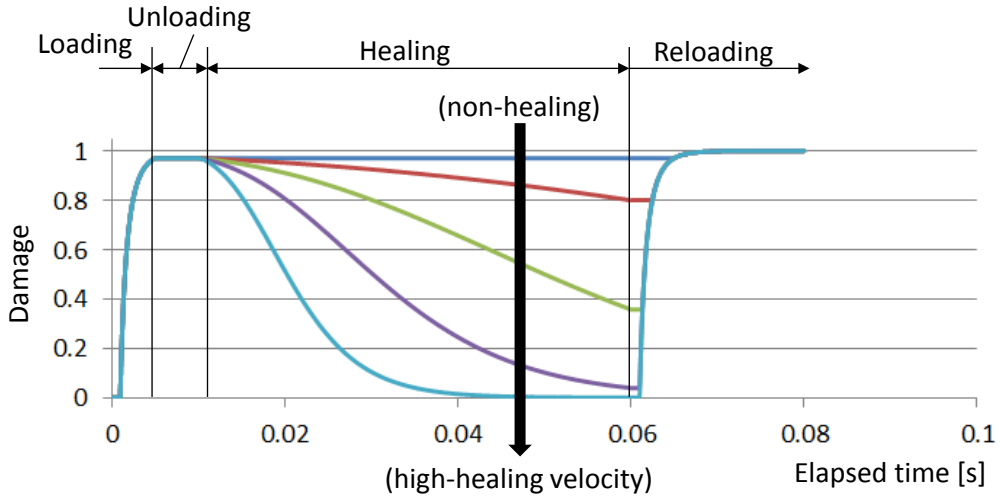
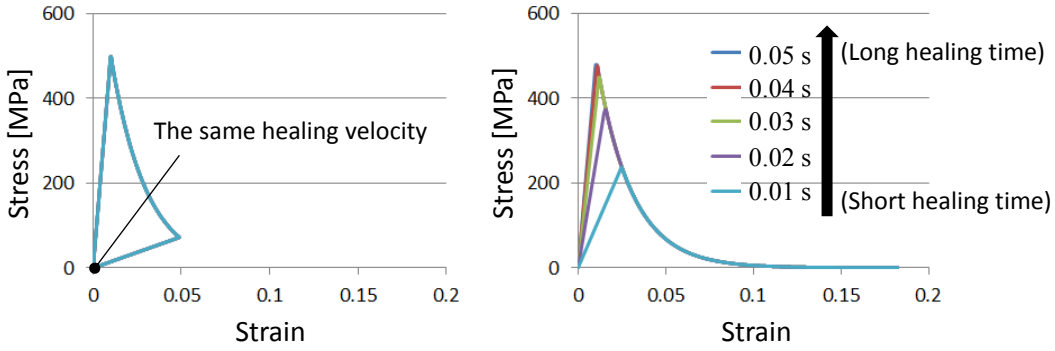


Figure 2: Variation of damage variable D during loading-rest-reloading.

Figure 3(a) and 3(b) show the effect of healing time on stress-strain relation under cyclic loading. Figure 3(a) and 3(b) present result of "loading-unloading-rest" and "reloading" states, respectively, where the average velocity of self-healing is the same for each condition. It can be confirmed from Figure 3 that the fracture strength recovers in accordance with the healing time.



(a) Loading-unloading- rest (b) Reloading

Figure 3: Effect of healing time on stress-strain relation under cyclic loading.

3.2 Analysis of unit cell model of fiber reinforced ceramic

Next, we analyze the damage and self-healing processes of ceramic matrix composite. Figure 4 shows the FE analysis model and imposed boundary conditions, where eight-node solid elements are adopted for discretization of ceramic matrix composite. In this study, we adopt the unit cell model of fiber reinforced ceramic matrix composite [2] for the analysis model. In the calculation, we apply the prescribed forced cyclic displacement to the side surface of unit cell. To reproduce the damage progress from matrix part, the initial defect is given as shown in Figure 4. Material parameters used for the analysis is listed in Table 2, while the fiber is assumed to an elastic material having $E = 80$ [GPa] and $\nu = 0.2$. Note that, to reasonably verify the mechanical response of proposed model, we adopt quite high values for the average velocity of self-healing, v_h.

Figure 5(a) and 5(b) show the relationship between nominal stress and nominal strain of unit cell under cyclic loading. Figure 5(a) and 5(b) present result of "normal damage model" and "self-healing and damage model", respectively, where the healing time is the same for each condition. Figure 6 shows the contour map of damage variable D during cyclic loading in the case of high healing velocity condition. As can be seen from Figures, the damaged region progresses from initial defect point and the stiffness of unit cell decreases during loading. After unloading, the self-healing occurs and the damaged region recovers to initial state as shown in Figure 6. Then, the fracture strength and stiffness are recovered according to the healing velocity v_h and the healing time. These results suggest that the proposed model can be applied to analyses of ceramic matrix composites having self-healing functions.

Table 2: Material parameters for fiber reinforced cermic.

E [MPa]	43200	k	10
G [MPa]	18000	h_e [mm]	0.008
ν	0.2	v_h [s^{-1}]	200, 2000
G_f [Nmm]	10	ξ_1	1.0
κ_0	0.0056	ξ_2	1.0

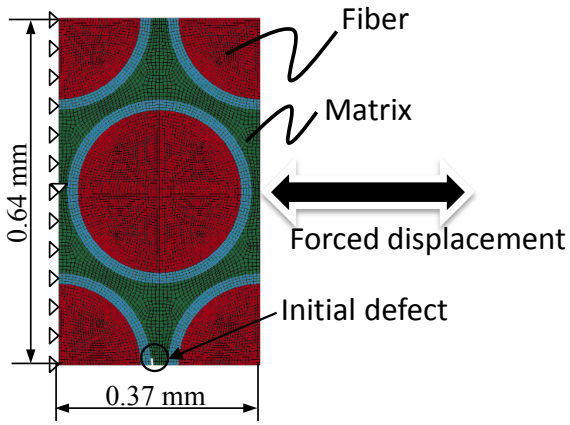
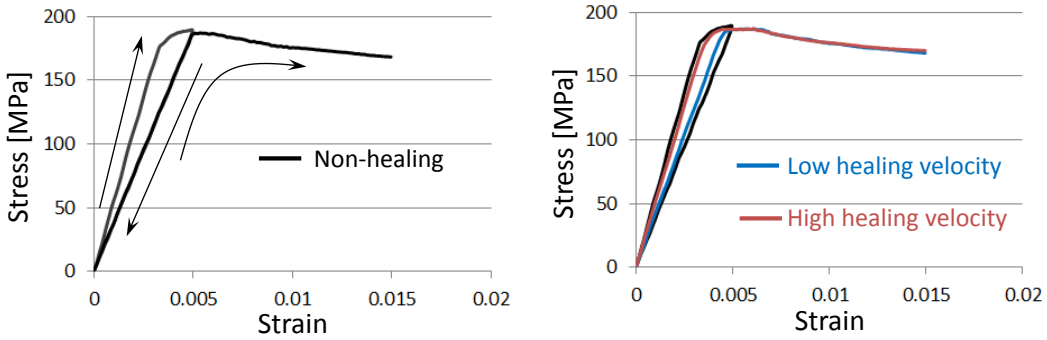


Figure 4: FE analysis model and boundary conditions of unit cell (thickness is 0.005 mm)



(a) Normal damage model (b) Self-healing and damage model

Figure 5: Relationship stress and strain of unit cell under cyclic loading.

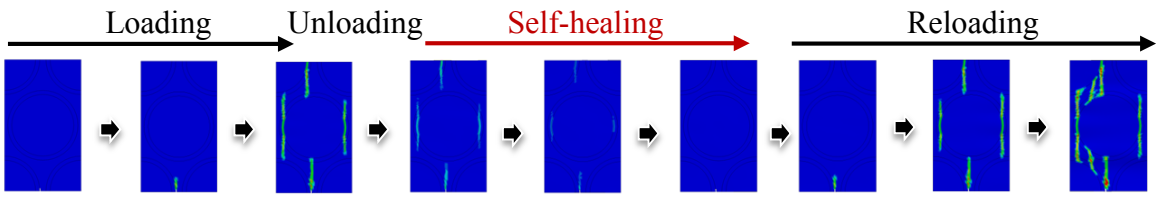


Figure 6: Contour map of damage variable during cyclic loading in the case of high healing velocity condition.

4 CONCLUSIONS

The present study formulated the self-healing and damage constitutive equation. We then applied it to the typical FE problems and then verified the formulation. The proposed model can describe the competition between the progress of damage and the self-healing by the unified formulation. Thus, the proposed model can be applied arbitrary deformation histories, including cyclic loading and strain holding. From the results of FE analyses, we can conclude that the present FE approach can be used to study the mechanical and material design of self-healing ceramic materials.

In the present FE analysis, however, the quite large values for healing velocity were adopted. Therefore, effects of temperature and oxygen conditions on healing velocity were not examined. Eq.(28) based on kinetics should be applied to FE analyses, in near future. Then, the verification of the concreate functions for the evolution rules should be performed by comparing with experiments.

ACKNOWLEDGMENT

This work has been supported by the Advanced Low Carbon Technology Research and Development Program (ALCA), JST ,Japan. We wish to thank Prof. Mao Kurumatani, Ibaraki University, for useful discussions and comments.

REFERENCES

[1] Osada, T., Nakao, W., Takahashi, K. and Ando, K. Kinetics of self-crack-healing of alumina/silicon carbide composite including oxygen partial pressure effect. *Journal of the American Ceramic Society* (2009) **92**:864-869.

[2] Nakao, W. and Haga, Y. *Dvelopment of selfhealing fiber reinforced ceramics matrix composite*, Japan Patent Kokai (2010) 2010-290947.

[3] Osada, T., Nakao, W., Takahashi, K. and Ando, K. Self-crack-healing behavior in ceramics matrix composite, *Advances in ceramics matrix composites* (2014) **17**: 410-441.

[4] Kurumatani, M., Terada, K., Kato, J., Kyoya, T. and Kashiyama, K. An isotropic damage model based on fracture mechanics for concrete and its evaluation, *Transactions of JSCES* (2013) No.20130015, *in Japanese*.

[5] de Vree, J.H.P., Brekelmans, W.A.M. and van Gils, M.A.J. Comparison of nonlocal approaches in continuum damage mechanics, *Comput. Struct.* (1995) **55**:581-588.

[6] LS-DYNA Ver.971, *Keywords manual & Theory manual*. LSTC (2012).

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis (Eds)

ROBUST OPTIMIZATION APPROACH FOR MIXED NUMERICAL/EXPERIMENTAL IDENTIFICATION OF ELASTIC PROPERTIES OF ORTHOTROPIC COMPOSITE PLATES

JANIS AUZINS[*], EDUARDS SKUKIS[†]

[*] Institute of Mechanics
Riga Technical University
Ezermalas street 6k, Riga, LV-1014, Latvia
e-mail: auzinsjp@latnet.lv

[†] Institute of Materials and Structures
Riga Technical University
Azenes street 16/20, Riga, LV-1048, Latvia
e-mail: Eduards.Skukis@rtu.lv

Key words: Metamodeling, Elastic properties, Vibration tests, Finite element method, Uncertainty.

Abstract. This paper describes a method for determination of elastic parameters (elastic moduli and Poisson's ratio) of orthotropic composite plate-type structural elements using the results of natural frequency measurements. The identification of parameter values is provided by minimization of weighted squared difference (discrepancy) between physically measured frequencies and natural frequencies calculated by Finite Element Method. The metamodels for the frequency dependence on the elastic parameters and other geometrical and physical parameters of test specimens, including parameters with uncertainty ("noisy constants") are built using experimental designs optimized according to the Mean Squared Error space filling criterion and third-order polynomial approximations. The minimum of weighted squared difference is found using the multistart random search method. The expressions for standard deviations of identified parameters depending on deviations of "noisy constants" are derived using linearized metamodels. The expressions for identification errors allow the statement of the identification task as a robust minimization problem by simultaneous minimization of the discrepancy function and standard deviations of the identified values by varying the values of unknown elastic parameters and weighting coefficients for different frequencies. The partial scaling of natural frequencies is used for the reduction of the uncertainty impact on the identification error. This allows to reduce the identification error of elastic moduli about two times and Poisson's ratio about 20 times in comparison with the results obtained by using dimensioned frequencies.

1 INTRODUCTION

The method of identification of elastic parameters (Young's moduli, shear moduli, Poisson's ratio) using eigenfrequency measurements of specimens is quite old. Currently there exists an extensive amount of literature on the identification of elastic properties of layered

composite materials using physical measurements and numerical calculations of natural frequencies, mostly using Finite Element Method (FEM) [1,2,3]. The traditional numerical-experimental identification procedure is based on the minimization of discrepancy between numerical and experimental results. There have been papers published suggesting that normal maximum likelihood is superior to weighted least squares for the determination of material elastic properties using the vibration method [4,5]. At the same time it has been proven that the three main estimation methods (normal maximum likelihood, weighted least squares and ridge regression) all have the same asymptotic covariance and that there is no gain in efficiency among them [6].

During the first years of using discrepancy minimization method, the main problem was the minimization of the discrepancy functional. FEM software cannot be used in identification, since the minimization of the differences between physical measurements and numerical results requires many thousands of FEM calculations that would take years to run. Therefore the metamodeling (also called surrogates) methodology is used [2,3]. The methodology consists in the creation of experimental designs for computer experiments with FEM software, carrying out computer experiments (100-300 experimental runs), building the approximate model for the dependence of natural frequencies of a given structural element on their geometrical and elastic parameters and finding the values of elastic parameters that minimize the difference between measured and calculated frequency values. The last step in this methodology is the verification of the result – repeated calculation of FEM results and validation – comparing identified elastic parameter values with values obtained by a different method (bending tests, for example).

Today, using modern numerical experimental designs and nonparametric approximation methods, the discrepancy minimization is not a hard task. However, the estimation of the variance of obtained identified parameters is a pressing problem. The errors of identification depend on errors introduced by material production, cutting testing specimens, physical measurement errors and errors caused by disregarding significant factors in the finite element model. A large amount of literature is devoted to analysis of the accuracy of FEM, but the influence of errors of physical experiments, caused by parameter variance during material production, specimen preparation and errors of registration and measurement of natural frequencies, is significantly less studied.

This paper will propose a methodology that allows both to minimize the difference between physically measured and numerically calculated natural frequency values, and to evaluate the robustness of the identification results and organize the identification process in a way that would allow obtaining the best possible identification precision.

2 VIBRATION-BASED IDENTIFICATION OF ELASTIC PARAMETERS

Identifiable elastic parameters usually are Young's moduli, shear moduli, Poisson's ratio. The classical idea is to find those values of elastic parameters of the mathematical FEM model which will give minimal discrepancy between calculated and physically measured natural frequencies.

We will designate the column vector of n physically measured natural frequencies as $\boldsymbol{f}^{\mathrm{EXP}}$

$$\boldsymbol{f}^{\mathrm{EXP}} = [f_1^{\mathrm{EXP}}, f_2^{\mathrm{EXP}}, \dots, f_L^{\mathrm{EXP}}]^{\mathrm{T}}, \tag{1}$$

column vector of numerically calculated (mostly using FEM) natural frequencies as $\boldsymbol{f}^{\mathrm{FEM}}$

$$\boldsymbol{f}^{\mathrm{FEM}} = [f_1^{\mathrm{FEM}}, f_2^{\mathrm{FEM}}, \dots, f_L^{\mathrm{FEM}}]^{\mathrm{T}} \tag{2}$$

and column vector of n physically identifiable parameters of elasticity as $\boldsymbol{E}$

$$\boldsymbol{E} = [E_1, E_2, \dots, E_n]^{\mathrm{T}}. \tag{3}$$

A superscript $^{\mathrm{T}}$ denotes the matrix transpose operation. The number n of identified parameters can be different, including elastic modulus and Poisson's ratio for different composite layers. The discrepancy Φ between measured and calculated natural frequencies is calculated as weighted sum of L squared differences:

$$\Phi = \sum_{i=1}^{L} w_i \left(f_i^{\mathrm{FEM}} - f_i^{\mathrm{EXP}}\right)^2, \tag{4}$$

where w_i – nonnegative weighting coefficient for i-th frequency.
A frequently used weighting method for discrepancy measure is the squared relative error:

$$\Phi = \sum_{i=1}^{L} \left(\frac{f_i^{\mathrm{FEM}} - f_i^{\mathrm{EXP}}}{f_i^{\mathrm{EXP}}}\right)^2. \tag{5}$$

The discrepancy minimization approach means that input parameters which give minimal value of functional Φ will be considered as identified values for unknown parameters $\boldsymbol{E}$:

$$\boldsymbol{E}^* = \underset{\boldsymbol{E}}{\arg\min}\, \Phi(\boldsymbol{E}). \tag{6}$$

When using FEM software, the minimization requires physical experiments as well as numerical experiments; therefore this approach is sometimes called Mixed Numerical-Experimental Technique (MNET) [2]. The traditional MNET steps are:

Step 1. Preparation of specimen samples, providing frequency measurements by resonance measurements or Fourier analysis of free oscillations registered after initial excitation.

Step 2. Design of numerical experiments for FEM software. The variable input factors for eigenfrequency calculations are identifiable elastic parameters. Mostly the Latin Hypercube (LH) type designs are used. Here we use LHs and non-LH type designs optimized according to the Mean Square Error space-filling criterion, introduced in [7]. The values for other input parameters (geometrical, mass, density, layer configuration and others) must correspond to specimens used in physical experiments. The number of runs for numerical experiments depends on the number m of identifiable parameters. For relatively simple plate-type specimens the calculations are fast enough to execute 100-300 trial runs in 15 minutes of computing time on PC with Intel quad-core i7 processor.

Step 3. Carrying out numerical experiments, registering and grouping the eigenfrequencies according to vibration modes.

Step 4. Building the metamodels (surrogate model) for the dependency of calculated eigenfrequencies $\boldsymbol{f}^{\mathrm{FEM}}$ on the input parameters $\boldsymbol{E}$.

$$\boldsymbol{f}^{\mathrm{FEM}} \approx \hat{\boldsymbol{f}}(\boldsymbol{E}), \tag{7}$$

where the "hat" above a function symbol signifies approximation. The software EDAOpt [8], created at Institute of Mechanics of Riga Technical University, was used for design of computer experiments, metamodel building and minimization of discrepancy functional.

Quality of approximation is estimated using leave-one-out cross-validation [9]. The relative cross-validation error *CVE* is calculated relatively to standard deviation from the mean value of responses. In cases where analytical thick plate models [9] are used instead of FEM, the prediction error was calculated in more than 10^6 test points and the mean error agrees very well with the cross-validation estimate.

Practice shows that almost in all cases third order polynomials give the approximation of frequencies with relative cross-validation error less than 0.02%. For some more complicated specimens the best accuracy can be obtained using nonparametric approximation methods: kriging, locally weighted polynomials.

Step 5. Finding the values of identifiable parameters by minimization of approximated discrepancy functional

$$\boldsymbol{E}^* = \underset{\boldsymbol{E}}{\arg\min} \sum_{i=1}^{n} w_i \left(\hat{f}_i(\boldsymbol{E}) - f_i^{\mathrm{EXP}}\right)^2 \tag{8}$$

The software EDAOpt uses a modified multi-start simulated annealing method [10] and always gives global minimums of the discrepancy functional. It must be noted that the polynomial metamodels are relatively simple: for evaluation of the objective function Eq. (8) even several millions of evaluations need only a few minutes of computing time.

Step 6. Traditionally, the next step is the recalculation of the metamodel in the sub-area near the found values of identified parameters, and analysis of the significance of different elasticity parameters for natural frequencies [2,3]. However, this analysis gives insufficient information for the estimation of accuracy of identified values. Therefore the present work proposes a method for accuracy estimation.

3 STANDARD DEVIATIONS OF IDENTIFICATION

3.1 Sources of composite material parameter deviations

The errors of elastic parameter identification using the vibration method depend on errors introduced by material production, cutting of testing specimens, physical measurement errors and errors of the finite element model. Practice shows that the mode measurements using the Polytec PSV-400-3D Laser Scanning Vibrometer (http://www.polytec.com/) have very high accuracy. The repeated measurements for a given specimen give almost the same results. At the same time the measurements of 4-6 different specimens from the experimental sample allow to estimate the standard deviations from the mean as about 0.5% up to 2%. This means that the variance of parameters introduced by production (elasticity, density, thickness uniformity, misalignment of the reinforcing fibre etc.) and errors introduced by the preparation of the sample (geometrical errors, errors of density and weight estimation, microdamages created by sample cutting) have determining influence on the variance of identified elastic parameters. Here we will use the term „noisy constants" for parameters that are taken as constants in FEM calculations, but that contain uncertainty for the actual physical specimens.

3.2 Discrepancy minimization approach in the presence of noisy constants

Let us define the following variables:

$\boldsymbol{E}$ – column vector of size n of elastic parameters whose numerical values must be identified. In the simplest case of orthotropic material plate, this vector consists of 4 components (n = 4): longitudinal modulus E_x, transverse modulus E_y, shear modulus G_{xy} and Poisson's ratio ν_{xy}.

$$\boldsymbol{E} = \left[E_x, E_y, G_{xy}, \nu_{xy}\right]^{\mathrm{T}}, \tag{9}$$

$\boldsymbol{P}$ – column vector of size m of *noisy constants* – parameters that are taken into account as constants both in FEM and in approximated analytical calculations of eigenfrequencies, but physically are parameters with uncertainty. In the simplest case of rectangular plate, the vector $\boldsymbol{P}$ consists of four components (m = 4): plate length a, plate width b, plate total thickness h and the average material density ρ:

$$\boldsymbol{P} = [a, b, h, \rho]^{\mathrm{T}}. \tag{10}$$

$\boldsymbol{f}^{\mathrm{EXP}}$ – column vector of measured L lowest natural frequencies f_i^{EXP}, i = 1,2,…, L.

$\hat{\boldsymbol{f}}$- column vector of the approximated model for the dependence of natural frequencies on material elastic parameters $\boldsymbol{E}$ and specimen parameters $\boldsymbol{P}$. This functional dependence can be obtained by metamodeling (approximation of FEM results) or using approximate analytical expressions:

$$\hat{\boldsymbol{f}} = \hat{\boldsymbol{f}}(\boldsymbol{E}, \boldsymbol{P}) \tag{11}$$

$\boldsymbol{W}$ – diagonal matrix of weighting coefficients:

$$\boldsymbol{W} = \begin{bmatrix} w_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & w_L \end{bmatrix} \tag{12}$$

The discrepancy function that measures the deviation between measured and numerically calculated frequencies in matrix form:

$$\Phi(\boldsymbol{E}, \boldsymbol{P}, \boldsymbol{W}) = (\boldsymbol{f}(\boldsymbol{E}, \boldsymbol{P}) - \boldsymbol{f}^{\mathrm{EXP}})^{\mathrm{T}} \boldsymbol{W} (\boldsymbol{f}(\boldsymbol{E}, \boldsymbol{P}) - \boldsymbol{f}^{\mathrm{EXP}}) \tag{13}$$

Most frequently used weighting coefficients vary inversely as the square of measured frequency $w_i = \left(1/f_i^{\mathrm{EXP}}\right)^2$, so that the discrepancy function is the sum of squared relative differences between measured and calculated frequencies [3,10 and others].

The discrepancy minimization approach means that for given values of specimen parameters $\boldsymbol{P}$ and weighting coefficients $\boldsymbol{W}$, input parameters $\boldsymbol{E}$ which give minimal value of functional Φ will be considered as identified values for parameters $\boldsymbol{E}^*$:

$$\boldsymbol{E}^* = \underset{\boldsymbol{E}}{\arg\min}\, \Phi(\boldsymbol{E}, \boldsymbol{P}, \boldsymbol{W}) \tag{14}$$

Because the simple third-order polynomial metamodels of FEM results give very small prediction error and the number of minimization variables is relatively small, the minimization of discrepancy function is easy and needs 1-3 seconds of PC processor time using the multistart random search method [8]. The main problem is not the discrepancy minimization, but the estimation and improvement of the accuracy of identification.

3.3 Calculation of the standard deviations of identified values

Let us assume that the sources of identification errors are:

1) Uncertainty of specimen's parameters $\boldsymbol{P}$. We assume that these uncertainties may be described by unbiased independent random errors conforming to normal probability density distribution with zero means and given standard deviations:

$$\boldsymbol{P} = \overline{\boldsymbol{P}} + \boldsymbol{\varepsilon}^P, \;\; \varepsilon_i^P = N(0, \sigma_i^P), \;\; i = 1,2,3,m \tag{15}$$

2) Additional errors of measured frequencies, caused by measurement errors and uncertainty of parameters which are not included in vector $\boldsymbol{P}$, for example dimensional inaccuracies like deviations of rectangularity, curvature, deviations of orthotropic axes and others. For simplicity, we consider these errors as unbiased independent random errors conforming to normal probability density distribution with zero means and given standard deviations:

$$\boldsymbol{f}^{\mathrm{EXP}} = \overline{\boldsymbol{f}}^{\mathrm{EXP}} + \boldsymbol{\varepsilon}^F, \;\; \varepsilon_i^F = N(0, \sigma_i^F), \;\; i = 1,2, \dots, L \tag{16}$$

3) Deviations of elasticity parameters themselves. We consider these deviations as unbiased independent random errors conforming to normal probability density distribution with zero means and given standard deviations:

$$\boldsymbol{E} = \overline{\boldsymbol{E}} + \boldsymbol{\varepsilon}^E, \;\; \varepsilon_i^E = N(0, \sigma_i^E) \tag{17}$$

4) Other sources of identification errors include the lack of correspondence between the mathematical model and the physical experiment. This may occur due to neglecting temperature effects, internal and external friction (air resistance and fastening effects), the use of inadequate FEM elements (homogeneous shell elements instead of layered structural shell elements for multilayer plate, etc.). These errors are inherently more systematic than random errors and can be discovered by analysis of means and standard deviations of the difference between measured and calculated values. Comparisons with the elastic parameter measurements obtained by other methods, for example bending tests are always useful.

5) Outliers - large errors made during physical and mathematical experimentation. The main source of this type of errors is the incorrect grouping of oscillation modes. All FEM software packages give the eigenfrequencies and corresponding modes in order of increasing frequency. The automatic mode recognition will be explained below.

After the minimization of the weighted discrepancy function Φ, the column vector of identified values $\boldsymbol{E}^*$ is found. For the analysis of the identification errors we will use the linearized model of the frequency dependence on elastic parameters $\boldsymbol{E}$ and specimen parameters $\boldsymbol{P}$ in the form

$$\boldsymbol{f}(\boldsymbol{E}, \boldsymbol{P}) = \boldsymbol{A}\boldsymbol{E} + \boldsymbol{B}\boldsymbol{P} + \boldsymbol{C} \tag{18}$$

where $\boldsymbol{A}$ – constant matrix $L \times n$, $\boldsymbol{B}$ – constant matrix $L \times m$, $\boldsymbol{C}$ – constant column vector of size L. Matrices $\boldsymbol{A} = \frac{\partial \hat{\boldsymbol{f}}}{\partial \boldsymbol{E}}(\boldsymbol{E}^*, \overline{\boldsymbol{P}})$, $\boldsymbol{B} = \frac{\partial \hat{\boldsymbol{f}}}{\partial \boldsymbol{P}}(\boldsymbol{E}^*, \overline{\boldsymbol{P}})$ and column vector $\boldsymbol{C} = \hat{f}(\boldsymbol{E}^*, \overline{\boldsymbol{P}}) - \boldsymbol{A}\boldsymbol{E}^* - \boldsymbol{B}\overline{\boldsymbol{P}}\boldsymbol{C}$ can be calculated analytically, if the metamodel is polynomial or by use of numerical differentiation for other types of metamodels, including kriging.

Then the discrepancy function assumes the form:

$$\boldsymbol{\Phi}(\boldsymbol{E},\boldsymbol{P},\boldsymbol{W}) = (\boldsymbol{f}(\boldsymbol{E},\boldsymbol{P}) - \boldsymbol{f}^{\mathrm{EXP}})^{\boldsymbol{T}}\boldsymbol{W}(\boldsymbol{f}(\boldsymbol{E},\boldsymbol{P}) - \boldsymbol{f}^{\mathrm{EXP}}) = \\ = (\boldsymbol{AE} + \boldsymbol{BP} + \boldsymbol{C} - \boldsymbol{f}^{\mathrm{EXP}})^{\boldsymbol{T}}\boldsymbol{W}(\boldsymbol{AE} + \boldsymbol{BP} + \boldsymbol{C} - \boldsymbol{f}^{\mathrm{EXP}}) \tag{19}$$

This is a second-order polynomial regarding components of vector $\boldsymbol{E}$, therefore the minimum of the discrepancy function can be found by equating partial derivatives to zero:

$$\frac{\partial \boldsymbol{\Phi}}{\partial \boldsymbol{E}}(\boldsymbol{E}^*,\boldsymbol{P},\boldsymbol{W}) = 0 \tag{20}$$

This gives a system of linear algebraic equations regarding the vector of parameters to be identified $\boldsymbol{E}^*$

$$\boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{W}\boldsymbol{A}\boldsymbol{E}^* + \boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{K}\boldsymbol{C} + \boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{W}\boldsymbol{B}\boldsymbol{P} - \boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{W}\boldsymbol{f}^{\mathrm{EXP}} = 0 \tag{21}$$

Introducing the designation for system matrix $\boldsymbol{R} = \boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{K}\boldsymbol{A}$, we obtain the expression for vector of identified values:

$$\boldsymbol{E}^* = \boldsymbol{R}^{-1}(\boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{W}\boldsymbol{f}^{\mathrm{EXP}} - \boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{W}\boldsymbol{C} - \boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{W}\boldsymbol{B}\boldsymbol{P}) \tag{22}$$

The size of the symmetrical matrix $\boldsymbol{R}$ is equal to the number of parameters to be identified. The determinant and condition number of matrix $\boldsymbol{R}$ is very important. Small value of the determinant signifies an ill-conditioned identification problem, which can occur by using small number of frequencies and by including simultaneously non-identifiable parameters (like density, ply thickness and elastic moduli) in the vector $\boldsymbol{E}$.
We introduce symbols for matrices:

$$\boldsymbol{Q} = \boldsymbol{R}^{-1}(\boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{W}\boldsymbol{B}) \tag{23}$$

and

$$\boldsymbol{S} = \boldsymbol{R}^{-1}(\boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{W}) \tag{24}$$

Taking into account assumptions about normally dispersed parameter errors and additional frequency errors, the standard deviation of k-th identifiable parameter can be calculated according to the statistical law for linear sum of random variables [11].

$$\mathrm{STD}(E_k^*) = \sqrt{\sum_{i=1}^{m}(Q_{ik}\sigma_i^P)^2 + \sum_{i=1}^{L}(S_{ik}\sigma_i^F)^2 + \sigma_k^E}\ , \quad k = 1,\ldots, n, \tag{25}$$

It also is not difficult to take in account the correlation between frequency and parameter errors, if the covariance matrices can be estimated.

4 PARAMETER IDENTIFICATION WITH ROBUSTNESS CONTROL

The identification consists of two-stage optimization. During the first stage, the metamodel is built for the dependence of natural frequencies on the elastic parameters $\boldsymbol{E}$ and noisy constants $\boldsymbol{P}$, and the unconstrained minimization of the discrepancy function is provided using fixed weighting coefficients $\boldsymbol{W}$ for frequencies. The metamodel is created on the basis of computer experiments, which are carried out according to space-filling Latin hypercube experimental designs [7]. Usually third-order polynomial approximations are used. The range of variation of noisy constants in the computer experiment should not exceed the real uncertainty level more than twice, because the following analysis of robustness assumes small

parameter deviations. During the creation of the metamodel, accurate mode recognition must be provided to exclude large errors of responses (outliers). In most cases the modal assurance criterion (MAC) based on Pearson correlation coefficient between two mode shapes allows fast and accurate mode grouping.

During the second stage, the robustness of the identification is checked. The calculation of the standard deviations of identified parameter values is based on the assumed values of deviations of noisy constants and natural frequency measurements. In reality these values are known only very approximately or not at all. In any case, the estimation of the deviations of parameters, used in physical measurements and FEM calculations, should be obtained with replicated measurements using a sufficiently large number of specimens. If the deviations of noisy constants (and the probability density functions for them) would be known with high precision, then the step of robust optimization could include the minimization of the standard deviation of identification by variation of weighting coefficients. Since in reality the deviations of noisy constants are known rather approximately, then in practice robust minimization in this case means simply the choice – to use or not to use some measured frequency in the discrepancy function. But the main significance of robustness analysis is the ability to compare the precision of different methods (using different modes, non-dimensionalization and scaling of frequencies) with approximately identical deviations of measured parameters.

After the analysis of the standard errors of identification and the choice of a different frequency set for identification, the first stage must be repeated. This process (discrepancy minimization – STD minimization) usually converges quickly.

5 FREQUENCY APPROXIMATION AND NONDIMENSIONALIZATION

The frequently used simple approximate analytical formula for natural frequencies of orthotropic plates with completely free boundary conditions is given by Dickinson [9]. He applied characteristic beam functions in Rayleigh's method to obtain an approximate formula for the flexural vibration of specially orthotropic plates. (Specially orthotropic means that the orthotropic axes are parallel to rectangular plate edges).

Using the four elasticity parameters (E_x, E_y, G_{xy}, ν_{xy}), three dimensional parameters a, b, h and average density parameter ρ, the formula assumes the form:

$$f_{ij} = \frac{\pi}{2\sqrt{12}\sqrt{1-\nu_x\nu_x E_y/E_x}} \cdot \frac{h}{\sqrt{\rho}} \left(\frac{G_i^4 E_x}{a^4} + \frac{G_j^4 E_y}{b^4} + \frac{2H_i H_j \nu_x E_y}{a^2 b^2} + \frac{4J_i J_j G_{xy}\left(1-\nu_x\nu_x E_y/E_x\right)}{a^2 b^2} \right)^{1/2} \quad (26)$$

where f_{ij} is the natural frequency, measured in Hz, i and j are the number of halfwaves of mode shape in x and y direction respectively. The dimensionless parameters G, H, and J are functions of the indices i and j and can be taken from literature [9]. Comparison with the FEM calculations shows that the error of analytical approximation is about 4% for the case of homogenous plate. The analytical expressions of natural frequencies therefore cannot be used for accurate identification of elastic properties, but they can help to analyze the nature and sources of identification errors.

As can be seen from equation (26), all frequencies have the coefficient $h/\sqrt{\rho}$ and the multiplication of all three elasticity moduli with any constant coefficient c causes the

increasing of frequency $\sqrt{c}$ times. From this follows that it is not possible to simultaneously identify plate material density or thickness and elasticity parameters. This effect is called isospectral family – systems with different geometrical and elasticity parameters but equal all natural frequencies [12].

5.1 Partial nondimensionalization of natural frequencies.

Let us assume that independent standard deviations of frequencies are proportional to mean values of frequencies. As can be seen, in approximated analytical expressions (26) the plate thickness gives a multiplier $\sqrt{h^3}$ for all frequencies, if the density is calculated as division of mass by volume $\rho = M/abh$. Therefore it would be very desirable to use not the frequency values, but the ratios between them, for example dividing all first L frequency values by the sum of frequencies:

$$\tilde{f}_i = \frac{f_i}{\sum_{j=1}^{L} f_j}, i = 1,2,\dots,L \tag{27}$$

This nondimensionalization can be carried out for physically measured and numerically calculated frequency values. Obviously, the nondimensional frequencies of this type (at least for analytical model Eq.(26)) do not depend on the plate thickness and are not affected by thickness errors. The drawback is that the elastic moduli cannot be determined uniquely – the frequencies $\tilde{\boldsymbol{f}}$ build an isospectral family. As can be seen from expression (26), multiplying the values of moduli E_x, E_y, G_{xy} by the same coefficient will give exactly the same values of scaled frequencies $\tilde{\boldsymbol{f}}$. Numerical experiments with FEM metamodels showed that the use of scaled frequencies gives very accurate identification of Poison's ratio and the values of elastic moduli are determined with an accuracy of common unknown coefficient. This means that the values of nondimensional ratios E_y/E_x and G_{xy}/E_x can be determined very accurately. Therefore for practical use in identification, the lowest frequency was used without scaling:

$$\tilde{f}_1 = f_1,\ \tilde{f}_i = \frac{f_i}{\sum_{j=1}^{L} f_j}, i = 2,3,\dots,L \tag{28}$$

The lowest frequency for the Dickinson's analytical frequency model [9] is the torsional mode $f_{2,2}$

$$f_1 = f_{2,2} = \frac{1.10279h\sqrt{G_{xy}}}{ab\sqrt{\rho}} = \frac{1.10279\sqrt{h^3}\sqrt{G_{xy}}}{\sqrt{Mab}} \quad . \tag{29}$$

6 EXAMPLE OF IDENTIFICATION

The designs of computer experiments optimized according to the MSE space filling criterion [7], the polynomial and kriging metamodels and the optimization software EDAOpt [8] have been used for the identification of composite material properties for more than 10 years [3]. In the present study we demonstrate the use of the proposed method for the example problem and frequency measurements of P. Pedersen and P.S. Frederiksen [14], that have been used by many authors [4,10,13]. They measured the first ten natural frequencies of a thin glass/epoxy composite laminate with a stacking sequence of $[0,-40,40,90,40,0,90,-40]_s$. We seek the four in-plane ply-elastic constants of a thin composite laminate: E_x, E_y, G_{xy} and ν_{xy}. The rectangular plate had the dimensions (length a, width b and thickness h) given in Table 1.

Normal probability density distributions were used for uncertainties of noisy constants a, b, h, ρ. The plate was attached by two strings which were assumed to be modeled appropriately by free boundary conditions. The measured frequencies are provided in Table 2.

Table 1: Plate properties: length (a), width (b), thickness (h) and density (ρ) and their standard deviations

Parameter	a (mm)	b (mm)	h (mm)	ρ (kg/m^3)
Mean value	209	192	2.59	2120
Standard deviation	0.25	0.25	0.01	10.6

Table 2: Measured natural frequencies [12] and assumed standard deviations [4]

Frequency	f_1	f_2	f_3	f_4	f_5	f_6	f_7	f_8	f_9	f_{10}
Value (Hz)	172.5	250.2	300.6	437.9	443.6	760.3	766.2	797.4	872.6	963.4
STD %	0.5	0.52	0.54	0.58	0.585	0.686	0.687	0.697	0.721	0.75

The 301-point 8-factor sequential MSE-optimized design [7] was used for metamodel building. The 8 factors were four elastic parameters E_x, E_y, G_{xy}, ν_{xy} and four noisy constants a, b, h, ρ.

For FEM calculations the software ANSYS was used with elements SHELL 281. Third-order polynomial approximations were used for metamodeling. The oscillation modes 4-5 and 6-7 were mixed up (interchanged) for many points of experimental design in the output file of ANSYS. The relative cross-validation errors for frequencies 4,5, and 6, 7 with mixed modes were 3.9% and 10.2% respectively. After automatic mode recognition the cross-validation errors decreased to 0.03%, except for modes 2,3 and 9,10 which had CVE of about 0.2%. It must be noted that standard deviation from the mean value of frequency responses are about 25 times smaller than the mean value. Therefore if the relative error would be calculated in relation to the mean value, the numbers would be correspondingly 25 times smaller, that is, approximately 0.008%. This means that the approximation error is at least 100 times smaller than the error of physical measurements.

Using partially nondimensionalized frequencies the mean difference between measured and calculated frequencies for identified parameter values were about ±0.2%.

Table 3 shows the comparison of identification results obtained by different authors. As can be seen, the estimated standard deviation of Poison's ratio when using dimensioned frequencies is 59%. This means that using discrepancy minimization with dimensioned Hz frequencies it is not possible to identify the Poison's ratio if the frequency measurement standard deviation exceeds 0.5%. At the same time, using the partially scaled frequency approach, the estimated standard deviation for Poison's ratio is only 2.6%, which means that the accuracy of identification is very high.

In the publications of other authors regarding the identification of this plate, the evaluation of the standard deviation of identified parameters is given only in [4]. The present results of identification within 1-2 standard deviations overlap with the results given in [4], but the results obtained with the partial scaling method have 2 times smaller standard deviation for elasticity moduli and 5 times smaller STD for Poison's ratio. Therefore identification with the

robust optimization approach is much more precise if it is assumed that the FEM model gives sufficiently accurate frequency calculations without systematic error.

Table 3: Comparison of identification results obtained by different authors

Method	E_x STD CV%	E_y STD CV%	G_{xy} STD CV%	ν_{xy} STD CV%
Present Hz	62.81 4.41 7.01	20.97 2.88 13.7	9.47 2.39 25.2	0.246 0.14 59
Present scaled	65.40 1.15 1.75	22.63 0.42 1.84	8.39 0.21 2.51	0.186 0.005 2.6
[14]	61.3 ----- ------	21.4 ------ -----	9.8 ----- -----	0.280 ----- -----
[4]	60.8 1.85 3.05	21.3 1.16 5.46	9.87 0.59 5.96	0.27 0.034 12.2
[13]	56.5 ----- ------	20.8 ------ -----	11.8 ----- -----	0.349 ----- -----
[10]	57.2 ----- -----	21.4 ------ -----	11.3 ----- -----	0.300 ----- -----

7 CONCLUSIONS

- The use of the natural frequency metamodel with „noisy constants" included in the set of input variables and weighted discrepancy minimization method allows obtaining the estimated values for standard deviations of identified parameters.
- The use of calculated standard deviations of identified parameters allows the formulation of the identification problem in the form of robust minimization – simultaneous minimization of the discrepancy function and the standard deviations of determined values by variation of unknown values of elastic parameters and weighting coefficients.
- The third-order polynomial approximations for the dependence of natural frequencies (calculated by FEM) on the elastic parameters and „noisy constants" gives the prediction error of metamodel, which is approximately 100 to 200 times less than the frequency errors caused by uncertainty of geometrical and physical parameters of the composite material specimens.
- The use of partially scaled frequency values allows to reduce the standard deviations of determined values of elastic moduli about two times and the STD value of Poisson's ratio about 20 times in comparison with the use of dimensioned frequency values.
- The relative standard deviations of all identified values of elastic parameters are typically about 3-4 times larger than the relative standard deviations of physically measured natural frequencies. The accuracy of elastic moduli identification was about 4%, Poisson's ratio – about 14% for 95% confidence level. The method proved that the Poisson's ratio for plate type specimens can be determined using only out-of-plane bending modes with about 5% error for 95% confidence level, when frequency standard deviations are about 1%.

Acknowledgements

This work has been supported by the European Social Fund within the Project No. 2013/0025/1DP/1.1.1.2.0/13/APIA/VIAA/019 "New "Smart" Nanocomposite Materials for Roads, Bridges, Buildings and Transport Vehicles".

REFERENCES

[1] Larrson, D. Using modal analysis for estimation of anisotropic material constants. *Journal of Engineering Mechanics*, Vol. **123**, No. **3**, 222-229, (1997).

[2] Lauwagiea, T., Solb, H.W., Roebbenc, G. Determination of the in-plane elastic properties of the different layers of laminated plates by means of vibration testing and model updating. *Journal of Sound and Vibration*, **274**, 529–546, (2004).

[3] Rikards, R. and Auzins, J. Response Surface Method for Solution of Structural Identification Problems, *Inverse Problems in Science and Engineering*, Vol. **12**, No. 1, , pp. 59-70. (2004).

[4] Gogu, C., Haftka, R., Le Riche, R., Molimard, J. Effect of approximation fidelity on vibration-based elastic constants identification. *Structural and Multidisciplinary Optimization*, Vol. **42**, Issue **2**, (August 2010), 293-304.

[5] Smarslok, B.P., Haftka, R.T., Ifju, P.G. Correlation model for composite elastic properties using only measurements from a single specimen. *Probabilistic Engineering Mechanics*, **29**, (2012), pp. 64-69.

[6] Withers, C.S. and Nadarajah, A. Normal maximum likelihood, weighted least squares, and ridge regression estimates. *Probability and mathematical statistics*, Vol. **32**, fasc. 1 (2012), pp. 11–24.

[7] Auzins, J. Direct Optimization of Experimental Designs, 10th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference, AIAA Paper 2004-4578, CD-ROM Number 17, (2004).

[8] Auzins, J., Janushevskis, A., Janushevskis, J. and Skukis, E. Software EDAOPT For Experimental Design, Analysis And Multiobjective Robust Optimization. OPT-i International Conference on Engineering and Applied Sciences Optimization, M. Papadrakakis, M.G. Karlaftis, N.D. Lagaros (eds.) Kos Island, Greece, 4-6 June 2014. CD-edition, ISSN: 2241-9098, ISBN: 978-960-99994-5-8, (2014).

[9] Dickinson, S.M. The buckling and frequency of flexural vibration of rectangular isotropic and orthotropic plates using Rayleigh's method, *Journal of Sound and Vibration*, Vol. **61**, No. **1**, 1-8, 1978.

[10] Mota Soares, C.M., Moreira de Freitas, M., Araújo, A.L., Pedersen, P. Identification of material properties of composite plate specimens, *Composite Structures*, Vol. **25**, 277-285, (1993).

[11] DeVore, J.L. *Probability and Statistics for Engineering and Science*, 8th edition, (Duxbury Press, 2011).

[12] Gladwell, G.M.L. *Inverse Problems in Vibration*. Second Edition, Kluwer Academic Publishers, 2005.

[13] Rikards, R., Chate, A., Steinchen, W., Kessler, A., Bledzki, A.K. Method for identification of elastic properties of laminates based on experiment design. *Composites: Part B*, **30**, pp. 279–289. (1999).

[14] Pedersen, P., Frederiksen, P.S. Identification of orthotropic material moduli by a combined experimental/numerical method. *Measurement*, Vol. **10**, Issue **3**, (July–September 1992), 113–118.

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis(Eds)

SIMULATION OF SELF-HEALING PROCESSES IN MICROCAPSULE BASED SELF-HEALING POLYMERIC SYSTEMS

Steffen Specht*, Joachim Bluhm* and Jörg Schröder*

*Institute of Mechanics
University of Duisburg-Essen
Universitätsstr. 15, 45141 Essen, Germany
e-mail: steffen.specht@uni-due.de, web page: http://www.uni-due.de/mechanika

Key words: Self-healing materials, Theory of Porous Media, Phase transition

Abstract. Self healing materials are becoming more and more important for the construction of mechanical components due to their ability to detect and heal failures and cracks autonomously. Especially in polymers and polymer-composites, where the component can loose a high rate of strength and durability due to micro cracks, those damages are nearly impossible to repair from outside. Thus, self healing ability is a very effective approach to extend the lifetime of polymer-made components.

In view of the numerical simulation of such self healing effects we develop a thermodynamically consistent macroscopic 5-phase model within the theoretical framework of the Theory of Porous Media. The model consists of the following different phases: solid (matrix material) with dispersed catalysts, liquid (healing agents), healed material and gas (air inside the cracks). The increase of damage is driven by a discontinuous damage evolution equation. Furthermore, a mass exchange between the liquid-like healing agents and the solid-like healed material, i.e. the change of the aggregate state from liquid healing to solid healed material, describes the healing process. The onset of the healing process is associated with the break open of the microcapsules in connection with the subsequent motion of the liquid healing agents. A numerical example of the simulation of damage and healing processes in polymers, is presented in order to show the applicability of the model.

1 INTRODUCTION

Self-healing materials have the ability to repair damages autonomously, i.e., no manual intervention is necessary. Such a behavior can be realized for example with a multiphase material, which consists of a solid matrix material (epoxy) with dispersed catalysts (Grubbs' catalysts) and microcapsules, which are filled with liquid healing agents (dicyclopentadiene). If a crack propagates through the body, the capsules break open and release the healing agents into the crack, where they react with the catalysts, polymerize and close the crack, cf. [1]. Such a self-healing approach is very useful for components where the detection and repair is very difficult

and/or very costly like in aerospace applications. For an overview of different self-healing materials can be found in [2], whereas [3] and [4] show the different principles of self-healing.

In view of the numerical simulation of self-healing materials, several examples can be found in literature. Some of these publications are listed in the following:

Self healing effects of polymers in an analytical manner where analyzed by [5]. Regarding fiber reinforced composites the self healing behavior was investigated and simulated by [6–8]. In view of the development of thermodynamically consistent models it is referred to [9–11]. In [12] the simulations base on Continuum Damage Mechanics (CDM), and [13] taking continuous damage and healing variables into account. The model of [14] is based on the Mixture Theory.

In this contribution we focus on the numerical simulation of self-healing materials using the Theory of Porous Media, as described in the following.

2 THEORY

For the description of coupled multiphase problems, the Theory of Porous Media (TPM) can be used. This macroscopic continuum mechanical approach is basically a combination of the Mixture Theory and Concept of Volume Fractions.

Using the the Mixture Theory, one is able to describe, for example, all κ constituents φ^α of a multiphase problem by their own independent motion function. Furthermore, the superposition of the phases is assumed, i.e., all different constituents appear in a spatial point $\mathbf{x}$ simultaneously, see Figure 1. Due to that, all geometrical and physical quantities are defined as statistically averages of the real quantities in the observed body.

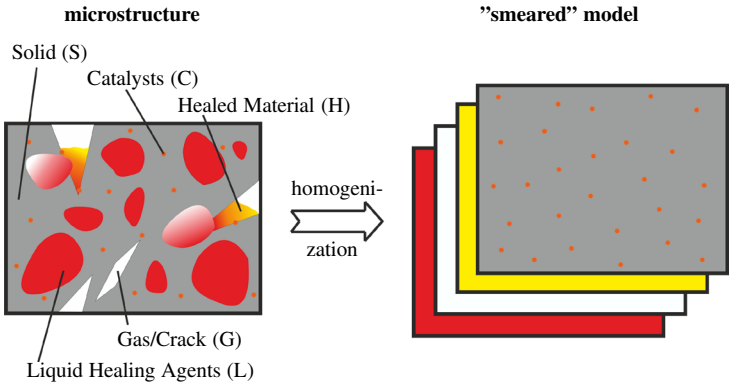


Figure 1: Homogenization of the microstructure.

Additionally, with help of the concept of volume fractions, the different constituents φ^α in a material point $\mathbf{x}$ can be identified by the volume fraction of the corresponding phase, which relates the real quantities with the partial quantities. For example, the partial density ρ^α of a certain constituent can be directly related to its real density $\rho^{\alpha\mathrm{R}}$ using the corresponding volume

fraction $\mathrm{n}^\alpha = \mathrm{dv}^\alpha/\mathrm{dv}$, such that $\rho^\alpha = \mathrm{n}^\alpha \rho^{\alpha\mathrm{R}}$. The sum over all κ volume fractions n^α in a material point $\mathbf{x}$ is restricted by the so called saturation condition

$$\sum_{\alpha=1}^{\kappa} \mathrm{n}^\alpha = 1\,. \tag{1}$$

For a detailed introduction into the Theory of Porous Media, the interested reader is referred to the publications [15–20].

3 SIMPLIFIED FIVE-PHASE MODEL

In the following the developed five phase model for the description of the considered self healing material will be presented. It consists of the solid matrix material (S) with dispersed catalysts (C), the liquid healing agents (L), the solid like healed material (H), and the gas phase (G), which represents the air. In order to build up the model, some assumptions and simplifications are made: 1) the whole is treated as isothermal; 2) dynamic effects are neglected; 3) mass transition will be considered only between the liquid like healing agents and the solid like healed material, in order to describe the phase transition of the healing material; 4) all phases are assumed to be incompressible ($\rho^{\alpha\mathrm{R}} = \text{const.}$), except the gas phase which is compressible ($\rho^{\mathrm{GR}} \neq \text{const.}$); 5) the volume fraction of the catalysts is neglected with respect to the saturation condition, due to the fact that it is very small in comparison with the other phases; 6) the velocities of the solid and the solidified healed material are assumed to be identical, except at an initial solid motion, i.e., before the healing mechanism is activated. This leads to a multiplicative decomposition of the deformation gradient, depicted in Figure 2, in the form

$$\mathbf{F}_\mathrm{S} = \mathrm{Grad}\,\boldsymbol{\chi}_\mathrm{S} = \frac{\partial \boldsymbol{\chi}_\mathrm{S}}{\partial \mathbf{X}_\mathrm{S}} = \mathbf{F}_\mathrm{H}\,\mathbf{F}_{\mathrm{S}0}\,, \tag{2}$$

which can be found in [21–23]. Due to the multiplicative decomposition of the deformation gradient, three different right Cauchy-Green deformation tensors (corresponding to the whole deformation, the deformation before and after the healing is activated) are available,

$$\mathbf{C}_\mathrm{S} = \mathbf{F}_\mathrm{S}^\mathrm{T}\,\mathbf{F}_\mathrm{S}\,, \qquad \mathbf{C}_{\mathrm{S}0} = \mathbf{F}_{\mathrm{S}0}^\mathrm{T}\,\mathbf{F}_{\mathrm{S}0}\,, \qquad \hat{\mathbf{C}}_\mathrm{H} = \mathbf{F}_\mathrm{H}^\mathrm{T}\,\mathbf{F}_\mathrm{H}\,. \tag{3}$$

3.1 Field Equations

Considering the above mentioned assumptions and simplifications, the field equations are given by the balance equation of mass for the solid, healed material, liquid healing agents, catalysts and gas,

$$(\mathrm{n}^\mathrm{S})'_\mathrm{S} + \mathrm{n}^\mathrm{S}\,\mathrm{div}\,\mathbf{x}'_\mathrm{S} = 0\,, \qquad (\mathrm{n}^\mathrm{H})'_\mathrm{S} + \mathrm{n}^\mathrm{H}\,\mathrm{div}\,\mathbf{x}'_\mathrm{S} = \frac{\hat{\rho}^\mathrm{H}}{\rho^{\mathrm{HR}}}\,, \tag{4}$$

$$(\mathrm{n}^\mathrm{L})'_\mathrm{L} + \mathrm{n}^\mathrm{L}\,\mathrm{div}\,\mathbf{x}'_\mathrm{L} = -\frac{\hat{\rho}^\mathrm{H}}{\rho^{\mathrm{LR}}}\,, \qquad \mathrm{n}^\mathrm{S}\,(\mathrm{c}^\mathrm{C})'_\mathrm{S} - \mathrm{div}\,(\mathrm{n}^\mathrm{S}\,\mathrm{c}^\mathrm{C}\,\mathbf{w}_{\mathrm{CS}}) = \hat{\rho}^\mathrm{C}\,,$$

$$(\mathrm{n}^\mathrm{G})'_\mathrm{G} + \mathrm{n}^\mathrm{G}\,\mathrm{div}\,\mathbf{x}'_\mathrm{G} + \frac{\mathrm{n}^\mathrm{G}}{\rho^{\mathrm{GR}}}\,(\rho^{\mathrm{GR}})'_\mathrm{G} = 0\,,$$

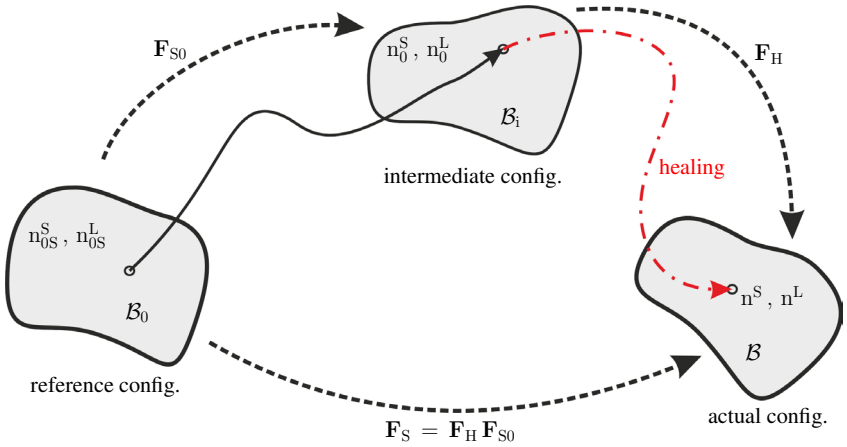


Figure 2: Illustration of the multiplicative decomposition of the deformation gradient of the solid phase.

the balance equations of momentum for the mixture as well as for the liquid and gas phases,

$$\operatorname{div} \mathbf{T}^{\mathrm{SHLCG}} + \rho^{\mathrm{SHLCG}} \mathbf{b} = -\hat{\rho}^{\mathrm{H}} \mathbf{w}_{\mathrm{LS}} , \qquad \operatorname{div} \mathbf{T}^{\mathrm{L}} + \rho^{\mathrm{L}} \mathbf{b} = -\hat{\mathbf{p}}^{\mathrm{L}} , \qquad \operatorname{div} \mathbf{T}^{\mathrm{G}} + \rho^{\mathrm{G}} \mathbf{b} = -\hat{\mathbf{p}}^{\mathrm{G}} , \tag{5}$$

and the material time derivative of the saturation condition with respect to the solid phase,

$$\operatorname{div} (\mathrm{n}^{\mathrm{L}} \mathbf{w}_{\mathrm{LS}} + \mathrm{n}^{\mathrm{G}} \mathbf{w}_{\mathrm{GS}} + \mathbf{x}_{\mathrm{S}}') + \frac{\mathrm{n}^{\mathrm{G}}}{\rho^{\mathrm{GR}}} (\rho^{\mathrm{GR}})_{\mathrm{G}}' - \hat{\rho}^{\mathrm{H}} (\frac{1}{\rho^{\mathrm{HR}}} - \frac{1}{\rho^{\mathrm{LR}}}) = 0 . \tag{6}$$

With the symbol $(\dots)_\alpha'$, the material time derivative of the expression with respect to the corresponding constituent φ^α is indicated and $\hat{(\dots)}$ denotes the direct production terms, in this case of mass and momentum, respectively. The relative velocities $\mathbf{w}_{\zeta\mathrm{S}} = \mathbf{x}_\zeta - \mathbf{x}_\mathrm{S}$ $(\zeta = \mathrm{L}, \mathrm{C}, \mathrm{G})$ are the difference velocities between the phases ζ and the solid phase. The value $\mathrm{c}^{\mathrm{C}} \in [0, 1]$ indicates the concentration of catalysts and $\mathbf{T}^\alpha = (\mathbf{T}^\alpha)^{\mathrm{T}}$ are the symmetric Cauchy stress tensors for the different constituents. The expressions $\mathbf{T}^{\mathrm{SHLCG}}$ and ρ^{SHLCG} describe the sum of the corresponding Cauchy stresses and partial densities, respectively, of the individual phases.

3.2 Constitutive Relations

In order to be able to solve the problem, constitutive relations for the stresses and the total production terms of mass and momentum are needed. In the following, just the final equations are given. For the derivation of these equations, it is referred to [24].

The constitutive relations for the Cauchy stresses read

$$\begin{aligned} \mathbf{T}^{\mathrm{SH}} = &- \mathrm{n}^{\mathrm{SH}} \lambda \mathbf{I} + \frac{1}{\mathrm{J_S}} \{ (1 - D^{\mathrm{S}}) [2 \mu^{\mathrm{S}} \mathbf{K}_{\mathrm{S}} + \lambda^{\mathrm{S}} (\log \mathrm{J_S}) \mathbf{I}] + \\ &+ \epsilon^{\mathrm{H}} \mathrm{n}^{\mathrm{H}} \mathrm{J_S} (1 - D^{\mathrm{H}}) [2 \mu^{\mathrm{H}} \mathbf{K}_{\mathrm{H}} + \lambda^{\mathrm{H}} (\log \mathrm{J_H}) \mathbf{I}] \} , \end{aligned} \tag{7}$$

$$\mathbf{T}^{\mathrm{L}} = - \mathrm{n}^{\mathrm{L}} \mathrm{p}^{\mathrm{LR}} \mathbf{I} , \quad \mathbf{T}^{\mathrm{G}} = - \mathrm{n}^{\mathrm{G}} \mathrm{p}^{\mathrm{GR}} \mathbf{I} , \quad \mathbf{T}^{\mathrm{C}} = - \mathrm{n}^{\mathrm{C}} \mathrm{p}^{\mathrm{CR}} \mathbf{I} .$$

In Eq. $(7)_1$, the variables $\mathrm{J_S} = \det \mathbf{F}_\mathrm{S}$ and $\mathrm{J_H} = \det \mathbf{F}_\mathrm{H}$ represent the volume deformations of the solid and healed material, respectively. The expressions $\mathbf{K}_\mathrm{S} = \frac{1}{2}(\mathbf{B}_\mathrm{S} - \mathbf{I})$ and $\mathbf{K}_\mathrm{H} = \frac{1}{2}(\mathbf{B}_\mathrm{H} - \mathbf{I})$ are the Karni-Reiner strain tensors, whereat $\mathbf{B}_\mathrm{S} = \mathbf{F}_\mathrm{S}\mathbf{F}_\mathrm{S}^\mathrm{T}$ and $\mathbf{B}_\mathrm{H} = \mathbf{F}_\mathrm{H}\mathbf{F}_\mathrm{H}^\mathrm{T}$ are the left Cauchy-Green strain tensors of the solid and healed material, and $\mathbf{I}$ is the identity tensor. The material parameters μ^S, μ^H and λ^S, λ^H are the Lamé constants. In order to ensure that the healed material part of the Cauchy stresses get its full influence only if the liquid is completely transformed into healed material, the parameter ϵ^H is chosen such that the product $\epsilon^\mathrm{H}\mathrm{n}^\mathrm{H}$ is equal to one if n^H reaches its maximum. The variables D^S and D^H are damage variables in order to describe the isotropic discontinuous damage behavior of the solid and also the healed material. The so called $(1 - D)$ approach was originally introduced in [25] and further discussed, e.g. in [26, 27]. The Lagrange parameter λ in Eq. $(7)_1$ is defined as

$$\lambda = \mathrm{p}^\mathrm{GR} - \mathrm{p}^\mathrm{h}, \tag{8}$$

where p^GR is the real gas pressure given by the nonlinear gas law

$$\mathrm{p}^\mathrm{GR} = -\Theta\,\mathrm{R}^\mathrm{G}\,\rho_\mathrm{0G}^\mathrm{GR}\,\log\frac{\rho_\mathrm{0G}^\mathrm{GR}}{\rho^\mathrm{GR}} + \mathrm{p}_0^\mathrm{GR}, \tag{9}$$

and an additional pressure p^h, which is alligned to the capillary pressure presented in [28],

$$\mathrm{p}^\mathrm{h} = k_\mathrm{h}^\mathrm{L}\,\mathrm{s}^\mathrm{L}\left[\log\left(\frac{\mathrm{s}_0^\mathrm{L}}{\mathrm{s}^\mathrm{L}} - \mathrm{s}_0^\mathrm{L}\right) - \log\left(1 - \mathrm{s}_0^\mathrm{L}\right)\right]. \tag{10}$$

The absolute temperature is given by Θ, R^G denotes the specific gas constant, $\rho_\mathrm{0G}^\mathrm{GR}$ and p_0^GR are initial real gas density and the initial real gas pressure. The constants k_h^L and s_0^L are material parameters and the liquid saturation $\mathrm{s}^\mathrm{L} = \mathrm{n}^\mathrm{L}/(\mathrm{n}^\mathrm{L} + \mathrm{n}^\mathrm{G})$ is the ratio of liquid with respect to the whole hollow space inside the observed body. Furthermore, the real liquid pressure p^LR, appearing in Eq. $(7)_2$, is given by

$$\mathrm{p}^\mathrm{LR} = \mathrm{p}^\mathrm{GR} + k_\mathrm{h}^\mathrm{L}\left[\log\left(\frac{\mathrm{s}_0^\mathrm{L}}{\mathrm{s}^\mathrm{L}} - \mathrm{s}_0^\mathrm{L}\right) - \log\left(1 - \mathrm{s}_0^\mathrm{L}\right)\right], \tag{11}$$

see also [28]. The pressure part of Eq. $(7)_3$ is the real pressure of catalysts,which is given by

$$\mathrm{p}^\mathrm{CR} = -k^\mathrm{C}\,\log\frac{1}{\mathrm{c}^\mathrm{C}} + \mathrm{p}_0^\mathrm{CR}, \tag{12}$$

whereat the concentration c^C is defined as the quotient of the volume fraction of the catalysts with respect to the volume fraction of the solid $\mathrm{c}^\mathrm{C} = \mathrm{n}^\mathrm{C}/\mathrm{n}^\mathrm{S}$.

For the description of mass exchange between the liquid healing agents and the solid healed material $\hat{\rho}^\mathrm{H}$, see equation $(4)_{2,3}$, the production function proposed in [29] is used and modified such that it depends on the concentration of catalysts,

$$\hat{\rho}^\mathrm{H} = \hat{\rho}_\mathrm{m}^\mathrm{H}\left(\frac{\mathrm{c}^\mathrm{C} - \mathrm{c}_0^\mathrm{C}}{\mathrm{c}_m^\mathrm{C}}\right)^2 \exp\left[1 - \left(\frac{\mathrm{c}^\mathrm{C} - \mathrm{c}_0^\mathrm{C}}{\mathrm{c}_\mathrm{m}^\mathrm{C}}\right)^2\right]. \tag{13}$$

Therein, $\hat{\rho}^{\mathrm{H}}_{\mathrm{m}}$ is the maximum value of $\hat{\rho}^{\mathrm{H}}$, and $\mathrm{c}^{\mathrm{C}}_{0}$ is the maximum value of the concentration. The parameter $\mathrm{c}^{\mathrm{C}}_{\mathrm{m}}$ defines the value of concentration where $\hat{\rho}^{\mathrm{H}}$ becomes its maximum. Due to the fact that the amount of catalysts decreases in areas where healing occur, the total production term of mass for the catalysts $\hat{\rho}^{\mathrm{C}}$ in Eq. $(4)_4$ is set to be a negative and constant value.

Furthermore, due to the evaluation of the entropy inequality, the direct production terms of momentum for liquid and gas are given by

$$\begin{aligned} \hat{\mathbf{p}}^{\mathrm{L}} &= \lambda \operatorname{grad} \mathrm{n}^{\mathrm{L}} - \mathrm{p}^{\mathrm{h}} \operatorname{grad} \mathrm{n}^{\mathrm{G}} + \hat{\mathbf{p}}^{\mathrm{L}}_{\mathrm{E}} \,, \\ \hat{\mathbf{p}}^{\mathrm{G}} &= \left(\lambda + \mathrm{p}^{\mathrm{h}} \right) \operatorname{grad} \mathrm{n}^{\mathrm{G}} - \hat{\mathbf{p}}^{\mathrm{G}}_{\mathrm{E}} \,. \end{aligned} \tag{14}$$

In Eq. (14) the vectors $\hat{\mathrm{p}}^{\mathrm{L}}_{\mathrm{E}}$ and $\hat{\mathrm{p}}^{\mathrm{G}}_{\mathrm{E}}$ denote the effective parts of the direct production terms of momentum, which are defined as

$$\hat{\mathbf{p}}^{\mathrm{L}}_{\mathrm{E}} = -\gamma^{\mathrm{L}}_{\mathbf{w}_{\mathrm{LS}}} \mathbf{w}_{\mathrm{LS}} - \gamma^{\mathrm{L}}_{\mathbf{w}_{\mathrm{GS}}} \mathbf{w}_{\mathrm{GS}} \,, \qquad \hat{\mathbf{p}}^{\mathrm{G}}_{\mathrm{E}} = -\gamma^{\mathrm{G}}_{\mathbf{w}_{\mathrm{GS}}} \mathbf{w}_{\mathrm{GS}} - \gamma^{\mathrm{G}}_{\mathbf{w}_{\mathrm{LS}}} \mathbf{w}_{\mathrm{LS}} \,, \tag{15}$$

whereat the occurring material parameters are restricted by

$$\gamma^{\mathrm{L}}_{\mathbf{w}_{\mathrm{LS}}} \geq 0 \,, \qquad \gamma^{\mathrm{G}}_{\mathbf{w}_{\mathrm{GS}}} \geq 0 \,, \qquad \gamma^{\mathrm{L}}_{\mathbf{w}_{\mathrm{GS}}} + \gamma^{\mathrm{G}}_{\mathbf{w}_{\mathrm{LS}}} = 0 \,. \tag{16}$$

4 NUMERICAL EXAMPLE

In order to show the applicability of the developed model, a numerical simulation of a real experiment, cp. [30], is carried out. The dimensions of the specimen and the damaged virtual specimen are depicted in Figure 3. It is discretized with 142 linear eight-nodular brick elements and the total number of degrees of freedom is 2492. On both flanks a displacement of $\mathrm{u} = 0.6$ mm is applied in y-direction. Moreover, the boundary surface at the beginning of the notch is open for the gas phase, i.e., air can flow in and out.

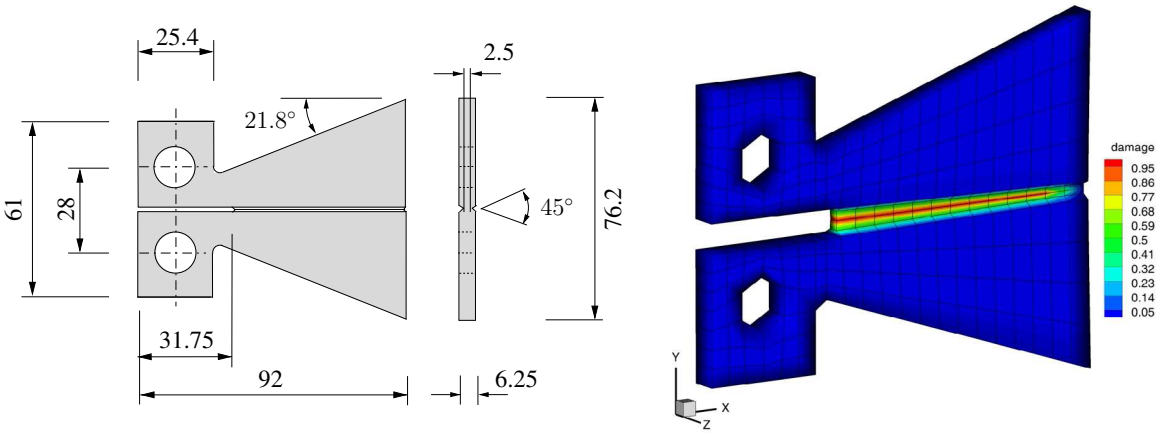


Figure 3: TDCB geometry, cp. [30] (left); damaged virtual specimen (right).

During the loading the TDCB fails. After the loading process the specimen is unloaded, because in the real experiment the specimen is able to heal, only if the crack faces come into

contact. Then, the TDCB gets 48 hours resting time before it is reloaded. As it is depicted in Figure 4, the results of the numerical simulation is qualitatively in a very good agreement to the experimental results of [30].

Table 1: Initial material parameters.

	S	H	L	G	C	unit
Young's modulus E^α	3.0e+9	3.0e+9	–	–	–	Pa
Poisson's ratio ν^α	0.2	0.2	–	–	–	–
real density $\rho^{\alpha\mathrm{R}}_{0\alpha}$	1200.0	980.0	980.0	1.0	–	kg/m^3
Darcy parameter k^α_{Darcy}	–	–	9.0e-9	5.0e+2	–	m^4/N s
Parameter associated with healing k^{L}_h	–	–	5.0e+1	–	–	Pa
initial volume fraction n^α	0.7	0.0	0.2	0.1	–	–
initial concentration c^α_0	–	–	–	–	1.0	$\times 100\%$
initial saturation $\mathrm{s}^{\mathrm{L}}_0$	–	–	0.9	–	–	–

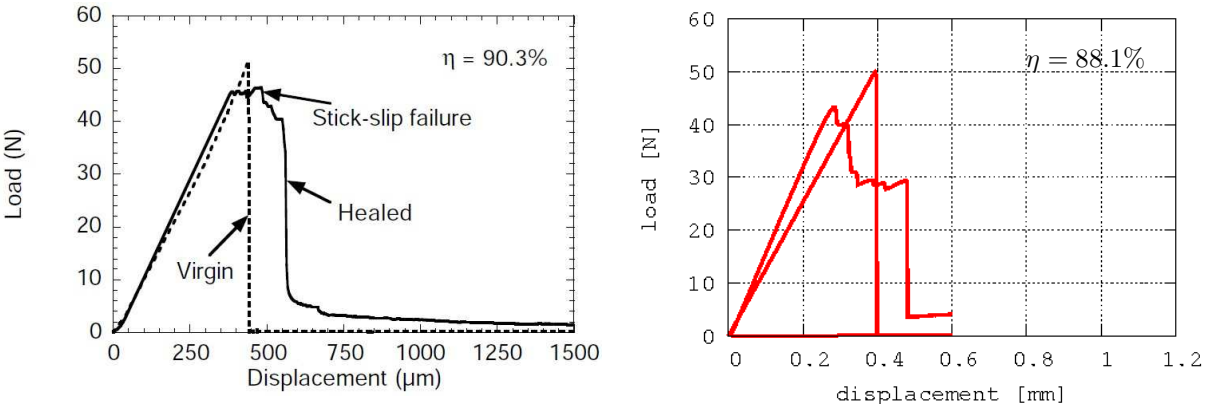


Figure 4: Experimental result, cp. [30] (left); result of the numerical simulation (right).

5 CONCLUSION

The presented work concentrates on the numerical simulation of damage as well as healing effects in a self-healing polymer composite. As the underlying theoretical framework the Theory of Porous Media is used. The developed multiphase model consists of the solid matrix material with dispersed catalysts, the liquid like healing agents, the solid like healed material, and the gas phase. For the separate description of damage for the solid and the healed material, two different damage functions are introduced based on the $(1 - D)$ approach. In order to describe the healing mechanism, a phase transition between the liquid healing agents and the solid healed material is considered.

To show the applicability of the developed model, the numerical simulation of a tapered double cantilever beam (TDCB) is compared with the experimental result from [30]. The simulation shows a qualitatively good agreement with the experimental observation, even for the healing efficiency. The divergence between both results can be explained, e.g., due to the fact that the healing of the real specimen depends on different factors, like wetting, different distributions of microcapsules and catalysts in the damaged area, etc.. Hence, it can be assumed that the resulting load-displacement curves of different specimens vary.

Acknowledgements This work has been supported by the German Research Society (DFG) within the Priority Program SPP 1568 "Design and Generic Principles of Self-healing Materials" under the grant number BL 417/7-1.

REFERENCES

[1] White, S.R., Sottos, N.R., Geubelle, P.H., Moore, J.S., Kessler, M.R., Sriram, S.R., Brown, E.N. and Viswanathan, S. Autonomic healing of polymer composites. *Nature* (2001) **409**:794–797.

[2] Grigoleit, S. *Überblick über Selbstheilende Materialien*. Tech. Report, Frauenhofer-Institut für Naturwissenschaftlich-Technische Trendanalysen (INT), (2011).

[3] Yuan, Y.C., Yin, T., Rong, M.Z. and Zhang, M.Q. Self healing polymers and polymer composites. Concept, realization and outlook: a review. *eXPRESS Polymer Letters* (2008) **2**:238–250.

[4] Hager, M.D., Greil, P., Leyens, C., van der Zwaag, S. and Schubert, U.S. Self-healing materials. *Advanced Materials* (2010) **22**:5424–5430.

[5] Zemskov, S.V., Jonkers, H.M. and Vermolen, F.J. Two analytical models for the probability characteristics of a crack hitting encapsulated particles: application to self-healing materials. *Computational Materials Science* (2011) **50**:3323–3333.

[6] Sanada, K., Itaya, N. and Shindo, Y. Self-healing of interfacial debonding in fiber-reinforced polymers and effect of microstructure on strength recovery. *The Open Mechanical Engineering Journal* (2008) **2**:97–103.

[7] Barbero, E.J. and Ford, K.J. Characterization of self-healing fiber-reinforced polymer-matrix composite with distributed damage. *Journal of Advanced Materials* (2007) **39**:20–27.

[8] Privman, V., Dementsov, A. and Sokolov, I. Modeling of self-healing polymer composites reinforced with nanoporous glass fibers. *Journal of Computational and Theoretical Nanoscience* (2007) **4**:190–193.

[9] Voyiadjis, G.Z., Shojaei, A. and Li, G. A thermodynamic consistent damage and healing model for self healing materials. *International Journal of Plasticity* (2011) **27**:1025–1044.

[10] Yagimli, B. and Lion, A. Experimental investigations and material modelling of curing processes under small deformations. *Zeitschrift fr angewandte Mathematik und Mechanik* (2011) **91**:342–359.

[11] Schimmel, E.C. and Remmers, J.J.C. *Development of a constitutive model for self-healing materials*. Tech. Report, Delft Aerospace Computational Science, (2006).

[12] Voyiadjis, G.Z., Shojaei, A., Li, G. and Kattan, P.I. A theory of anisotropic healing and damage mechanics of materials. *Proceedings of the Royal Society London A* (2012) **468**:163–183.

[13] Mergheim, J. and Steinmann, P. Phenomenological modelling of self-healing polymers based on integrated healing agents. *Computational Mechanics* (2013) **52**:681–692.

[14] Henson, G.M. Engineering models for synthetic vascular materials with interphase mass, momentum and heat transfer. *International Journal of Solids and Structures* (2012), **50**:2371–2382.

[15] Bowen, R.M. Incompressible porous media models by use of the theory of mixtures. *International Journal of Engineering Science* (1980) **18**:1129–1148.

[16] Bowen, R.M. Compressible porous media models by use of the theory of mixtures. *International Journal of Engineering Science* (1982) **20**:697–735.

[17] de Boer, R. *Trends in continuum mechanics of porous media*. Springer, (2005).

[18] de Boer, R. *Theory of porous media*. Springer, (2000).

[19] Ehlers, W. *Poröse Medien - ein kontinuummechanisches Modell auf der Basis der Mischungstheorie*. PhD.-Thesis, Universität - Gesamthochschule - Essen, (1989).

[20] Ehlers, W. and Bluhm, J. *(ed.)*, *Porous media*. Springer, (2002).

[21] Ateshian, G.A. and Ricken, T. Multigenerational interstitial growth of biological tissues. *Biomechanics and Modelling in Mechanobiology* (2010) **9**:689–702.

[22] Hunphrey, J. and Rajagopal, K. A constrained mixture model for growth and remodelling of soft tissues. *Mathematical Models and Methods in Applied Science* (2002) **12**:407–430.

[23] Rodriguez, E., Hoger, A. and McCulloch, A. Stress-dependent finite growth in soft elastic tissues. *Journal of Biomechanics* (1994) **27**:455–467.

[24] Bluhm, J., Specht, S., Schröder, J. Modeling of self-healing effects in polymeric composites. *Archive of Applied Mechanics* (2014) *DOI 10.1007/s00419-014-0946-7*.

[25] Kachanov, L.M. Time of the rupture process under creep conditions. *Izvestija Akademii Nauk Sojuza Sovetskich Socialisticeskich Republiki (SSSR) Otdelenie Techniceskich Nauk (Moskra)* (1958) **8**:26–31.

[26] Miehe, C. Discontinuous and continuous damage evolution in Ogden-type large-strain elastic materials. *European Journal of Mechanics, A/Solids* (1995) **14**:697–720.

[27] Balzani, D. *Polyconvex anisotropic energies and modeling of damage applied to arterial walls*. PhD.-Thesis, Universität Duisburg-Essen, (2006).

[28] Bluhm, J., Ricken, T. and Bloßfeld, W.M. Ice formation in porous media, in *Advances in extendet & multifield theories for continua*. Markert, B. *(ed.)* Springer, (2011).

[29] Michalowski, R.L. and Zhu, M. Frost heave modelling using porosity rate function. *International Journal for Numerical and Analytical Methods in Geomechanics* (2006) **30**:703–722.

[30] Brown, E.N., Sottos, N.R. and White, S.R. Fracture testing of a self-healing polymer composite. *Experimental Mechanics* (2002) **42**:372–379.

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis(Eds)

A FAST FREQUENCY SWEEP APPROACH WITH A PRIORI CHOICE OF PADÉ APPROXIMANTS AND CONTROL OF THEIR INTERVAL OF CONVERGENCE

ROMAIN RUMPLER* AND PETER GÖRANSSON†

*Creo Dynamics AB
Westmansgatan 37, 58216 Linköping, Sweden
e-mail: rumpler@kth.se - Web page: http://www.creodynamics.com

†The Marcus Wallenberg Laboratory for Sound and Vibration Research (MWL)
Department of Aeronautical and Vehicle Engineering, School of Engineering Sciences
KTH Royal Institute of Technology, SE-100 44 Stockholm, Sweden
e-mail: pege@kth.se - Web page: http://www.kth.se

Key words: Padé approximants; reduced-order model; frequency sweep; finite element analysis.

Abstract. In this work, a solution strategy based on the use of Padé approximants is investigated for efficient solution of parametric finite element problems such as, for example, frequency sweep analyses. An improvement to the Padé-based expansion of the solution vector components is proposed, suggesting the advantageous a priori estimate of the poles of the solution. This allows for the intervals of approximation to be chosen a priori in connection with the Padé approximants to be used. The choice of these approximants is supported by the Montessus de Ballore theorem, proving the convergence of a series of approximants with fixed denominator degrees. An acoustic case study is presented in order to illustrate the potential of the approach proposed by the authors.

1 INTRODUCTION

The use of Padé approximants in order to improve the performance of Finite Element (FE) solution procedures, e.g. for frequency sweeps in structural-acoustic applications, has been given substantial attention in the last decade. These expansion methods have shown to be very efficient for a wide range of problems such as for the control theory [1, 2], acoustics or structural dynamics problems [3–5], or coupled problems [6–10].

In the work initiated with the present paper, the authors build upon previous contributions related to both a modal approach for coupled structural-acoustic problems with dissipative media [11, 12], and its combination with the component-wise solution approximation by Padé approximants [6]. The main idea is to introduce an early step in the Padé

approximation procedure where poles of the response would be calculated. This step has a twofold objective: *i)* with an accurate estimate of the poles of the response, a solid foundation is provided for the Padé approximants to be calculated, thus ensuring a good convergence of these approximants, and *ii)* with the poles known early in the procedure, the choice of approximants in combination with their subdomains of convergence may be done a priori, thus ensuring an efficient reconstruction of the response over the entire domain of interest.

In the first part, the main steps of the component-wise Padé approximation procedure, as previously used by the authors [6], are recalled . In the following section, the addition of a priori knowledge of the poles of the response –an original step to best of the authors' knowledge– is detailed. It is then tested in the last part on an academic acoustic problem allowing for a detailed discussion of the potential of the proposed method. Further steps, currently under investigation by the authors, are presented in the conclusion.

2 UNIVARIATE SOLUTION EXPANSION USING PADÉ APPROXIMANTS

2.1 Numerical calculation of the Padé coefficients

The starting point of the component-wise Padé univariate sweep, as described in [6], is given by a linear system of equations in the following form,

$$\mathbf{K}(x)\mathbf{U}(x) = \mathbf{F}(x), \tag{1}$$

where x is the independent variable of the problem, e.g. the angular frequency ω for the problems of interest in this contribution. In an FE problem, $\mathbf{K}(x)$, $\mathbf{U}(x)$, and $\mathbf{F}(x)$ represent the system matrix of the discretized problem, the solution vector and the vector of externally applied loads, respectively.

A component-wise solution expansion of the solution vector may be sought as Padé approximants in the form

$$u(x_0 + \Delta x) \approx \frac{P_L(\Delta x)}{Q_M(\Delta x)}, \tag{2}$$

where the solution vector $\mathbf{U}(x_0)$, of which $u(x_0)$ is a component, is assumed to be known after solving the system in Eq. (1). $P_L(\Delta x)$ and $Q_M(\Delta x)$ are two truncated power series in the variable $\Delta x = (x - x_0)$, to the orders L and M respectively, and given by

$$P_L(\Delta x) = \sum_{k=0}^{L} p_k (\Delta x)^k, \tag{3a}$$

$$Q_M(\Delta x) = \sum_{k=0}^{M} q_k (\Delta x)^k. \tag{3b}$$

In previous works, it was shown that the coefficients of these power series may be deter-

mined from the coefficients of the Taylor series expansion [13]

$$A_{L+M}(\Delta x) = \sum_{k=0}^{L+M} a_k (\Delta x)^k, \tag{4}$$

where

$$a_k = \frac{u^{(k)}(x_0)}{k!}, \text{ with } u^{(0)}(x_0) = u(x_0) = a_0. \tag{5}$$

These coefficients p_k and q_k are indeed solutions of the system of linear equations resulting from equating the Padé approximant in Eq. (2) to the Taylor series expansion Eq. (4), such that

$$P_L(\Delta x) - A_{L+M}(\Delta x) Q_M(\Delta x) = 0, \tag{6}$$

where the identification of the coefficients of equal order in Δx enables the formation of a set of $(L + M + 1)$ equations, after normalizing the zero-order denominator coefficient, i.e. $q_0 = 1$. In a first step, this leads to the following system of linear equations with the denominator coefficients q_k as unknowns,

$$\begin{bmatrix} a_L & \cdots & a_{L-M+1} \\ \vdots & & \vdots \\ a_{L+M-1} & \cdots & a_L \end{bmatrix} \begin{bmatrix} q_1 \\ \vdots \\ q_M \end{bmatrix} = - \begin{bmatrix} a_{L+1} \\ \vdots \\ a_{L+M} \end{bmatrix} \text{ with } a_i := 0 \text{ if } i < 0. \tag{7}$$

The numerator coefficients may subsequently be determined in a second step by simple algebraic operations,

$$p_k = \sum_{i=0}^{M} q_i a_{(k-i)}, \tag{8}$$

with

$$\begin{cases} k = 0 \cdots L \\ a_j := 0 \text{ if } j < 0 \end{cases}.$$

The calculation of these Padé coefficients however depends on the ability to efficiently calculate the $L + M$ successive partial derivatives of the solution vector, at the reference point $x = x_0$, as implied by Eq. (5). This may be achieved via a recursive scheme, using a Leibniz formula resulting from the differentiation of Eq. (1) with respect to x, at order k, in x_0,

$$\sum_{j=0}^{k} \binom{k}{j} \mathbf{Z}^{(k-j)}(x_0)\, \mathbf{U}^{(j)}(x_0) = \mathbf{F}^{(k)}(x_0), \text{ for } k = 1, \ldots, (L+M), \tag{9}$$

where the zero-order derivatives correspond to the non differentiated functions, and the binomial coefficients are given by

$$\binom{k}{j} = \frac{k!}{j!(k-j)!}. \tag{10}$$

The recursive expression for $\mathbf{U}^{(k)}(x_0)$ follows from extracting the highest-order term from the summation in Eq. (9),

$$\mathbf{Z}(x_0)\,\mathbf{U}^{(k)}(x_0) = \mathbf{F}^{(k)}(x_0) - \sum_{j=0}^{(k-1)} \binom{k}{j} \mathbf{Z}^{(k-j)}(x_0)\,\mathbf{U}^{(j)}(x_0), \text{ for } k = 1, \ldots, (L+M), \quad (11)$$

This implies that the successive derivatives of $\mathbf{U}$ with respect to x, required for the determination of the Padé approximations, can be efficiently calculated as the solution of a full-sized system of equations, with multiple right-hand sides.

2.2 Comments on the procedure

In agreement with the theoretical background recalled in the previous section, a corresponding efficient approach to calculate the Padé-expansion around a solution vector $\mathbf{U}(x_0)$ (only one reference point at x_0 is considered in the present discussion) essentially consists of 5 steps:

1. Factorize the system matrix $\mathbf{Z}$ of Eq. 1 at x_0, and calculate the solution $\mathbf{U}(x_0)$ *(coefficient a_0)*.
2. Solve for the $L+M$ successive derivatives of $\mathbf{U}$ with respect to x using an iterative multiple right-hand-side procedure, Eq. (11) *(coefficients $a_1 \cdots a_{L+M}$)*.
3. For each component of interest in the solution vector, solve for the denominator coefficients via a small linear system of equations Eq. (7) *(coefficients $q_1 \cdots q_M$, $q_0 = 1$)*.
4. For each component of interest in the solution vector, evaluate the numerator coefficients via simple algebraic operations as presented in Eq. (8) *(coefficients $p_0 \cdots p_L$)*.
5. For each component of interest in the solution vector, evaluate the solution expansion around x_0 by evaluating Eq. (2).

Note that while the first two steps have to be performed at a global scale, involving the full size of the original problem, the last three steps may be limited to the degrees of freedom of interest for the solution. A few remarks may be made in relation with the five main steps detailed above:

- Step 2 involves a multiple right-hand-side problem with as many recursions as the number of numerator and denominator coefficients. In practice, there is a limit in the number of recursions that can be performed: the propagation of the small numerical error made at each recursion eventually affects the accuracy of the successive partial derivatives.

- More importantly, the system matrix involved in the solutions of Eq. (7) has the form of a Toeplitz matrix, which, although benefiting from very efficient algorithms, may become rapidly ill-conditioned with its increasing dimension. This implies that the solution of Eq. (7) may become sensitive to small errors of the right-hand-side, which consists of the Taylor coefficients, dependant on the accuracy of the successive partial derivatives.

- In addition to the previous points, the right-hand-side of Eq. (7), as well as the components of the system matrix, consist of the Taylor coefficients of highest order, i.e. those most costly to calculate (latest stages of the recursive procedure in step 2) and those with the most accumulated approximation error.

Even though the procedure presented in this section has proved to be very efficient for a wide variety of examples [4, 6, 7], it appears that some of its steps may be improved in light of the points aforementioned. The following section discusses one such possibility, which additionally opens for the possibility of an a priori choice of Padé approximants and their associated range of convergence.

3 A PRIORI CHOICE OF PADÉ APPROXIMANTS AND CONVERGENCE

3.1 Convergence of Padé approximants

The convergence of Padé approximants is a broad topic which has been given much attention by specialists for decades [13]. In the present work, a key result from the so-called Montessus de Ballore theorem, on the convergence of a series of Padé approximants [14], is highlighted. This theorem guarantees, assuming a meromorphic function within a disk (which is typically the case for the FE solutions of interest in this work), the uniform convergence, except at poles of this function, of a sequence of Padé approximants of increasing numerator degree, fixed denominator degree, and whose poles correspond to simple poles of the approximated function within the considered disk. Thus, this result implies that, given a set of simple poles of the function to be approximated, which define a disk of anticipated convergence, uniform convergence will be observed on this disk, except at the poles location, for a denominator degree corresponding to the number of poles, and an increasing numerator degree.

It is on the basis of this result that the following section suggests an improved approach in order to determine a priori, both the Padé approximants and their anticipated domain of convergence, thus enabling the systematic approximation of FE solutions involving parameter sweeps in a wide range.

3.2 An improved procedure with a priori control of the convergence

Following the observation made in the previous section, the poles of the Padé approximants play a key role in order to ensure a reliable reconstruction of the solution. Additionally, the poles of these Padé approximants should correspond to the zeros of the

characteristic equation of the FE system matrix. Thus, starting with a step at a global scale (full-sized system of equations) by determining the zeros of this characteristic equation within subsets of the range of parametric sweep, would allow to address the comments made in Section 2.2:

- Replace the vector-component-wise step of solving for the denominator coefficients with Eq. (7).
- Ensure to have the same poles from one component to the other of the approximated solution vector, and thus, in light of the convergence properties recalled in the previous section, to enforce the same domain of convergence for all components.
- Reduce the number of Taylor coefficients required to be recursively calculated with Eq. (11), or reallocate them in order to have numerator polynomials of higher degrees, thus ensuring a more accurate approximation over the domain of convergence of the approximants.
- Use the information provided by the location of these poles in order to decompose the range of the parametric sweep into subdomains corresponding to one reference point of solution expansion and its associated convergence interval.

For the sake of validation in the scope of this contribution, the system matrix $\mathbf{K}(x)$ in Eq. (1) is here supposed to be a quadratic function of x with real-valued eigenvalues, solution of a polynomial characteristic equation. The application to more general cases is an important extension for the validation of such an approach which will be presented in further contributions by the authors. Thus, Step 3 in Section 2.2 would be replaced by a step at the global scale, including the calculation of the roots $\{\hat{x}_{01}, \cdots, \hat{x}_{0M}\}$ closest to x_0, solution of the characteristic equation

$$det\left(\mathbf{K}\left(x\right)\right) = 0. \tag{12}$$

The steps of Section 2.2, for the Padé-based approximation of the solution around x_0 may therefore be updated to the following 6-step procedure aiming at the same domain of convergence (i.e. the same number of poles):

1. Solve for a set of roots $\{\hat{x}_1, \cdots, \hat{x}_n\}$, $n > M$, solution of the characteristic equation Eq. (12), including the subset of the M closest roots to x_0 denoted $\{\hat{x}_{01}, \cdots, \hat{x}_{0M}\}$.
2. Factorize the system matrix $\mathbf{Z}$ of Eq. 1 at x_0, and calculate the solution $\mathbf{U}(x_0)$ *(coefficient a_0)*.
3. Solve for the L' successive derivatives of $\mathbf{U}$ with respect to x using an iterative multiple right-hand-side procedure corresponding to Eq. (11) *(coefficients $a_1 \cdots a_{L'}$)*.

4. Once for all the components of interest in the solution vector, and assuming simple roots at this stage, identify the M coefficients $q_1, \cdots, q_M$ such that

$$\frac{(-1)^M}{\hat{x}_{01}\hat{x}_{02}\cdots\hat{x}_{0M}}\left(x-\hat{x}_{01}\right)\left(x-\hat{x}_{02}\right)\cdots\left(x-\hat{x}_{0M}\right) = 1+q_1x+q_2x^2+\cdots+q_Mx^M. \quad (13)$$

   (coefficients $q_1 \cdots q_M$, $q_0 = 1$).

5. For each component of interest in the solution vector, evaluate the numerator coefficients via simple algebraic operations as presented in Eq. (8) *(coefficients $p_0 \cdots p_{L'}$).*

6. For each component of interest in the solution vector, evaluate the solution expansion around x_0 by evaluating Eq. (2).

The first step becomes an important stage when considering a parametric sweep encompassing a large number of poles, as well as higher-multiplicity poles. An initial step where a sufficiently large number of poles may be determined –potentially all the poles in the range of interest– allows for an a priori decomposition in subintervals of Padé-based reconstructions. These intervals are typically defined by a fixed number of poles, or by pushing poles with a higher order multiplicity to their bounds. In contrast with the original procedure in Section 2.2, this latter point prevents the procedure from being potentially hampered by the appearance of a higher-multiplicity pole within an interval of reconstruction, which would concentrate the approximation at a singularity where no convergence can be expected in the sense of the Montessus de Ballore theorem. It may be argued that the pole-based bounds from one interval to the other can be chosen as distinct poles. Then, if we consider three consecutive intervals centered on the interval associated with x_0,

$$\cdots \; [\cdots, \hat{x}_{01-1}] \; [\hat{x}_{01}, \hat{x}_{0M}] \; [\hat{x}_{0M+1}, \cdots] \; \cdots \quad (14)$$

This suggestion is supported by the empirical evidence available for engineering applications [4, 6, 7] showing the slow divergence of Padé approximants outside of the convergence bounds, especially for smooth responses. The analysis of the reconstruction of solutions around multiple intervals is however left for a discussion beyond the scope of this contribution where the attention is turned towards the validation of this procedure on a single interval of reconstruction.

Another point to be highlighted from the upgraded procedure presented above concerns the fact that the numerator polynomial degree L' may be increased ($L' \geq L$) thanks to the reduced number of successive derivatives that need to be calculated in Step 3. In light of the Montessus de Ballore theorem on the convergence of a series of Padé approximants and the limitations associated with the number of successive derivatives of the global solution vector that may be iteratively calculated, this may imply the possibility to increase the range of convergence of each interval in the reconstruction procedure. As a consequence, the number of required intervals required to cover the entire parametric sweep may be subsequently reduced.

Finally, the reference point at x_0 may be placed in the close vicinity (x_0 should not coincide with a pole so as not to jeopardize the calculation of the successive derivatives of the solution vector) of the interval mid-point

$$x_0 = \frac{\hat{x}_{01} + \hat{x}_{0M}}{2}. \tag{15}$$

Note that if the poles are complex-valued, e.g. in the case of a frequency sweep for a damped dynamics problem, the mid-point may be calculated from the real part of the poles.

4 NUMERICAL EXPERIMENT AND VALIDATION

In this section, some points of comparison between the original procedure in Section 2.2 and the updated one are analysed on a simple validation case consisting of a conservative acoustic problem. A rigid acoustic cavity of dimensions $0.1 \times 0.15 \times 0.25$ m^3 is excited at the corner $(0, 0, 0)$ with a time-harmonic acoustic point source for consideration in the range f $= [500, 2250]$ Hz. The acoustic pressure fluctuation at the arbitrary position $(0.06\text{m}, 0.11\text{m}, 0.16\text{m})$ is used in order to compare the frequency sweeps obtained by using a Padé approximant on one interval with the reference solution. There are 11 eigenfrequencies of the acoustic cavity within the range of interest, see Table 1.

Table 1: Eigenfrequencies of the acoustic cavity in the range f $= [500, 2250]$ Hz.

Mode order (l, m, n)	$(0, 0, 1)$	$(0, 1, 0)$	$(0, 1, 1)$	$(0, 0, 2)$	$(1, 0, 0)$	$(0, 1, 2)$
Eigenfrequency (Hz)	686	1143	1333	1372	1715	1786
Mode order (l, m, n)	$(1, 0, 1)$	$(0, 0, 3)$	$(1, 1, 0)$	$(1, 1, 1)$	$(1, 0, 2)$	
Eigenfrequency (Hz)	1847	2058	2060	2172	2196	

First, anticipating a limitation in the Padé approximants orders mostly due to the loss in precision associated with the recursive multiple-right-hand-side solution step, a convergence test is made for a fixed denominator degree of 7 (at most seven poles in the solution can be captured). The two procedures are compared, for one Padé approximant with increasing orders of its numerator, which should imply a uniform convergence on an interval encompassing the 7 closest poles to the reference frequency. The reference frequency is first chosen at 1500 Hz, thus limiting the convergence to an interval including the eigenfrequencies $\{1143 \text{ Hz} \cdots 2058 \text{ Hz}\}$. Fig. 1 compares the convergence for these two procedures with Padé approximants of numerator expansion orders 6, 9 and 11. The main conceptual difference between the two procedures can be clearly seen in Figs. 1a and 1b as the original procedure relies on the scalar-component data in order to determine the pole of the Padé approximant when the updated approach relies on the eigenfrequencies of the full-size problem. Consequently, the latter approach enforces the position of the poles no matter which step of convergence is reached thanks to the polynomial order

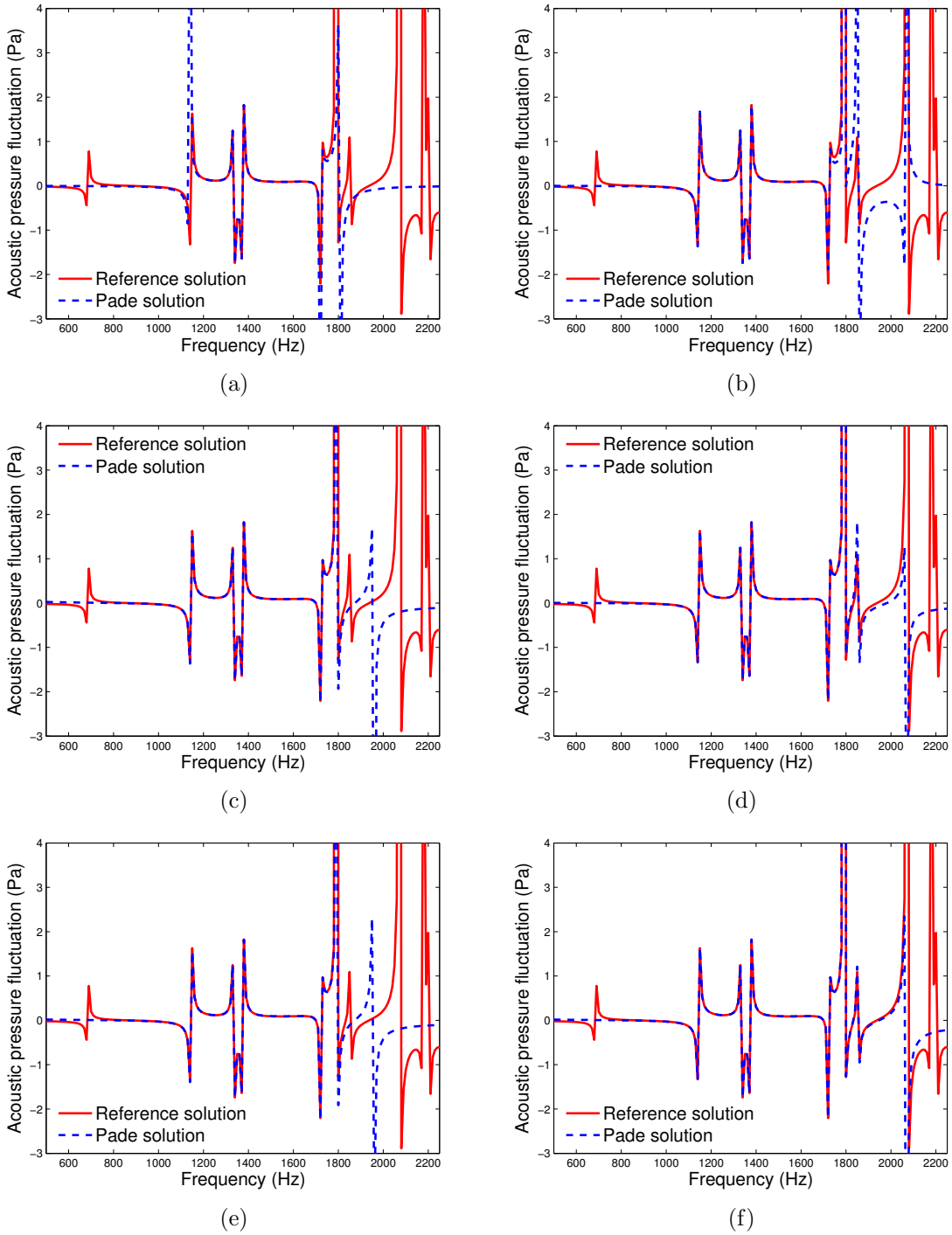


Figure 1: Convergence of the Padé-approximated solution with 7 poles ($M = 7$), reference point at 1500 Hz: original procedure (Left) and updated procedure (Right); (a)–(b) $L = 6$, (c)–(d) $L = 9$, (e)–(f) $L = 11$.

of the numerator. Thus, as soon as expansion order $L = 6$ for the numerator, the 7 expected poles are visible on the approximated solution of the updated procedure in Fig. 1b. In contrast, it appears that two poles are missing for the original procedure (at 1847 Hz and 2058 Hz). A typical behaviour is for the poles to be shifted away in such circumstances, in the higher frequency region in this case. Upon convergence with increasing numerator order, one of the missing poles is shifted down to the domain of interest, as can be seen in Figs 1c and 1e. It is however shifted from the eigenfrequencies of the cavity, and one pole remains uncaptured by the approximation. The updated procedure, however, manages to capture the dynamic content of the response in the entire interval, all 7 eigenfrequencies as well as the sound pressure level in between being accurately approximated after an expansion of the numerator to polynomial order 11.

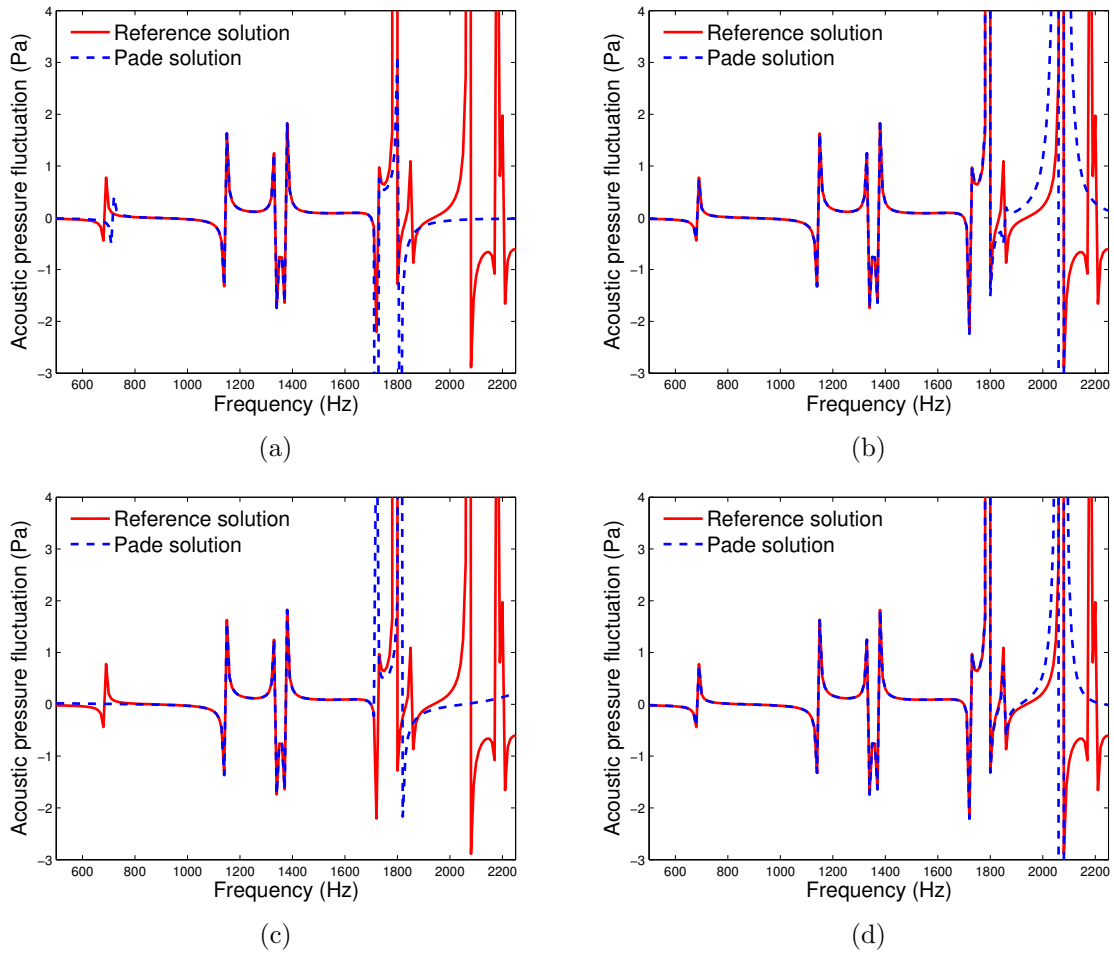


Figure 2: Convergence limits of the Padé-approximated solution with 9 poles ($M = 9$), reference point at 1425 Hz: original procedure (Left) and updated procedure (Right); (a)–(b) $L = 10$, (c)–(d) $L = 12$.

The convergence limit of the original procedure is further illustrated in Figs 2, where 9 poles are considered, with a reference frequency of 1425 Hz, thus limiting the a priori convergence to an interval including the eigenfrequencies $\{686\text{ Hz} \cdots 2060\text{ Hz}\}$. Note that this interval includes two very close eigenfrequencies (2058 Hz and 2060 Hz) at one bound of the domain, a challenge for the approximation. Upon reaching limitations due to a combination of the high order of recursive derivatives in Eq. (11) and the ill-conditioned nature of the problem in Eq. (7), the Padé-approximation based upon the original procedure fails to capture the eigenfrequency at 686 Hz, and collapses in the region $1715 - 1847$ Hz and above (see Figs 2a and 2c). Indeed, estimating the location of the 9 poles captured by the approximant $(L = 12; M = 9)$ in Fig. 2c gives two complex conjugate poles together with $\{1143\text{Hz}, 1210\text{Hz}, 1333\text{Hz}, 1374\text{Hz}, 1717\text{Hz}, 1809\text{Hz}, 2431\text{Hz}\}$. Note that only 4 eigenfrequencies are accurately captured as poles of the approximant. In contrast, the updated procedure allows to have a good representation over the entire domain anticipated, even giving a fair approximation of the response at the challenging upper bound, where two eigenfrequencies almost coincide (see Fig. 2d).

5 CONCLUSIONS

In this contribution, a procedure for the scalar-component Padé approximation approach is proposed as an alternative to the method where the Padé coefficients are derived exclusively from the polynomial coefficients of the power series expansion of the solution. This enhancement suggests that the denominator coefficients, i.e. the poles of the Padé approximant, may be be advantageously determined from the full system of equations. Subsequently, the numerator coefficients may be determined, for each scalar component of the solution vector, from the polynomial coefficients of its power series expansion.

Several points of interest may be highlighted for the potential of this alternative approach: *i)* by calculating the denominator coefficients from the full system of equations, the polynomial coefficients of the power series expansion are used exclusively for the calculation of the numerator coefficients of the approximant, thus allowing to increase the order of expansion of this Padé approximant; *ii)* the calculation of the poles of the solution from the full system of equation may ensure that the accurate poles are enforced for the Padé approximant in the neighbourhood of the expansion point considered, which opens for the possibility to rely fully on the convergence theorem by Montessus de Ballore; *iii)* consequently, the bounds of the intervals of convergence may be determined a priori, which allows for an optimal decomposition of the parametric range of interest into several intervals of convergence.

After the illustration on a conservative acoustic problem in this contribution, current work aims at extending the method to fully coupled dissipative problems and integrating it in a multi-point solution reconstruction.

References

[1] P. Feldmann and R.W. Freund. Efficient linear circuit analysis by Padé approximation via the Lanczos process. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*,

14(5):639–649, 1995.

[2] Christopher Beattie and Serkan Gugercin. Interpolatory projection methods for structure-preserving model reduction. *Systems & Control Letters*, 58(3):225–232, 2009.

[3] Marcus M. Wagner, Peter M. Pinsky, and Manish Malhotra. Application of Padé via Lanczos approximations for efficient multifrequency solution of Helmholtz problems. *The Journal of the Acoustical Society of America*, 113(1):313–319, 2003.

[4] P. Avery, C. Farhat, and G. Reese. Fast frequency sweep computations using a multi-point Padé-based reconstruction method and an efficient iterative solver. *International Journal for Numerical Methods in Engineering*, 69(13):2848–2875, 2007.

[5] U. Hetmaniuk, R. Tezaur, and C. Farhat. Review and assessment of interpolatory model order reduction methods for frequency response structural dynamics and acoustics problems. *International Journal for Numerical Methods in Engineering*, 90(13):1636–1662, 2012.

[6] R. Rumpler, P. Göransson, and J.-F. Deü. A finite element approach combining a reduced-order system, padé approximants, and an adaptive frequency windowing for fast multi-frequency solution of poro-acoustic problems. *International Journal for Numerical Methods in Engineering*, 97(10):759–784, 2014.

[7] J.-D. Chazot, B. Nennig, and E. Perrey-Debain. Harmonic response computation of poroelastic multilayered structures using zpst shell elements. *Computers and Structures*, 121:99–107, 2013.

[8] R. Rumpler and P. Göransson. A finite element solution strategy based on padé approximants for fast multiple frequency sweeps of multivariate problems. *Proceedings of Meetings on Acoustics*, 19, 2013.

[9] R. Rumpler, P. Göransson, and J.-F. Deü. A finite element approach combining a reduced-order system, padé approximants, and an adaptive frequency windowing for fast multi-frequency solution of poro-acoustic problems. *International Journal for Numerical Methods in Engineering*, 97(10):759–784, 2014.

[10] Marcos Souza Lenzi, Sanda Lefteriu, Hadrien Beriot, and Wim Desmet. A fast frequency sweep approach using Padé approximations for solving Helmholtz finite element models. *Journal of Sound and Vibration*, 332(8):1897–1917, 2013.

[11] R. Rumpler, J.-F. Deü, and P. Göransson. A substructuring fe model for structural-acoustic problems with modal-based reduction of poroelastic interface. *Proceedings of the 4th International Conference on Computational Methods for Coupled Problems in Science and Engineering, Coupled Problems 2011*, pages 802–813, 2011.

[12] R. Rumpler, P. Göransson, and J.-F. Deü. A residue-based mode selection and sorting procedure for efficient poroelastic modeling in acoustic finite element applications. *The Journal of the Acoustical Society of America*, 134(6):4730–4741, 2013.

[13] G.A. Baker and P.R. Graves-Morris. *Padé approximants*, volume 59. Cambridge University Press, 1996.

[14] R. de Montessus de Ballore. Sur les fractions continues algébriques. *Bulletin de la Société Mathématique de France*, 30:28 – 36, 1902.

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis (Eds)

PROPAGATION OF ACOUSTIC-GRAVITY WAVES IN INHOMOGENEOUS OCEAN ENVIRONMENT BASED ON MODAL EXPANSIONS AND HP-FEM

KOSTAS A. BELIBASSAKIS*, GERASSIMOS A. ATHANASSOULIS*, ANGELIKI E. KARPERAKI* AND THEODOSIOS K. PAPATHANASIOU†

* School of Naval Architecture and Marine Engineering
National Technical University of Athens
Heroon Polytechniou 9, Zografos 15773, Athens, Greece
e-mail: kbel@fluid.mech.ntua.gr, http://arion.naval.ntua.gr/~kbel/

† School of Applied Mathematical and Physical Science
National Technical University of Athens
Heroon Polytechniou 9, Zografos 15773, Athens, Greece
e-mail: papaththh@gmail.com

Key words: Acoustic-Gravity waves, Ocean Environment, Coupled-mode methods, FEM

Abstract. A coupled mode model is presented for the propagation of acoustic-gravity waves in layered ocean waveguides. The analysis extends previous work for acoustic waves in inhomogeneous environment. The coupled mode system is derived by means of a variational principle in conjunction with local mode series expansion, obtained by utilizing eigenfunction systems defined in the vertical section. These are obtained through the solution of vertical eigenvalue problems formulated along the waveguide. A crucial factor is the inclusion of additional modes accounting for the effects of spatialy varying boundaries and interfaces. This enhancement provides an implicit summation for the slowly convergent part of the local-mode series, rendering the series rapidly convergent, increasing substantialy the efficiency of the method. Particular aspects of the method include high order Lagrange Finite Element Methods for the solution of local vertical eigenvalue problems in the case of multilayered waveguides, and Gauss-type quadrature for the computation of the coupled-mode system coefficients. The above aspects make the present method quite efficient for long range propagation in extended waveguides, such as the ones found in geophysical applications, e.g. ocean basins, as only few modes are needed for the accurate representation of the wave field.

1 INTRODUCTION

Ocean waves generate acoustic modes in a wide range of acoustic frequencies, as a result of non-linear interactions of pairs of nearly opposing gravity waves having equal or nearly equal frequencies. In this case, the lower frequency part of the spectrum, nominally for frequencies lower than 2 Hz is caused by the nonlinearity of the hydrodynamic equations; see e.g., [1,2]. Also, in the low frequncy band, energetic acoustic–gravity waves appear as a result of seismic activity in the seabed and generation and propagation of tsunami waves; see, e.g., [3] and the references cited there. These waves propagating from the open sea to nearshore areas, interact

with various layers of the inhomogeneous ocean waveguide and the seabed, and are characterised by partial energy exchage of the modes, shoaling and scattering effects.

Wave propagation and scattering in an inhomogeneous waveguide is an interesting mathematical problem finding important applications, as, e.g., underwater acoustic propagation and scattering in shallow water and seismoacoustics [4, 5], atmospheric acoustics [6] and other. Several methods for treating this, generally non-separable, boundary value problem have been proposed, ranging from fully numerical, finite element and finite difference methods, to semi-analytical ones, like wavenumber integration, boundary integral equation and coupled-mode techniques, as well as various asymptotic models, like ray theory and the adiabatic and parabolic approximations; see e.g., [5, 7]. Fully numerical methods are computationally intense and thus, their use is more appropriate for short-range/low-frequency propagation and local scattering problems. In this work, a fast-convergent spectral model is presented for treating harmonic wave propagation and scattering problems in stratified, non uniform waveguides, governed by the Helmholtz equation. The method is based on a local mode series expansion, obtained by utilizing local eigenfunction systems defined through the solution of eigenvalue problems, formulated along the cross section of the waveguide. Following Belibassakis et al. [8] the local mode series are enhanced by including additional modes accounting for the effects of inhomogeneous waveguide boundaries and/or interfaces. The additional modes provide an implicit summation of the slowly convergent part of the local-mode series, rendering the remaining part to be fast convergent, increasing the efficiency of the method, especially in long-range propagation applications (see also [9] and [10]). Using the enhanced local mode expansion, in conjunction with an energy-type variational principle, a coupled-mode system of equations is derived for the determination of the unknown modal-amplitude functions. The numerical solution of the local vertical eigenvalue problems, in the case of multilayered waveguides, is obtained by means of *h*- and *p*-FEM (see, e.g., [8, 11]) exhibiting robustness and good rates of convergence. On the basis of the above, the coefficients of the coupled-mode system are calculated by numerical integration. Subsequently, the solution of the present coupled-mode system is obtained by using a finite difference scheme based on a uniform grid and using second-order central differences to approximate derivatives. Particular aspects of the method include high order Lagrange Finite Element Methods for the solution of local vertical eigenvalue problems in the case of multilayered waveguides, and Gauss-type quadrature for the computation of the coupled-mode system coefficients. The above aspects makes the present method quite efficient for long range propagation in extended waveguides, such as the ones found in geophysical applications, e.g. ocean basins, as only few modes are needed for the accurate representation of the unknown fields. Numerical examples are presented illustrating the efficiency of the present model, that can be naturally extended to treat propagation and scattering problems in more complicated 3D waveguides.

2 PROBLEM DESCRIPTION

Consider the multilayered waveguide of Fig.1. For simplicity we restrict ourselves to a 2D problem in an ocean acoustic environment, governed by Helmholtz equation. However, the present method and analysis can be naturally extended to more general 3D acoustic-gravity waveguides.

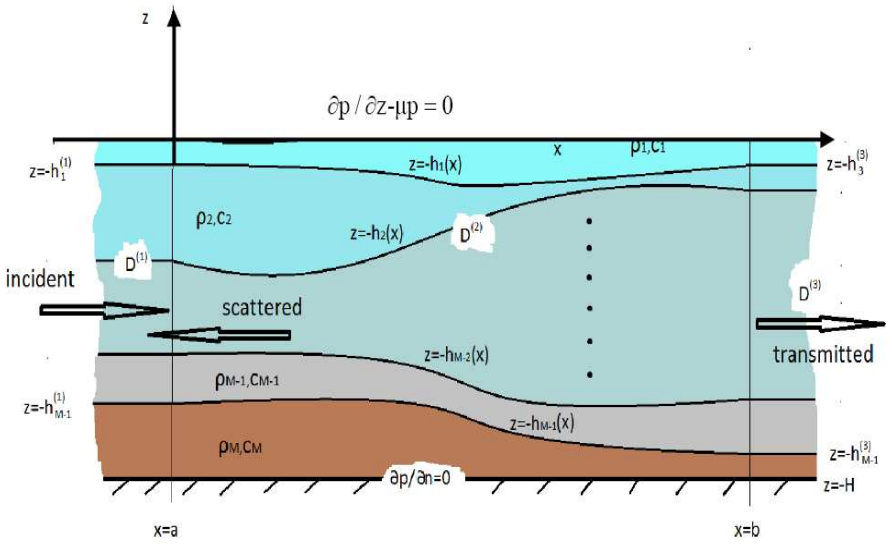


Figure 1: Multilayered two-dimensional acoustic waveguide.

The domain $D = D^{(1)} \bigcup D^{(2)} \bigcup D^{(3)}$ is decomposed into three parts $D^{(m)}$, $m = 1,2,3$ (see Fig. 1), as follows: $D^{(1)}$ is the subdomain characterized by $x_1 < a$ nand $D^{(3)}$ is the subdomain characterized by $x_1 > b$ $(b > a)$, and $D^{(2)}$ is the variable cross section subdomain lying between $D^{(1)}$ and $D^{(3)}$. A similar decomposition is also applied to the (upper and lower) boundaries, as well as to the internal interfaces.-The acoustic medium inside the domain is stratified. The physical properties of the layers, vary with respect to the (x,z) coordinates in the middle range-dependent subdomain $D^{(2)}$, and present only vertical variability in the two semi-infinite subdomains $D^{(1)}$ and $D^{(3)}$. Assuming that the whole domain consists of M layers, a total number of $M-1$ interfaces at $z = -h_j(x)$, $j = 1,2,...,M-1$, are considered, where $h_j(x)$ denotes the local depth of each interface (see Fig.1). The waveguide is terminated below by a perfectly rigid (acoustically hard) horizontal boundary, located at $z = -H$. On the other hand, the waveguide is terminated above by an acoustically soft boundary, located at $z = 0$, corresponding to the free surface.

The density ρ_j, $j = 1,2,..M$, of each layer is assumed to be constant within the layer, presenting possibly sharp discontinuities at the interfaces. Moreover, the sound speed $c_j(x,z)$, $j = 1,2,..M$, presents both vertical and horizontal variability in the middle subdomain $D^{(2)}$, and could also exhibit strong discontinuity at the interfaces. The sound speed becomes function only of z in the two semi-infinite subdomains $D^{(1)}$ and $D^{(3)}$, which are then range independent subdomains with respect to both geometry and physical parameters. This fact permits us to obtain complete expansions of the wave field in the above semi-infinite regions by means of separation of variables, and consistently formulate the conditions of wave incidence and transmission at $x=a$ and $x=b$, respectively.

2.1 Governing Equations

Restricting ourselves to monochromatic waves of angular frequency $\omega=2\pi f$, the acoustic harmonic wave propagation problem inside the present multi-layered waveguide is governed by the Helmholtz equation. The respective boundary value problem takes the form of finding the continuous function p representing the acoustic pressure such that

$$\nabla\cdot\left(\frac{1}{\rho}\nabla p\right)+\frac{k^2}{\rho}p=0\ ,\ \text{in}\quad D \tag{1}$$

where the wavenumber $k(x,z)=\omega/c(x,z)$ is a piecewise smooth function of the spatial coordinates, possibly presenting sharp discontinuities at the interfaces $z=-h_j(x),\quad j=1,2,...,M-1$. Eq. (1) is supplemented by the following boundary conditions

$$\partial p/\partial z-\mu p=0\,,\quad \text{on the free surface}\quad z=0\,, \tag{2}$$

$$\partial p/\partial n=\partial p/\partial z=0\,,\ \text{on the perfectly rigid (terminating) boundary at}\ \ z=-H\,, \tag{3}$$

where $\mu=\omega^2/g$ is the frequency parameter of gravity waves (g=9.81m/s^2), in conjunction with the interface conditions

$$\frac{1}{\rho_j}\frac{\partial p}{\partial n}=\frac{1}{\rho_{j+1}}\frac{\partial p}{\partial n}\quad\text{on}\quad z=-h_j(x),\quad j=1,2,...,M-1\,. \tag{4}$$

In the previous equations $\partial p/\partial n=\mathbf{n}\nabla p$ denotes the normal derivative, where $\mathbf{n}$ is the unit normal vector on each boundary and interface. We consider a transmission problem forced by plane waves propagating in the positive x direction. The waves are incident from $D^{(1)}$, and then they are refracted and scattered in the range dependent subdomain $D^{(2)}$, and finally transmitted in $D^{(3)}$. In order to treat the present problem in the infinite domain, complete normal-mode type representations of the wave field in the regions of incidence $D^{(1)}$ and transmission $D^{(3)}$ are derived by separation of variables. In particular, the expansion of the wavefield in $D^{(1)}$ consists of incident and reflected (scattered) waves is as follows,

$$p^{(1)}=\sum_{n=1}^{\infty}\left(A_n^{(1)}e^{ik_n^{(1)}x}+B_n^{(1)}e^{-ik_n^{(1)}x}\right)Z_n^{(1)}(z) \tag{5}$$

where the functions $Z_n^{(1)}(z)$ and the numbers $k_n^{(1)},\ n=1,2,3...,$ satisfy the following vertical eigenvalue problem in $D^{(1)}$

$$\frac{d^2Z_n^{(1)}}{dz^2}+\left[\left(k^{(1)}(z)\right)^2-(k_n^{(1)})^2\right]Z_n^{(1)}=0\,, \tag{6}$$

$$\frac{dZ_n^{(1)}(z=0)}{dz}-\mu Z_n^{(1)}(z=0)=0,\qquad \frac{dZ_n^{(1)}(z=-H)}{dz}=0\,, \tag{7a,b}$$

in conjunction with the interface condition,

$$Z_n^{(1)}(-h_j^{(1)}+0) = Z_n^{(1)}(-h_j^{(1)}-0), \quad j=1,2,\, M-1\,, \tag{8}$$

$$\frac{1}{\rho_j}\frac{\partial Z_n^{(1)}(-h_j^{(1)}+0)}{\partial z} = \frac{1}{\rho_{j+1}}\frac{Z_n^{(1)}(-h_j^{(1)}-0)}{\partial z}, \quad j=1,2,\, M-1\,, \tag{9}$$

where $k^{(1)}(z)=\omega/c^{(1)}(z)$. Similarly, the expansion of the acoustic wavefield in the region of transmission $D^{(3)}$, consists only of outgoing radiated waves, and is given by

$$p^{(3)} = \sum_{n=1}^{\infty}\left(A_n^{(3)}e^{ik_n^{(1)}x}\right)Z_n^{(3)}(z)\,, \tag{10}$$

where the eigenfunctions $Z_n^{(3)}(z)$ and the corresponding eigenvalues $k_n^{(3)}$, $n=1,2,3...$, are obtained by similar as above vertical eigenvalue problems formulated in $D^{(3)}$. From the properties of regular Sturm-Liouville problems ([12], [13]) the eigenvalues $\left\{\left(k_n^{(m)}\right)^2, n=1,2..\right\}$, m=1,3, are discrete, infinite, with continuously decreasing moduli, and thus, the corresponding parameters $\left\{k_n^{(m)},\, n=1,2,3...\right\}$, are subdivided into a finite real subset $\left\{k_n^{(m)},\, n=1,2,3...N_p^{(m)}\right\}$ and an infinite imaginary one $\left\{i\left|k_n^{(m)}\right|,\, n=N_p^{(m)}+1,....\right\}$, where $N_p^{(m)}$, denotes the number of propagating modes in $D^{(m)}$, $m=1,3$. The first eigenvalue (n=1) and corresponding eigenvector is essentially associated with the free-surface gravity mode which presents fast decay in depth.

Clearly, in order for the wave field to remain bounded at infinity, the coefficients of the expansion $A_n^{(1)}=0,\, n>N_p^{(1)}$. On the other hand the terms $A_n^{(1)}\exp\left(ik_n^{(1)}x\right)Z_n^{(1)}(z)$, $n\le N_p^{(1)}$, constitute the given data associated the incident wave field. Due to the linearity of the problem each one of the above terms could be separately considered as forcing of the present acoustic waveguide and the solutions is obtained by superposition of the responses. In addition, the terms $\exp\left(-ik_n^{(1)}x\right)Z_n^{(1)}(z)$, $n>N_p^{(1)}$, and $\exp\left(ik_n^{(3)}x\right)Z_n^{(3)}(z)$, $n>N_p^{(3)}$, are the evanescent modes in $D^{(m)}$, $m=1,3$, respectively. These modes decay exponentially at large distances from the inhomogeneity in the two semi-infinite strips. By exploiting the representations (5) and (10), the problem can be formulated as a transmission boundary value problem in the bounded subdomain $D^{(2)}$, satisfying Eq. (1), (2) (3) and (4) and the following matching conditions:

$$p^{(2)}(x,z)=p^{(1)}(x,z), \qquad \frac{\partial p^{(2)}}{\partial x^{(2)}}=\frac{\partial p^{(1)}}{\partial x^{(1)}}\,, \quad x=a\,, \quad -H<z<0, \tag{11a}$$

$$p^{(2)}(x,z)=p^{(3)}(x,z), \qquad \frac{\partial p^{(2)}}{\partial x}=\frac{\partial p^{(3)}}{\partial x}\,, \quad x_1=b, \quad -H<z<0\,. \tag{11b}$$

2.2 Variational Formulation

We proceed to formulate a functional $\mathscr{F}$ allowing us to state a variational formulation of the transmission problem. The admissible function space for the wave field $p^{(2)}(x,z) \in D^{(2)}$ (simply denoted from now on as p), consists of globally continuous and piecewise smooth functions, possessing continuous second derivatives in the interior of each layer, such that

$$\frac{\partial p(x,z=0)}{\partial z} - \mu p(x,z=0) = 0 \tag{12}$$

For this purpose, we consider the following energy-type functional

$$\begin{aligned} \mathscr{F}\left(p\ ,\left\{B_n^{(1)}\right\},\left\{A_n^{(3)}\right\}\right) = \ & \frac{1}{2}\int_{D^{(2)}}\left[\rho^{-1}\left(\nabla p\ \right)^2 - \rho^{-1}k^2\left(p\ \right)^2\right]dxdz + \frac{1}{2\rho_1}\int_{x=a}^{x=b}\mu p^2 dx + \\ -\int_{z=-H}^{z=\eta_3}\left(p - \frac{1}{2}p^{(3)}\left(\left\{A_n^{(3)}\right\}\right)\right)\frac{\partial p^{(3)}\left(\left\{A_n^{(3)}\right\}\right)}{\partial x}dz \ & + \int_{z=-H}^{z=\eta_1}\left(p - \frac{1}{2}p^{(1)}\left(\left\{B_n^{(1)}\right\}\right)\right)\frac{\partial p^{(1)}\left(\left\{B_n^{(1)}\right\}\right)}{\partial x}dz\ . \end{aligned} \tag{13}$$

The present problem admits of an equivalent variational formulation expressed by

$$\delta\mathscr{F}\left(p\ ;\left\{B_n^{(1)}\right\},\left\{A_n^{(3)}\right\}\right) = 0 \tag{14}$$

Using Green's theorem in conjunction with Eq. (12), the above variational equation takes the form

$$\begin{aligned} & -\int_{D^{(2)}}\left(\nabla\cdot\left(\rho^{-1}\nabla p\right) + \rho^{-1}k^2 p\right)\delta p\ \ dxdz + \int_{z=-H}^{z=\eta_1}\left(p\ - p^{(1)}\right)\delta\frac{\partial p^{(1)}}{\partial x}dz + \\ & -\int_{z=-H}^{z=\eta_3}\left(p\ - p^{(3)}\right)\delta\frac{\partial p^{(3)}}{\partial x}dz - \int_{z=-H}^{z=\eta_1}\left(\frac{\partial p}{\partial x} - \frac{\partial p^{(1)}}{\partial x}\right)\delta p\ dz + \\ & -\int_{z=-H}^{z=\eta_3}\left(\frac{\partial p}{\partial x} - \frac{\partial p^{(3)}}{\partial x}\right)\delta p\ \ dz - \sum_{j=1}^{M-1}\int_{z=-h_j(x)}\left(\frac{1}{\rho_j}\frac{\partial p}{\partial N} - \frac{1}{\rho_{j+1}}\frac{\partial p}{\partial N}\right)\delta p\,dx = 0\,, \end{aligned} \tag{15}$$

where $\left.\frac{\partial p}{\partial N}\right|_{z=-h_j(x)} = \frac{\partial p}{\partial z} + \frac{dh_j}{dx}\frac{\partial p}{\partial x}$. The usefulness of the above variational principle hinges on the fact that it leaves us the freedom to choose any particular representation for the unknown field $p(x,z) \in D^{(2)}$. In this way, a variety of possible algorithms for the numerical solution of the present wave problem can be constructed.

3 ENHANCED LOCAL MODE REPRESENTATION

Inside the bounded domain $D^{(2)}$, the solution $p(x,z)$ may be set into a standard spectral-type representation, based on local-mode series, as follows

$$p\ (x,z)=\sum_{n=1}^{\infty}U_n(x)Z_n(z;x)\,. \tag{16}$$

The family of local vertical basis functions $Z_n(z;x)$ appearing in the above expansion, which are parametrically dependent on x, is obtained by formulating and solving local, vertical Sturm-Liouville problems in the z-intervals $[-H,\,0]$, for each horizontal position $a<x<b$. However, any finite truncation of the series (16) is incompatible with any of the sloping interface conditions, whenever $dh_j(x)/dx\neq 0$, $j=1,2,...M-1$, rendering the above series to converge only in an L^2-sense, and the coefficients U_n to decay like $O\left(n^{-2}\right)$; see Belibassakis et al [8]. To remedy this inconsistency, an additional mode associated with each interface is introduced, denoted by $U_j(x)Z_j(z;x)$, $j=-M+2,..,-1,0$. These modes are called sloping-interface modes. Thus, we obtain the following enhanced local-mode series

$$p(x,z)=\sum_{n=-M+2}^{0}U_n(x)Z_n(z;x)+\sum_{n=1}^{\infty}U_n(x)Z_n(z;x)\,. \tag{17}$$

The vertical structure of the sloping-interface modes, for every horizontal position $a<x<b$, is any globally continuous and piecewise smooth function defined with support in the local vertical intervals $[-h_{M-1}(x),-h_{M-2}(x)],....,[-h_1(x),0]$, satisfying the following condition(s)

$$\left.\frac{1}{\rho_j}\frac{\partial Z_n}{\partial z}\right|_{z=-h_j}-\left.\frac{1}{\rho_{j+1}}\frac{\partial Z_n}{\partial z}\right|_{z=-h_j}=1,\quad j=1,2,...,M-1\,. \tag{18}$$

Moreover, the function $Z_0(z;x)$ should satisfy the homogeneous Dirichlet condition at $z=\eta(x)$. Consequently, the M-1 terms $U_j(x)Z_j(z;x)$, $j=-M+2,..,-1,0$, are additional degrees of freedom in the bounded subdomain $D^{(2)}$, permitting the consistent satisfaction of all interface conditions. The amplitude of the additional modes is given by

$$U_{j-1}(x)=\left.\frac{1}{\rho_j}\frac{\partial p}{\partial z}\right|_{z=-h_j}-\left.\frac{1}{\rho_{j+1}}\frac{\partial p}{\partial z}\right|_{z=-h_j},\quad j=1,2,..M-1\,. \tag{19}$$

From this last relation, it is evident that no extra mode needs to be introduced in the last layer terminated in the lower flat boundary, where a homogeneous Neumann condition is satisfied.

The important effect of the additional modes is to significantly increase the rate of decay of Z_n – Fourier coefficients of the acoustic wave potential (modal amplitudes). In this case, the modes associated with the enhanced series exhibit a rapid decay rate: $\left|U_n(x)\right|\le C(x)\,n^{-4}$, $n\to\infty$, $\forall x\in[a,b]$. The bound $C(x)$ is a continuous function on $[a,b]$ and, thus, the previous estimate is global: ; see also [9,10]. If the additional modes are not included, then the

rate of decay of the modes in the standard series (16) is only $|U_n| = O\left(n^{-2}\right)$. The above result is obtained by means of repetitive use of integration by parts, in conjunction with the properties of the local Sturm-Liouville system, and will be illustrated through appropriate numerical examples in a subsequent section.

4 FEM SOLUTION OF THE LOCAL VEP

In this section we will describe the finite element method, as applied to the solution of the vertical eigenvalue problem (vep). In the following we adopt the notation by Hughes [11]. We assume that for all x in the interval $a<x<b$, it is $\rho(z;x),\ k^2(z;x) \in L^\infty(-H,0)$.

Let us introduce the (Sobolev) function spaces $H^1_{|E(x)} \triangleq \left\{u : u \in H^1(-H,0),\ \mathrm{x} \in [\mathrm{a,b}]\right\}$. The continuous vertical eigenvalue problem, at each horizontal position, can now be stated in variational form as follows

$$\text{Find } (\lambda, p) \in \mathbb{R} \times H^1_{0|E(x)} \text{ such that } \ a(w,p) = \lambda b(w,p)\ ,\ \forall w \in H^1_{|E(x)}\,, \tag{20}$$

Where

$$a(w,p) = \int_{-H}^{0} \rho^{-1} \frac{dw}{dz}\frac{dp}{dz} dz - \int_{-H}^{0} k^2 \rho^{-1} wp dz + \rho^{-1}\mu\left[wp\right]_{z=0}\,, \tag{21a}$$

$$b(w,u) = -\int_{-H}^{0} \rho^{-1} wp dz\,, \tag{21b}$$

Assume a partition of $[-H,0]$, of the form $-H = z_1 < z_2 < ..... < z_{N+1} = 0$, with $N \in \mathbb{N}$ and $N > M$ (M being the number of layers in the waveguide). The partition is such that $\forall x$ the interface positions coincide with $M-1$ nodes. We introduce the sequence of finite element sub-spaces,

$$V^h \triangleq \left\{u^h \in H^1(-H,0) : u^h\Big|_{[z_i, z_{i+1}]} \equiv \mathrm{P}_\ell(\mathrm{z}),\ \ i = 1,2,...,N,\ \ell \in \mathbb{N}, \mathrm{x} \in [\mathrm{a,b}]\right\}$$

where $P_\ell(z)$ is a polynomial of degree ℓ. Obviously $V^h \subset H^1_{|E(x)}$. The discrete variational formulation of the vertival eigenvalue problem takes the following form:

$$\text{Find } (\lambda^h, p^h) \in \mathbb{R} \times V^h \text{ such that } a(w^h, p^h) = \lambda^h b(w^h, p^h)\ ,\ \forall w^h \in V^h \tag{22}$$

In the following analysis we assume piecewise linear, quadratic and cubic interpolations for the finite element solution, i.e. $\ell = 1, 2 \text{ and } 3$. Regardless of the interpolation degree, the desired solution has the form,

$$p^h = \sum_{j=1}^{N} c_j N_j(z)\,. \tag{23}$$

where $N_j \in V^h$. Introducing the above expansion in Eq. (22), the discrete variational formulation finally becomes an eigenvalue matrix equation of the form:

$$\mathbf{A}\,\mathbf{u} = \lambda\,\mathbf{B}\,\mathbf{u}\ , \tag{24}$$

where the elements of the $N \times N$ matrices $\mathbf{A}$ and $\mathbf{B}$ are $a_{ij} = a(N_i, N_j)$ and $b_{ij} = b(N_i, N_j)$, $i, j = 1,2,..,N$, respectively.

4.1 Numerical solution of VEP

As demonstrative examples we consider the case of an acoustic environment of total thickness (depth) $H = 1000m$, consisting of two layers of thickness 100m (upper) and 900m (lower) and equal thickness 500m, respectively. In the first case the position of the internal interface is at a depth $h_1 = 100m$, and in the second at $h_1 = 500m$. Moreover, two frequencies have been considered $f = 2Hz$ and $f = 0.08Hz$. Both the density and speed of sound are assumed to be constant within each layer. For the upper layer corresponding to water these quantities are $\rho_1 = 1g/cm^3$, $c_1 = 1500m/\sec$, and for the lower layer $\rho_2 = 1.5g/cm^3$, $c_2 = 1700\ m/\mathrm{s}$, corresponding to sediment. Details concerning the exact analytical solution of this problem are included in the APPENDIX 1.

The first 5 eigenfunctions for *f*=2Hz are shown in Fig. 2. A zoom on the upper 10m is shown in the right subplot, illustrating the layered structure of the free-surface (gravity) mode (*n*=1), which presents exponetial decay in depth. Corresponding resuts for the lower frequency *f*=0.08Hz are plotted in Fig.3, for two positions of the interface at 100m (left subplot) and 500m (right subplot), which is indicated by using thick dashed lines. In the above plots the first eigenfunction (*n*=1) associated with the free-surface (gravity) mode is normalizad with its maximum value at z=0, while the rest of the modes (*n*=2,3,4…) are normalizad with respect to their values at the bottom (*z*=-*H*). With increasing frequency parameter μ, a positive sway of the first eigenvalue, is observed leading to the formation of a boundary layer visible in the sublot of Fig. 2. The eigenvalues as computed with the finite element method, using $N = 20, 40, 80$ elements, are in perfect match when compared against the exact solution for the first modes.A series of numerical results are shown for the case of placing the interface at $h_1 = 100m$ and $f = 0.08Hz$. In Fig. 4 a comparison between the first exact and computed eigenvalues is shown for different number of elements, $N = 20, 40 \text{ and } 80$, and two different interpolation degrees, $p = 1 \text{ and } 2$. It is observed that the error of the numerical solution for the eigenvalues is found to increase with increasing eigenvalue numbers. The convergence of the finite element solution for the 5th, 10th and 15th eigenvalues is demonstrated in Fig. 5. Convergence rates are observed for increasing interpolation degree, $p = 1, 2 \text{ and } 3$ and are found to be 2, 4 and 6 respectively. Enhanced rates have been obtained by raising the degree ℓ of the piecewise polynomials.

5 THE COUPLED - MODE SYSTEM

Having obtained the eigenfunctions associated with the local vertical problem, $\forall x \in [a,b]$, we proceed to the calculation of the mode amplitudes $\{U_j(x),\ \ j = -M+2,..,-1,0,1,2,.....\}$. To this respect we substitute the enhanced local mode representation (17) in the variational

principle (15), and express the variation of the unknown field $p(x,z) \in D^{(2)}$, through the variations of the modal amplitudes

$$\delta p(x,z) = \sum_{n=-M+1}^{\infty} Z_n(z;x)\, \delta U_n(x) \quad . \tag{25}$$

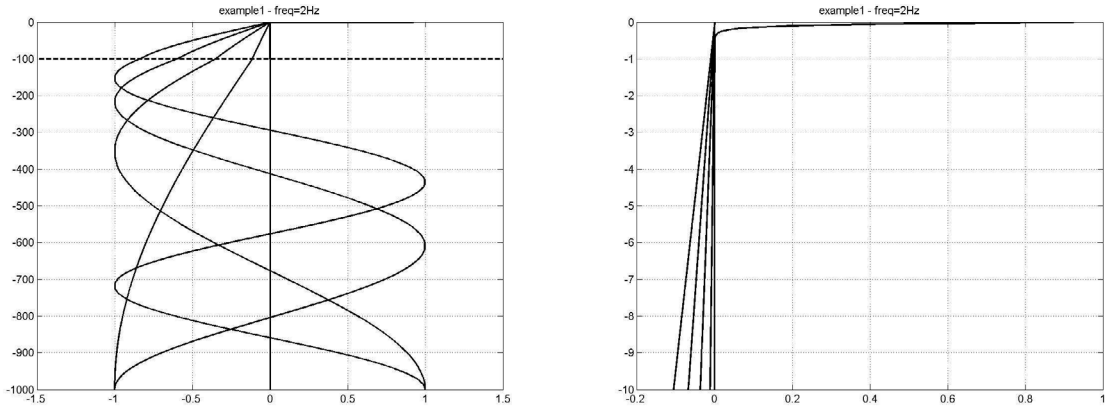


Figure 2: First 5 eigenfunctions for *f*=0.2Hz. The position of the interface is at 100m is indicated in the left subplot by thick dashed lines. A zoom on the upper 10m is shown in the right subplot, illustrating the layered structure of the free-surface (gravity) mode (*n*=1).

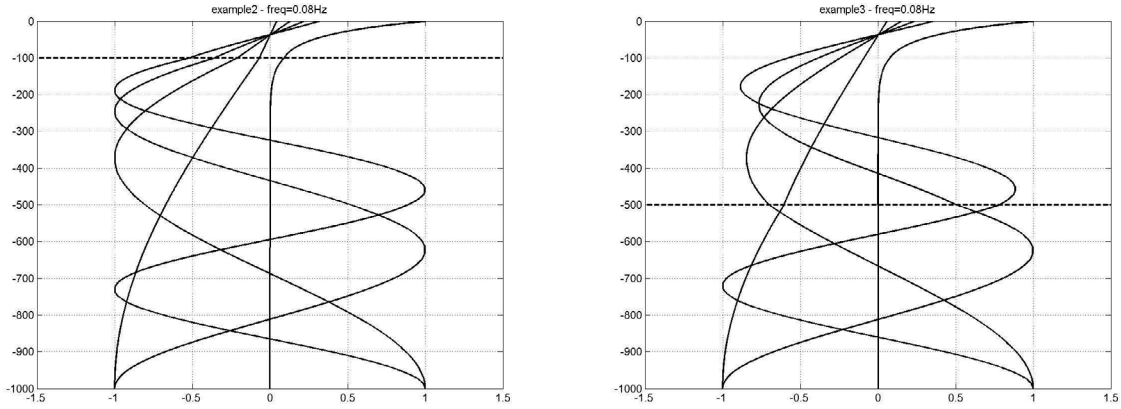


Figure 3: First 5 eigenfunctions in the low frequency case *f*=0.08Hz. The position of the interface at 100m (left subplot) and 500m (right subplot) is indicated by thick dashed lines.

Next, by considering only the variations $\delta U_n(x),\ n = -M+2,..,0,1,...$, in $a < x < b$, we obtain from the first term in the left hand side of Eq. (16) the following coupled-mode system (CMS) of second-order ordinary differential equations, with respect to the mode amplitudes, $U_n(x), n = -M+2,..,0,1,2,..$,

$$\sum_{n=-M+2}^{\infty} a_{mn}(x)\frac{d^2U_n(x)}{dx^2} + b_{mn}(x)\frac{dU_n(x)}{dx} + c_{mn}(x)U_n(x) = 0\,, \tag{26}$$

where $m = -M+2,..,0,1,2,...$ The x-dependent coefficients a_{mn}, b_{mn} and c_{mn} are defined in terms of $Z_n(z;x)$ in $a < x < b$ and are given by

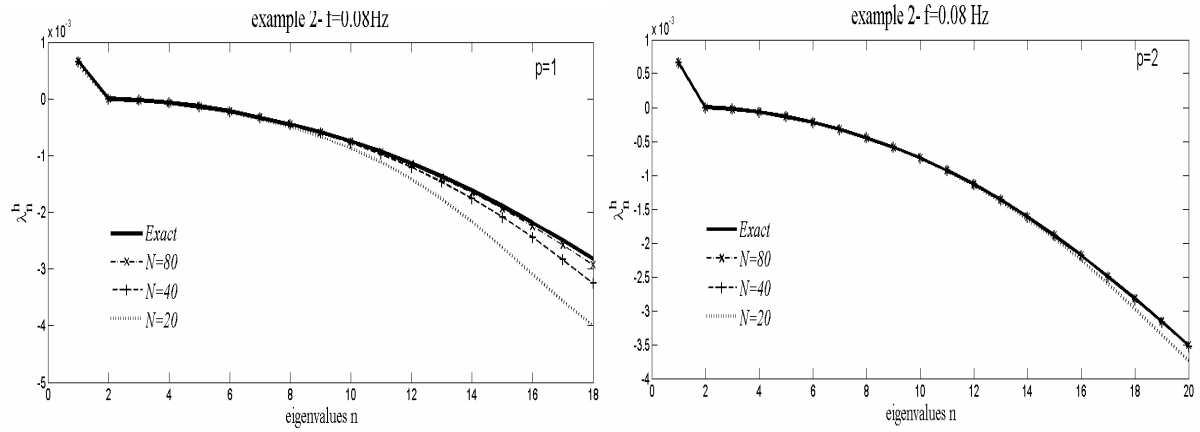


Figure 4: Comparison between the computed eigenvalues against the exact solution (thick line) for $p=1$ (left) and $p=2$ (right).

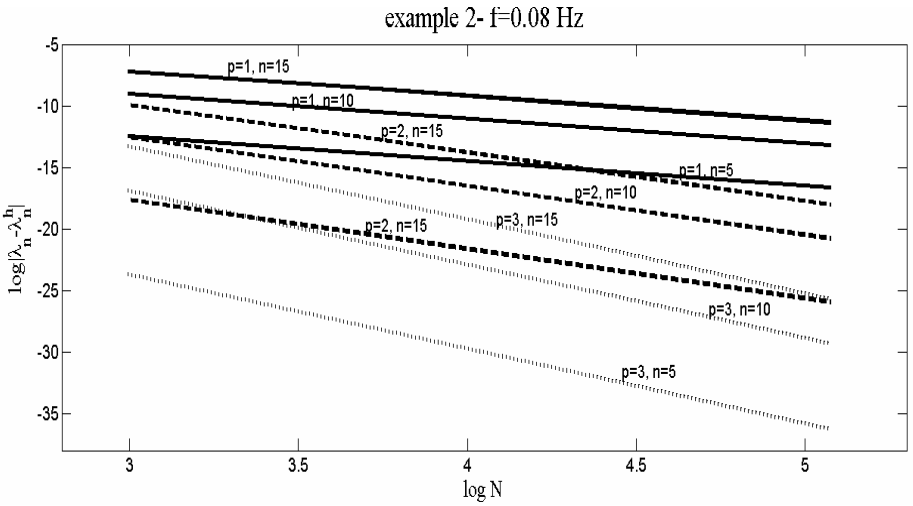


Figure 5: Convergence of the finite element solution for the 5th, 10th and 15th eigenvalues, using $p=1,2$ and 3.

$$a_{mn} = \langle Z_n, Z_m \rangle, \tag{27}$$

$$b_{mn} = 2\left\langle \frac{\partial Z_n}{\partial x}, Z_m \right\rangle + \sum_{j=1}^{M-1} \left(\frac{1}{\rho_j} - \frac{1}{\rho_{j+1}} \right) \frac{dh_j}{dx} Z_n(-h_j) Z_m(-h_j), \tag{28}$$

$$c_{mn} = \left\langle \frac{\partial^2 Z_n}{\partial x^2} + \frac{\partial^2 Z_n}{\partial z^2} + k^2 Z_n, Z_m \right\rangle + \sum_{j=1}^{M-1} \left([\![Z_{n,z}]\!] + \frac{dh_j}{dx} [\![Z_{n,x}]\!] \right) Z_m(-h_j). \tag{29}$$

In the above relations, $\langle f,g\rangle := \int_{-H}^{\eta} \rho^{-1} f(z)g(z)dz$ is the weighted inner product of $L^2(-H,0)$ function spaces, for all x in $a < x < b$. Further, the quantities $[\![Z_n]\!]_x$ are defined by

$$[\![Z_{n,w}]\!] = \left[\frac{1}{\rho_j} \frac{\partial Z_n}{\partial w}\bigg|_{z=-h_j^+} - \frac{1}{\rho_{j+1}} \frac{\partial Z_n}{\partial w}\bigg|_{z=-h_j^-} \right], \; j = 1,2,...,M-1 . \tag{30}$$

From the last four terms in the left-hand side of the variational equation (16), defined on the vertical interfaces at $x = a$ and $x = b$, respectively, we obtain the following end-conditions for the mode amplitudes $U_n(x)$,

$$C_n^{(m)} dU_n / dx_1 + D_n^{(m)} U_n = F_n^{(m)}, \quad n = 0,1,2,..., m = 1,3, \tag{31}$$

where the coefficients $C_n^{(m)}, D_n^{(m)}, F_n^{(m)}$, $m = 1,3$, are defined in terms of the physical parameters at the end points $x = a$ and $x = b$.

6 NUMERICAL RESULTS AND DISCUSSION

In the present work, the numerical solution of the above coupled-mode system is obtained by truncating the series (17) and using a finite difference scheme based on a uniform grid and second-order central differences to approximate derivatives. In order to further enhance the efficiency of the present model, future work is focused on the application of *p*-Finite Element Methods, in conjunction with grid adaptation techniques based on the spatial variability of the system coefficients a_{mn}, b_{mn} and c_{mn}.

As an example, we consider underwater acoustic propagation in a coastal environment, characterized by variable seabed boundary. As before, the upper layer (layer 1) is sea water of density and speed of sound $\rho_1 = 1g/cm^3$, $c_1 = 1500\,m/\mathrm{s}$. The lower layer (layer 2) corresponds to sand-silt-clay sediment with properties $\rho_2 = 1.5g/cm^3$, $c_2 = 1700\ m/\mathrm{s}$, terminated at the impermeable (rigid) bottom which is located at a depth $z = -100m$. The geometry of the internal interface is defined as

$$h_1(x) = 500 - 450 \tanh\left[2\pi \left(\left(\frac{x-3000}{4000} \right) - 0.5 \right) \right], \quad a \le x \le b \quad , \tag{32}$$

where *a*=2800m and *b*=7200m.

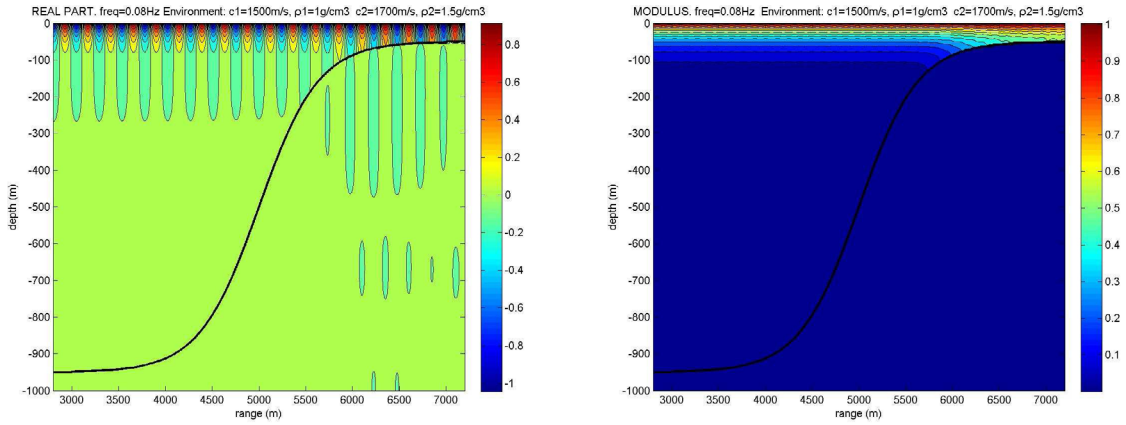


Figure 6: Acoustic pressure (real part). Frequency 0.08 Hz. Excitation by the first surface (gravity) mode, *n*=1. Parameters: $\rho_1 = 1g/cm^3, c_1 = 1500m/\mathrm{s}$, $\rho_2 = 1.5g/cm^3, c_2 = 1700m/\mathrm{s}$.

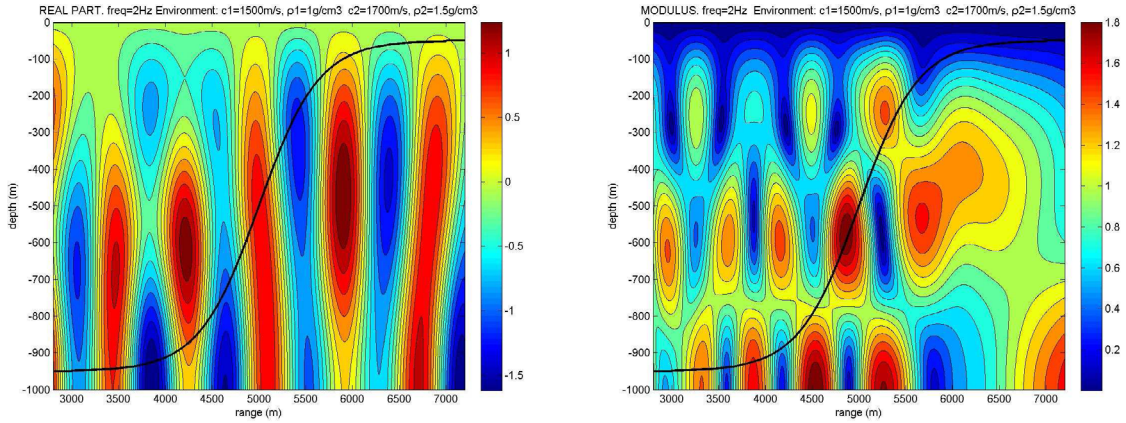


Figure 7: Acoustic pressure (modulus). Frequency 2 Hz. Excitation by the second interior (acoustic) mode *n*=2, and acoustic parameters same as in Fig.6.

Numerical results concerning the real part of the calculated wave field and its modulus, forced by the first incident mode *n*=1, when the waveguide is excited at frequency $f = 0.08Hz$ and by the second incident mode *n*=2, for $f = 2Hz$ are shown in Fig. 6 and 7, respectively, as obtained by present method. In the case of Fig.4 there is only one propagating mode (*n*=1), while in Fig.7 the number of propagating modes is 3 (*n*=1,2,3) both in the regions of incidence and transmission. The local-mode series is truncated by keeping 5 totally modes, and the coupled-mode system is discretized using 1200 segments, which were proved to be enough for numerical convergence. It can be seen that in the low frequency case $f = 0.08Hz$ the gravity mode interacts very little with the rest of the modes. Near the shallow end of the domain small part of the energy is transmitted from the upper medium (water) to the lower medium (sediment). Also, in the case of higher frequency $f = 2Hz$ the interaction of the first acoustic mode (*n*=2) with the gravity mode (*n*=1) is very small and the generated free-surface elevation is negligible. Thus, the acoustic mode is difficult to be observed at the free surface, however this could become possible at lower depths under the free surface.

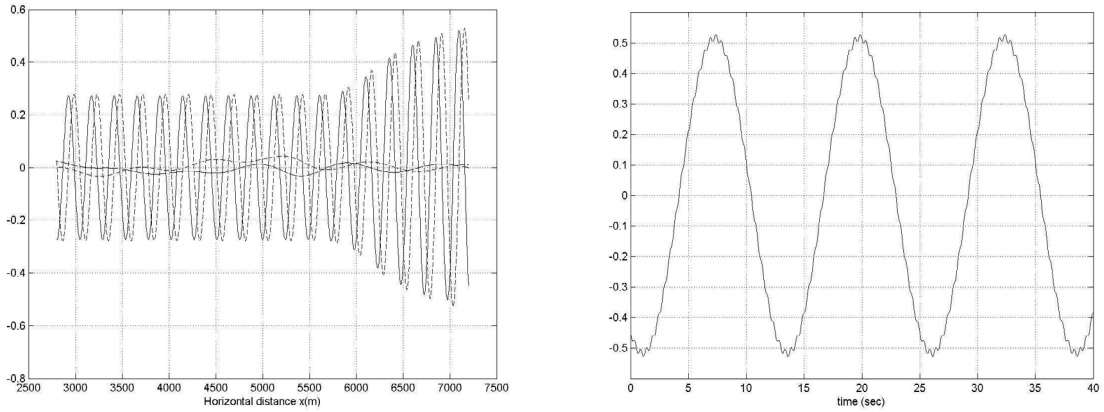


Figure 8: (a) Pressure at depth 50m for frequency f=0.08Hz with excitation by the free surface (gravity) mode n=1 of unit amplitude, and for f=2 Hz with excitation by the second interior (acoustic) mode n=2 of 10 times smaller amplitude. (b) Pressure signal at depth 50m at the shallow end of the domain b=7200m.

As an example we present in Fig.6 the pressure distribution at depth 50m (location of the seabed in the shallow end of the domain) calculated for frequency f=0.08Hz with excitation by the free surface (gravity) mode n=1 of unit amplitude, and for f=2 Hz with excitation by the second (acoustic) mode n=2 of 10 times smaller amplitude, which is in compatibility with the spectrum characteristics of tsunami waves generated by bottom dislocation. For this example the pressure signal at depth 50m at the shallow end of the domain is plotted in the right subplot of Fig.8, where the acoustic-gravity waves associated with the mode n=2 are clearly observable as high frequency ripples. The latter travel at significant higher speed that the free surface waves associated with the first mode at f=0.08Hz leaving an option for an advance warning sufficient for evacuation and protection.

7 CONCLUSIONS

In this work an improved coupled-mode method is presented for the efficient solution of the problem of time-harmonic propagation and scattering of acoustic-gravity waves in a non uniform stratified waveguide. The problem is governed by the Helmholtz equation, with variable coefficients, in conjunction with the linearized free-surface boundary condition associated with gravity waves. Our method is based on an enhanced local-mode series for the representation of the wave field, including additional modes, accounting for the effects of the inhomogeneous interfaces. In the case of multilayered waveguides, the local vertical eigenvalue problems are treated by h- and p-FEM, exhibiting robustness and good rates of convergence. In order to further enhance the efficiency of the present model, current work is focused on the application of hp-FEM for the solution of the coupled systen on thehorizontal plane, in conjunction with grid adaptation techniques based on the spatial variability of the system coefficients. Among several other advantages, the present method can be naturally extended to treat wave propagation and scattering problems in 3D multi-layered waveguides.

REFERENCES

[1] Ardhuin, F. Stutzmann E., Schimmel, M. and Mangeney A. Ocean wave sources of seismic noise. *J. Geophys. Res.*, (2011) **116**: C09004, doi:10.1029/2011JC006952.

[2] Ardhuin F., and Herbers, T.H.C. Noise generation in the solid Earth, oceans and atmosphere, from nonlinear interacting surface gravity waves in finite depth. *J. Fluid Mech.* (2013) **716**:316–348.

[3] Stiassnie, M. Tsunamis and acoustic-gravity waves from underwater Earthquakes. *J. Eng Math* (2010) **67**:23–32, doi: 10.1007/s10665-009-9323-x.

[4] Boyles, C.A. *Acoustic waveguides. Applications to ocean science*. J.Wiley & Sons, (1984).

[5] Jensen, F., Kupperman, W., Porter, M. and Schmidt, H. *Computational Ocean Acoustics*, AIP Press, (1994).

[6] Salomons, E.M. *Computational Atmospheric Acoustics*, Kluwer Academic Publishers, Dordrecht (2001).

[7] Lee, D. and Schultz, M.H. *Numerical Ocean Acoustic Propagation in Three Dimensions*, World Scientific, Singapore (1995).

[8] Belibassakis, K.A., Athanassoulis, G.A, Papathanasiou, T.K., Filopoulos, S.P. and Markolefas S. Acoustic wave propagation in inhomogeneous, layered waveguides based on modal expansions and hp-FEM. *Wave Motion* (2014) **51**:1021–1043.

[9] Athanassoulis, G.A and Belibassakis, K.A. A consistent coupled-mode theory for the propagation of small-amplitude water waves over variable bathymetry regions. *J. Fluid Mech.* (1999) **389**: 275-301.

[10] Athanassoulis, G.A., Belibassakis, K.A., Mitsoudis, D.A., Kampanis, N.A. and Dougalis, V.A. Coupled-mode and finite-element solutions of underwater sound propagation problems in stratified acoustic environments. *J. Comput. Acoust.* (2008) **16**(1): 83-116.

[11] Hughes, T.J.R. *The Finite Element Method, Linear Static and Dynamic Finite Element Analysis,* Dover Publications INC, (1987).

[12] Coddington, E. and Levinson, N. *Theory of Ordinary Differential Equations*, McGraw Hill, (1955).

[13] Tichmarsh, E.G. *Eigenfunction expansions*, Calderon Press, (1962).

APPENDIX

In the case of two layers $1,2$, with constant physical properties $\rho_1, c_1, k_1 = \omega/c_1$ and $\rho_2, c_2, k_2 = \omega/c_2$ respectively, the exact analytical solution of the vertical eigenvalue problem is given by

$$Z_{(1)}(z) = B_n\left[b_n \cos(\lambda_{n1} z) + \sin(\lambda_{n1} z)\right] \quad , \qquad Z_{(2)}(z) = B_n K_n \cos\left[\lambda_{n2}(z+H)\right] \quad , \qquad \text{(A1)}$$

where

$$\lambda_{n1} = \sqrt{k_1^{\,2} - k_n^{\,2}} \quad , \qquad \lambda_{n2} = \sqrt{k_2^{\,2} - k_n^{\,2}} \quad , \tag{A2}$$

and

$$b_n = \frac{\lambda_{n1}}{\mu} , \qquad K_n = \frac{b_n \cos(-\lambda_{n1} h_1) + \sin(-\lambda_{n1} h_1)}{\cos\left(\lambda_{n2} h_2\right)} . \tag{A3}$$

In this case, the eigenvalues k_n are found as the roots of the equation

$$\frac{\rho_2 \lambda_{n1}}{\rho_1 \lambda_{n2}} \left[b_n \sin(\lambda_{n1} h_1) + \cos(-\lambda_{n1} h_1) \right] = \tan(\lambda_{n2} h_2) \left[\sin(\lambda_{n1} h_1) - b_n \cos(-\lambda_{n1} h_1) \right], \tag{A4}$$

which expresses the continuity of $\rho^{-1} \partial Z / \partial z$ across the interface at $z = -h_1$. The remaining constants B_n, $n = 1, 2, \ldots.$, of the above solution can be fixed by appropriate normalization.

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis(Eds)

A SENSITIVITY BASED OPTIMIZATION APPROACH FOR AEROACOUSTIC PROBLEMS

STEFANIE NOWAK*, DÖRTE C. STERNEL* AND MICHAEL SCHÄFER*

*Institute of numerical Methods in Mechanical Engineering, TU Darmstadt,
Dolivostrasse 15, 64293 Darmstadt, Germany
e-mail: nowak@fnb.tu-darmstadt.de, web page: http://www.fnb.tu-darmstadt.de/

Key words: aeroacoustics, gradient-based optimization, sensitivity equations, finite-volume method

Abstract. This paper proposes an approach for the sensitivity computation of an aeroacoustic problem via the solution of the continuous sensitivity equations. The equations are derived by differentiating the coupled system (Navier-Stokes equations and the linearized Euler equations) with respect to design parameters. The obtained acoustic sensitivity information can be used to analyze the flow control variable's influence on the acoustic field or to calculate the gradient of an objective functional within aeroacoustic optimization problems. The components of the coupled sensitivity solver are systematically verified.

1 INTRODUCTION

The prediction and optimization of noise generated by fluid flow is of high interest in numereous fields in industry and current research. Many occurances in the aeroacoustics field are related to low-speed flows, e.g. wind turbines or automobile's noise. Numerical simulations of aeroacoustic problems support practical experiments and provide eligible measurements for noise reduction.
The sound prediction can be based on the simulation of the compressible Navier-Stokes equations (direct noise computation, DNC). DNC is inefficient for low Mach number flows, because of the different spatial and temporal scales in the fluid flow and the acoustics. Hybrid approaches consider the flow and the acoustic field separately and, therefore, are cost saving. Lighthill's acoustic analogy [1] and his further work [2] were important milestones in the aeroacoustics discipline. Alternatively to the acoustic analogies splitting approaches are considered. In [3] the acoustic/viscous splitting is introduced. Here, the compressible field is considered as the superposition of an incompressible flow with an acoustic perturbation. The flow is computed via the incompressible Navier-Stokes equations. Shen and Sørensen proposed a modified version of the method in [4], where the sound field is obtained with the linearized Euler equations.

However, not only the behavior of the flow and the originated sound is of interest, but how to affect the flow and consequently the sound to reach given aims. An objective could be, for instance, to minimize the sound intensity in a given area.
Optimization methods for the resolution of minimization problems can be divided into strategies without and with determination of the cost functional's derivative. An efficient possibility of computing this gradient is to solve the continuous sensitivity equations (CSE). These equations arise from differentiating the state equations with respect to a design parameter. In [5] an outline of sensitivity analysis of flow problems is presented and Gunzburger [6] discusses sensitivity analysis in the context of flow control and optimization. In addition to the CSE method, there is the discrete sensitivity equation approach, in which the total derivative of the flow approximation with respect to the design parameter is calculated. In Kleiber et al. [7] a discussion of the two approaches is found.
A method for computing transient flow sensitivities with the CSE method is presented by Ilinca et al. [8] and Hristova et al. [9]. After extensive research the authors have not found any application of the continuous sensitivity equation method for the LEE in any scientific contents.
This paper presents the theoretical background of the CSE of a coupled aeroacoustics problem. The coupling of the sensitivity equations and verification test cases of each component of the coupled solver are considered.

2 GOVERNING EQUATIONS

The aeroacoustic equations are derived via employment of the expansion about incompressible flow proposed by Shen and Sørensen [4] and are further developed by Kornhaas [10]. Here the compressible flow field at low Mach numbers is composed of an incompressible backround flow (superscript inc) with a superimposed acoustic perturbation (superscript ac). The Einstein summation convention is used throughout the paper. The composition is given by

$$\rho = \rho^{inc} + \rho^{ac}, \qquad u_i = u_i^{inc} + u_i^{ac}, \qquad p = p^{inc} + p^{ac}, \tag{1}$$

with density ρ, velocity u_i and pressure p. The following paragraphs describe the models for the incompressible flow and the acoustics.

Flow equations The incompressible and unsteady fluid flows are modelled by the Navier-Stokes equations (NSE) [11]. The conservation equations of mass and momentum can be written as

$$\frac{\partial u_i^{inc}}{\partial x_i} = 0, \tag{2}$$

$$\rho^{inc}\frac{\partial u_i^{inc}}{\partial t} + \rho^{inc}\frac{\partial u_i^{inc} u_j^{inc}}{\partial x_j} = \frac{\partial \tau_{ij}^{inc}}{\partial x_j} - \frac{\partial p^{inc}}{\partial x_i} + \rho^{inc} f_i, \tag{3}$$

with the time t, Cartesian coordinates x_i and the external body forces f_i. With the dynamic viscocity μ the stress tensor is given by

$$\tau_{ij}^{inc} = \mu \left(\frac{\partial u_i^{inc}}{\partial x_j} + \frac{\partial u_j^{inc}}{\partial x_i} \right). \tag{4}$$

Acoustic equations Under the assumption that the acoustic variables are small compared to the flow variables, the linearized Euler equations (LEE) [12] are used to obtain the acoustic variables.

The LEE are given by

$$\frac{\partial \rho^{ac}}{\partial t} + \rho^{inc} \frac{\partial u_i^{ac}}{\partial x_i} + u_i^{inc} \frac{\partial \rho^{ac}}{\partial x_i} = 0, \tag{5}$$

$$\rho^{inc} \frac{\partial u_i^{ac}}{\partial t} + \rho^{inc} u_j^{inc} \frac{\partial u_i^{ac}}{\partial x_j} + \frac{\partial p^{ac}}{\partial x_i} = 0, \tag{6}$$

$$\frac{\partial p^{ac}}{\partial t} + c^2 \rho^{inc} \frac{\partial u_i^{ac}}{\partial x_i} + c^2 u_i^{inc} \frac{\partial \rho^{ac}}{\partial x_i} = \frac{\partial p^{inc}}{\partial t}, \tag{7}$$

with the speed of sound c.

3 GRADIENT BASED OPTIMIZATION

A general aeroacoustic optimization problem is composed of a given objective functional $\mathcal{J}$ depending on flow state variables ϕ, for example, fluid velocity or pressure and acoustic state variables ψ, for example, the sound pressure or the particle velocity. Both state variables depend on the design variables $\mathbf{a}$, for example, the inlet velocity of the flow or shape parameters. Mathematically an aeroacoustic optimization problem can be formulated as

$$\min_{\mathbf{a} \in \mathbb{R}^{\mathbf{n}}} \mathcal{J}(\psi(\mathbf{a}), \mathbf{a}) \quad \text{subject to} \quad \begin{cases} \mathbf{F}(\phi(\mathbf{a}), \mathbf{a}) = \mathbf{0} \\ \mathbf{A}(\psi(\mathbf{a}), \phi(\mathbf{a}), \mathbf{a}) = \mathbf{0}. \end{cases} \tag{8}$$

Here, the constraints $\mathbf{F} = \mathbf{0}$ and $\mathbf{A} = \mathbf{0}$ are the governing equations of the fluid and the acoustics. Other side constraints are possible, like restrictions to the inlet velocity. The objective function $\mathcal{J}$ is an essential concept in optimization problems. It describes what is to be optimized, for instance, the sound intensity. $\mathcal{J}$ is usually not directly dependent on the design variables, but via the acoustic state variables $\mathcal{J}$ depends indirectly on the design variables.

Calculation of the objective function's gradient Gradient-based optimization methods use information about the objective function's gradient and its evaluations. The *i-th* component of the gradient of the objective function can be written as

$$\frac{d\mathcal{J}}{da_i} = \frac{\partial \mathcal{J}}{\partial \psi_i} \frac{\partial \psi_j}{\partial a_i} + \frac{\partial \mathcal{J}}{\partial a_i}. \tag{9}$$

$d\mathcal{J}/da_i$ denotes the total derivative of the objective function with respect to the design parameter a_i. In this case, the objective function depends only on the acoustic variables and the control variables. In Eq. (9) the term $\partial\psi_j/\partial a_i$, the sensitivity with respect to a design parameter a_i, is rather complicated to compute. There are several procedures for computing the sensitivities of flow and acoustics. The next section shows one of these methods for the continuous sensitivities, starting with the theoretical background of the coupling, and thereafter, the derivation of the continuous aeroacoustic sensitivity equations.

4 COUPLED SENSITIVITY EQUATION SYSTEM

To obtain the CSE system the constraints in (8) have to be differentiated with respect to the design parameters. This leads to the following equations

$$\frac{\partial \mathbf{F}}{\partial \mathbf{a}} + \frac{\partial \mathbf{F}}{\partial \phi}\frac{d\phi}{d\mathbf{a}} = 0, \tag{10}$$

$$\frac{\partial \mathbf{A}}{\partial \mathbf{a}} + \frac{\partial \mathbf{A}}{\partial \phi}\frac{d\phi}{d\mathbf{a}} + \frac{\partial \mathbf{A}}{\partial \psi}\frac{d\psi}{d\mathbf{a}} = 0. \tag{11}$$

For the purposes of simplicity, the following abbreviation is introduced

$$\phi_k = \frac{\partial \phi}{\partial a_k}. \tag{12}$$

Differentiation of the unsteady Navier-Stokes equations (2)-(3) with respect to the k-th design parameter a_k results in the unsteady continuous sensitivity equations of the unsteady NSE. In consideration of the chain rule, permutation of the differential operators and notation (12) the equations are given by

$$\frac{\partial u^{inc}_{j,k}}{\partial x_j} = 0, \tag{13}$$

$$\rho\frac{\partial u^{inc}_{i,k}}{\partial t} + \frac{\partial}{\partial x_j}\left(\rho u^{inc}_{i,k}u^{inc}_j + \rho u^{inc}_i u^{inc}_{j,k}\right) - \frac{\partial}{\partial x_j}\left[\mu\left(\frac{\partial u^{inc}_{i,k}}{\partial x_j} + \frac{\partial u^{inc}_{j,k}}{\partial x_i}\right)\right] = \frac{\partial p^{inc}_k}{\partial x_i}, \tag{14}$$

assuming, that the material properties ρ and μ are constant and the volume force is insensitive to external influences. The boundary conditions for the CSE are obtained by differentiation of the corresponding boundary conditions for the NSE with respect to the design parameter a.
The sensitivity equations of the linearized Euler equations are obtained via differentiation of the linearized Euler equations with respect to the *k-th* design variable a_k. Hence

differentiation of (5)-(7) yields to

$$\frac{\partial \rho_k^{ac}}{\partial t} + \rho^{inc}\frac{\partial u_{i,k}^{ac}}{\partial x_i} + u_i^{inc}\frac{\partial \rho_k^{ac}}{\partial x_i} = -u_{i,k}^{inc}\frac{\partial \rho^{ac}}{\partial x_i}, \tag{15}$$

$$\rho^{inc}\frac{\partial u_{i,k}^{ac}}{\partial t} + \rho^{inc}u_j^{inc}\frac{\partial u_{i,k}^{ac}}{\partial x_j} + \frac{\partial p_k^{ac}}{\partial x_i} = -\rho^{inc}u_{j,k}^{inc}\frac{\partial u_i^{ac}}{\partial x_j}, \tag{16}$$

$$\frac{\partial p_k^{ac}}{\partial t} + c^2\rho^{inc}\frac{\partial u_{i,k}^{ac}}{\partial x_i} + c^2 u_i^{inc}\frac{\partial \rho_k^{ac}}{\partial x_i} = \frac{\partial p_k^{inc}}{\partial t} - c^2 u_{i,k}^{inc}\frac{\partial \rho^{ac}}{\partial x_i}. \tag{17}$$

The obtained equations have the same form as the LEE. Additional terms can be considered as additional source terms. The coupling between sensitivity equations for flows and acoustics is realized by these terms using flow sensitivities and acoustic variables.

5 NUMERICAL METHODS

The numerical solution of the coupled system is realized as an integrated procedure, see Figure 1. The acoustic sensitivities do not influence the flow sensitivities.

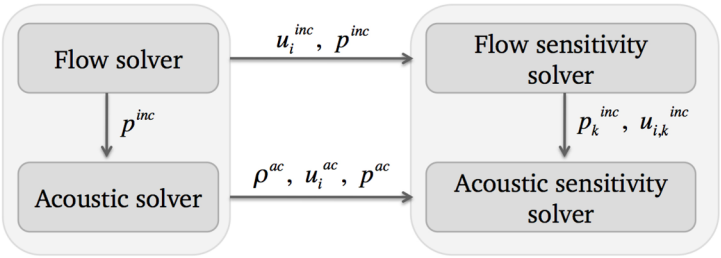


Figure 1: Numerical realization of the coupled approach as an integrated procedure

Before computing the flow sensitivities and the acoustic sensitivities, the NSE (2)-(3) and LEE (5)-(7) have to be solved. This is done by using our in-house solver FASTEST which applies a fully conservative finite-volume approach to solve the incompressible NSE and the linearized Euler-equations on a collocated, block structured and cell centered grid [12, 13]. After each time-step the flow and acoustic quantities are transferred to the sensitivity solver via MPI (Message Passing Interface).

Flow sensitivity solver Analogously to the Navier-Stokes equations the spatial discretization of the sensitivity equations utilizes the finite-volume method. The discretization of the sensitivity equations of the NSE for steady problems can be read in detail in [15] and [16]. The unsteady sensitivity equations contain one additional term:

$$\rho\frac{\partial u_{i,k}^{inc}}{\partial t}. \tag{18}$$

The time discretization is performed with an implicit Euler scheme of first and second order.

Acoustic sensitivity solver Analogously to the linearized Euler equations the acoustic sensitivities are computed via a high-resolution scheme, which solves a Riemann problem at the cell faces. The high-resolution scheme combines the second-order Lax-Wendroff method and the first order Godunov method with the help of flux-limiters. Kornhaas presents the high-resolution scheme in detail [10]. The time discretization is done with first and second order implicit Euler scheme.

6 VERIFICATION

To verify the implemented method, the numerical order achieved is compared to the analytical order for each method of the coupled code. At first, the order of the unsteady continuous sensitivity equation's discretization method is presented followed by the order of the sensitivity equations of the linearized Euler equations method.
To assess the solution quality, the error

$$e_h = p^{ac}_{num,h} - p^{ac}_{ana} \tag{19}$$

is considered, where p^{ac}_{num} is the numerically computed sound pressure and p^{ac}_{ana} is the analytical sound pressure. The discrete L_2-norm of the error is defined by

$$\|e\|_{h,L_2} = \sqrt{\frac{1}{N}\sum_{i=1}^{N} e_h^2}, \tag{20}$$

where N is the number of control volumes.

6.1 Verification of the acoustic sensitivity equations

The first verification test case is calculated without background flow ($u_i^{inc} = 0$). The computational domain and the initial location of the sound wave are presented in Figure 2. The computations are done on four grid levels with 16384, 8192, 4096 and 2046 control volumes in x-direction.
The initial values of the acoustic quantities are as follows:

$$p^{ac} = 2^{-1200(x-0.25)^2}, \tag{21}$$

$$u^{ac} = 0.001 \cdot 2^{-1200(x-0.25)^2}, \tag{22}$$

$$v^{ac} = 0, \tag{23}$$

$$\rho^{ac} = 0. \tag{24}$$

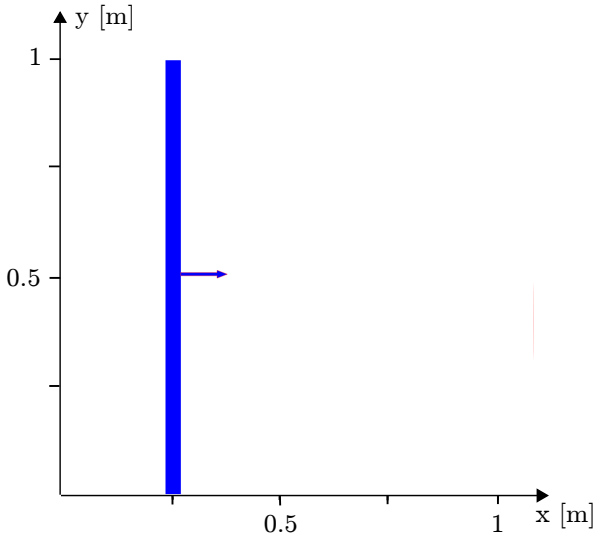


Figure 2: Domain and initial location of the sound wave.

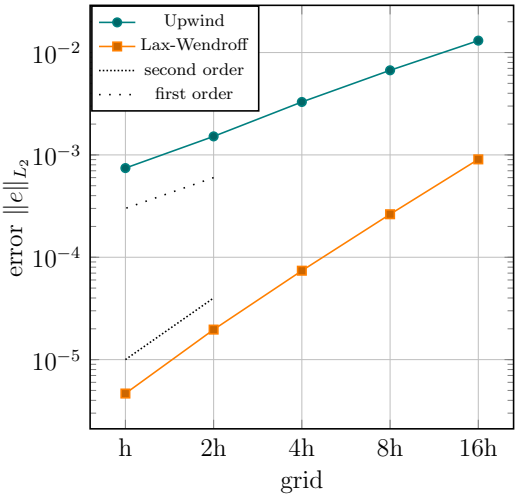


Figure 3: L_2-error depending on the grid for different methods

The wave starts at $x = 0.25$ m. The density is $\rho^{inc} = 1000$ kg/m^3 and the dynamic viscosity is $\mu^{inc} = 10^{-5}$ kg/m s. The corresponding speed of sound is $c = 1$ m/s. The resulting orders are illustrated in Figure 3. One can see that the upwind-scheme converges to an order of one, whereas the Lax-Wendroff scheme converges to an order of two.

6.2 Verification of the flow sensitivity equations

In this section the time discretization scheme implemented in the flow sensitivity solver is verified using the method of manufactured solutions (MMS) [17]. The order of the spatial discretization for the steady NSE is presented in [15]. Direct differentiation of the manufactured solution for the flow provides closed-form expressions for the sensitivities. A time convergence study is performed to assess the temporal accuracy of the sensitivity solutions.

We choose the following exact solution of the unsteady NSE, which is also the inflow boundary condition and depends on the design parameter a:

$$u_{in}(x,y) = a\cos(5t)\sin\left(\frac{\pi x}{2}\right)\cos\left(\frac{\pi y}{2}\right), \tag{25}$$

$$v_{in}(x,y) = -a\cos(5t)\cos\left(\frac{\pi x}{2}\right)\sin\left(\frac{\pi y}{2}\right). \tag{26}$$

$$\tag{27}$$

Differentiation with respect to the design parameter a gives the sensitivities

$$u_k(x,y) = \cos(5t)\sin\left(\frac{\pi x}{2}\right)\cos\left(\frac{\pi y}{2}\right), \tag{28}$$

$$v_k(x,y) = -\cos(5t)\cos\left(\frac{\pi x}{2}\right)\sin\left(\frac{\pi y}{2}\right). \tag{29}$$

$$\tag{30}$$

The computational domain and the sensitivities are shown in Figure 4 and Figure 5.

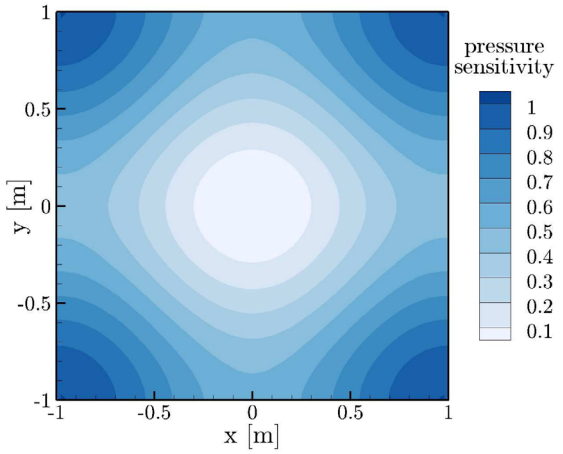


Figure 4: Domain and pressure sensitivity for $t = 0$ s.

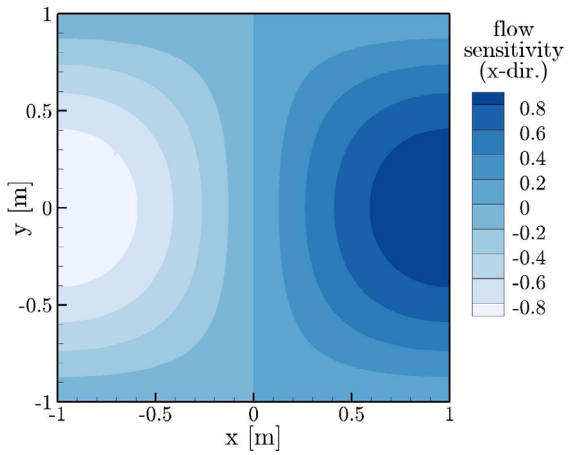


Figure 5: Domain and flow sensitivity in x-direction for $t = 0$ s.

Table 1 shows the time step sizes and the results for the order of the flow sensitivities for the first order implicit Euler scheme. The order of the pressure computation is illustrated at the left side, whereas on the right sight the order of the velocity computation is shown. The order of the pressure and the order of the velocity converge to one. In Table 2 the orders of the pressure and the velocity for the second order implicit Euler scheme are presented. Both orders converge to two.

Table 1: Order of accuracy of the flow sensitivities for the first order implicit Euler scheme

time step size	pressure	order
0.005	0.46531836731	
0.0025	0.47096579783	
0.00125	0.47380853834	0.9903
0.000625	0.47523497813	0.995
0.0003125	0.47594950138	0.998

time step size	u velocity	order
0.005	-0.78628698376	
0.0025	-0.78631729909	
0.00125	-0.78633378171	0.89
0.000625	-0.78634240443	0.93
0.0003125	-0.78634681860	0.97

Table 2: Order of accuracy of the flow sensitivities for the second order implicit Euler scheme

time step size	pressure	order
0.005	0.47917358818	
0.0025	0.47838279892	
0.00125	0.47846168824	2.43
0.000625	0.47847809934	2.27
0.0003125	0.47848192570	2.10

time step size	u velocity	order
0.005	-0.78633065815	
0.0025	-0.78636059225	
0.00125	-0.78636832215	1.95
0.000625	-0.78637006368	2.15
0.0003125	-0.78637047477	2.08

6.3 Comparison of the acoustic sensitivities with the corresponding difference quotient

The acoustic sensitivity solver can be verified by estimating the gradients of the acoustic quantities with respect to a using finite differences (FD). When computing FD the design parameter a is changed by a small amount δa and the solution is recomputed. In this test case only the sensitivity of the acoustic pressure and its difference quotient with respect to a are considered. As a first step, the flow sensitivities are left out. A pressure pulse and its sensitivities are prescribed as input for the sensitivity equations of the LEE:

$$p^{inc} = 10^3 e^{-10^2(x^2+y^2)} cos(200\pi t)\ a, \tag{31}$$

$$p_k^{inc} = 10^3 e^{-10^2(x^2+y^2)} cos(200\pi t). \tag{32}$$

The reference FD acoustic sensitivities are estimated by

$$\left(\frac{\partial p^{ac}}{\partial a}\right)_{FD} = \frac{p^{ac}(a+\delta a) - p^{ac}(a-\delta a)}{2\delta a} + \mathcal{O}(\delta a^2). \tag{33}$$

In the following comparison $\delta a = 10^{-3}$ is chosen. The computational domain is illustrated in Figure 6. As boundary conditions an acoustic outlet is used on each boundary. The grid consists of 128 control volumes in x- and y-direction. The Courant number for the acoustic computation is 0.1. The acoustic pressure sensitivity and the acoustic pressure are monitored at point $\mathrm{P}(0, 0.5)$. The difference quotient and the sensitivity of the acoustic pressure in P are shown in Figure 7. The comparison shows a good agreement between the sensitivity and the difference quotient. Only minor differences of the gradients are perceptible.

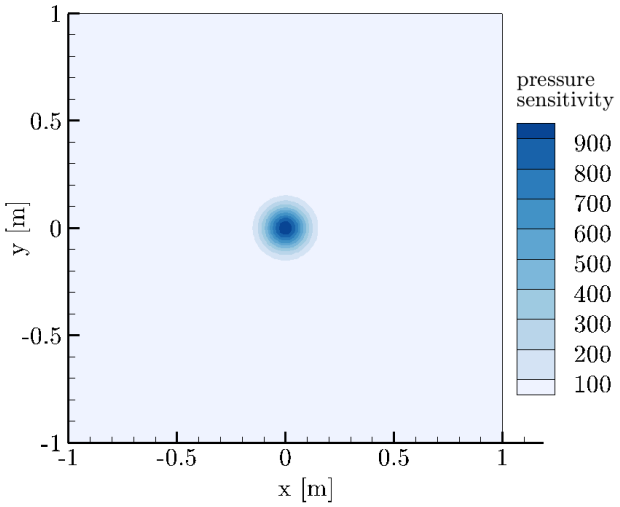


Figure 6: Domain and pressure sensitivity for $t = 0$ s.

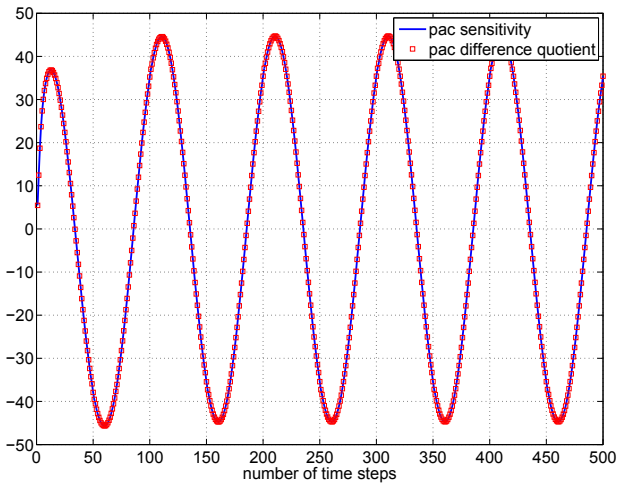


Figure 7: Difference quotient and the sensitivity of the acoustic pressure in P after $t = 0.05$ s.

7 CONCLUSION

A general sensitivity based optimization approach for aeroacoustic problems has been presented. We introduced the basic equations describing flow and acoustics, a general optimization problem, and the gradient computation with a coupled sensitivity solver. Furthermore, we presented the mathematical background for the derivation of the unsteady sensitivity equations from the Navier-Stokes equations, and the linearized Euler equations. Hereafter, the mathematical coupling of the solver was shown. In a first step

the numerical results of the sensitivity computation could be verified for each component of the coupled solver via investigation of the order and the comparison with a finite difference approach. The obtained sensitivity values enable a deeper understanding of the whole system and provide information about the influence of all flow parameters on the acoustics. For further investigations, a test case of acoustic sensitivities based on flow sensitivity computation will be considered and an optimization test case with sensitivities will be computed.

REFERENCES

[1] Lighthill, M. *On Sound Generated Aerodynamically. I. General Theory.* Proceedings of The Royal Societey London A, 1952

[2] Lighthill, M. *On Sound Generated Aerodynamically. II. Turbulence as a source of sound.* Proceedings of The Royal Societey London A, 1954

[3] Hardin, J.C. and Pope, D.S. *An Acoustic/Viscous Splitting Technique for Computational Aeroacoustics.* Theoretical and Computational Fluid Dynamics 6, 1994

[4] Shen, W. and Soerensen, J. *Comment on the aeroacoustic formulation of Hardin and Pope.* AIAA Journal, 1999

[5] Stanley, L.G. and Stewart, D.L. *Design sensitivity analysis.* Society for Industrial and Applied Mathematics, Philadelphia, 2002

[6] Gunzburger, M.D. *Perspectives in flow control and optimization.* Society for Industrial and Applied Mathematics, Philadelphia, 2003

[7] Kleiber, M. and Antunez, H. and Hien, T.D. *Parametric Sensitivity in Nonlinear Mechanics.* Whiley, New York, 1997

[8] Ilinca, F. and Hay, A. and Pelletier, D. *Shape sensitivity analysis of unsteady laminar flow past a cylinder in ground proximity.* AIAA Fluid Dynamics Conference and Exhibit, 2006.

[9] Hristova, H. and Etienne, S. and Pelletier, D. and Borggaard, J. *A continuous sensitivity equation method for time-dependent incompressible laminar flows.* International Journal for Numerical Methods in Fluids, 2006

[10] Kornhaas, M. *Effiziente numerische Methoden fr die Simulation aeroakustischer Probleme mit kleinen Machzahlen.* Ph.D. dissertation, TU Darmstadt, 2012

[11] Schäfer, M. *Computational Engineering - Introduction to Numerical Methods.* Springer, 2006

[12] Kornhaas, M. and Sternel, D.C. and Schäfer, M. *Efficient numerical simulation of aeroacoustics for low Mach number flows interacting with structures.* Computational Mechanics, to appear

[13] Durst, F. and Schäfer, M. *A parallel block-structured multigrid method for the prediction of incompressible flows.* 1996

[14] Leveque, R.J. *Finite Volume Methods for Hyperbolic Problems.* Cambridge Texts in Applied Mathematics, 2002

[15] Michaelis, J. *Parallele Mehrgitter-Strategien zur Sensitivitätsanalyse und Strömungsoptimierung.* Ph.D. dissertation, TU Darmstadt, 2012

[16] Siegmann, J. *Sensitivitätsanalyse zur gradientenbasierten Formoptimierung von Strömungsproblemen.* Ph.D. dissertation, TU Darmstadt, 2012

[17] Salari, K. and Knupp, P. *Code Verification by the Method of Manufactured Solutions.* Technical Report SAND2000 - 1444, Sandia National Laboratories, 2000 International Journal for Numerical Methods in Fluids, 1996

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis(Eds)

COUPLED INVERSE PROBLEMS AND VISUALIZATION OF ATMOSPHERE-OCEAN SYSTEM

SERGEY I. KABANIKHIN[*,†] **AND OLGA I. KRIVOROTKO**[*,†]

[*]Institute of Computational Mathematics and Mathematical Geophysics SB RAS
(ICM&MG SB RAS)
Prospect Ak. Lavrentjeva 6, 630090, Novosibirsk, Russia
e-mail: kabanikhin@sscc.ru, web page: http://labmpg.sscc.ru/kabanikh.html

[†]Novosibirsk State University (NSU)
Pirogova street 2, 630090, Novosibirsk, Russia
e-mail: krivorotko.olya@mail.ru, web page: http://labmpg.sscc.ru/krivorotko.html

Key words: Shallow Water Equations, Dirichlet Problem, Conjugate Gradient Method, Combined Data, Visualization System.

Abstract. We investigate two different inverse problems of determining the tsunami source using two different additional data, namely underwater measurements and satellite wave-form images, and combination of these two inverse problems. We investigate gradient-type methods for inverse problem solutions and show that combination of two types of data allows one to increase stability and convergence of numerical inverse problem solutions. Results of numerical experiments of the tsunami source reconstruction are presented and discussed. We present the 3D GIS visualization and information atmosphere-ocean system with embedded described mathematical tools of simulation of processes in atmosphere and ocean.

1 INTRODUCTION

Atmosphere-Ocean system is mathematically described by systems of hyperbolic equations. The parameter identification of Atmosphere-Ocean system using combined additional measurements is called coupled inverse problem for hyperbolic equations. These problems are ill-posed, *i.e.* their solutions are not unique or/and unstable, and should be regularized [1, 2].

As an example we consider coupled inverse source problem for the linear shallow water equations that use for describing long waves (tsunamis). Most suitable physical models related to simulation of tsunamis are based on shallow water equations (1), [3, 4]. There exist many numerical approaches for solving shallow water equations such as finite-difference, finite volume, finite element, *etc* [5]. An overview of methodologies and techniques related

to estimation of tsunami source characteristics are given in [6, 7, 8, 2]. The most of them consists in determining the tsunami source using additional measurements of a passing wave (this problem is often called inverse tsunami problem) such as DART (Deep-ocean Assessment and Reporting of Tsunamis) buoys positioned on the ocean floor, tide gauges measurements, satellite wave-form images, *etc.* Our goal is to reconstruct the tsunami source using a combination of two types of data: DART buoys and satellite wave-form image. We show that using a combination of two types of data allows one to increase the stability and efficiency of tsunami source reconstruction [9, 10].

The paper is organized as follows. In Section 2 we describe the statement of three inverse problems. In Section 3 we consider a variational formulation of the inverse tsunami problem for two types of measured data: DART data (inverse problem 1) and satellite image data measured on the part of the water surface (inverse problem 2), and then we consider the combined inverse problem (inverse problem 3). We compare two inverse problems and their combination and show the benefits of usage of combined data. Results of numerical experiments of the tsunami source reconstruction are presented in Section 4. In Section 5 we present 3D GIS visualisation and information software for tsunami simulation and run-ups with real bathymetry for the specified sea coast, as well as modelling of earthquakes, floods and other natural hazards.

2 STATEMENT OF THE PROBLEMS

The ocean domain being considered is bounded from above by the free water surface $\eta(x, y, t)$, and from below, by the bottom relief $H(x, y) > 0$. We assume that the computational time T is not large enough for the wave to reach the edges of the domain, and therefore we can set homogeneous boundary conditions at the boundary of the domain $\Omega := (0, L_x) \times (0, L_y)$ (figure 1). We formulate the initial boundary-value problem in the Cartesian coordinate system

$$\begin{cases} \mathcal{L}\eta := \eta_{tt} - \operatorname{div}(gH(x,y)\operatorname{grad}\eta) = 0, & t \in (0,T); \\ \eta|_{t=0} = q(x,y), \quad \eta_t|_{t=0} = 0, & (x,y) \in \Omega; \\ \eta|_{\partial\Omega_T} = 0, & \Omega_T := \Omega \times (0,T) \end{cases} \tag{1}$$

for the linear equations of shallow water theory in terms of the free surface without external forces, *e.g.* the Coriolis force and bottom friction [4]. Here $H \in \mathbf{H}^1(\Omega)$ is a known function describing the bottom relief (bathymetry), $q \in \mathbf{H}^2(\Omega)$ is a tsunami source which is supposed to have a compact support belonging to Ω, $g = 9.8\ [m/s^2]$. Further, we will use notation $c(x, y) = \sqrt{gH(x, y)}$ that describes the tsunami propagation velocity according to the long-wave theory.

The direct tsunami problem (1) consists in determining of a function $\eta \in C(\Omega_T; \mathbf{H}^2(\Omega))$ in the domain Ω by known functions $H(x, y)$ and $q(x, y)$.

Let us consider three inverse problems for linear shallow water equations:

Inverse problem 1 (IP 1): find $q(x, y)$ from (1) using function $H(x, y)$ and data

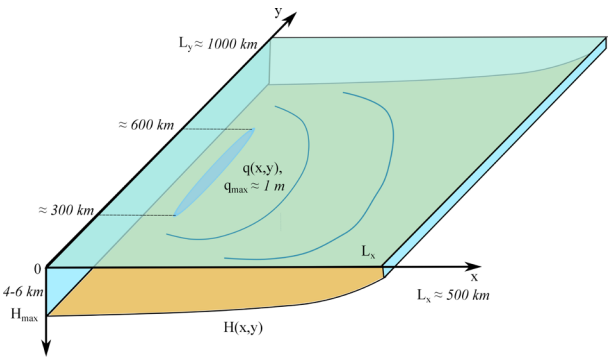


Figure 1: Domain of calculation of direct and inverse problems.

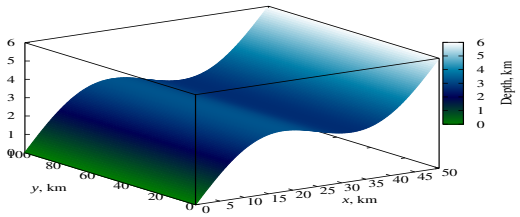


Figure 2: The 1D bottom relief $H(x)$.

$f_m^\varepsilon(x,y,t)$ from underwater systems (DART buoys, tide gauges measurements) at ε-neighborhoods of points $(x_m, y_m) \in \Omega$

$$\eta(x,y,t) = f_m^\varepsilon(x,y,t), \qquad x \in (x_m - \varepsilon, x_m + \varepsilon),\ y \in (y_m - \varepsilon, y_m + \varepsilon),\ \varepsilon > 0, \\ t \in (T_m^{(1)}, T_m^{(2)}),\ m = 1, 2, \ldots, M. \tag{2}$$

Inverse problem 2 (IP 2): find $q(x,y)$ from (1) using function $H(x,y)$ and satellite altimeters data $F_2(x,y)$

$$\eta(x,y,T) = F_2(x,y), \quad (x,y) \in \omega \subset \Omega,\ T > 0. \tag{3}$$

Here $\omega := (l_x^{(1)}, l_x^{(2)}) \times (l_y^{(1)}, l_y^{(2)})$ is a subset of Ω.

Inverse problem 3 (combined IP 3): find $q(x,y)$ from (1)-(3) using function $H(x,y)$, measured data $f_m^\varepsilon(x,y,t)$, $m = 1, \ldots, M$, and $F_2(x,y)$.

Let us present inverse problems 1, 2 and 3 in the operator form: $A_i q = F_i$, $i = 1,2,3$. Here $A_1 : \mathbf{H}^2(\Omega) \mapsto C(\Omega_T; E^M)$, $A_2 : \mathbf{H}^2(\Omega) \mapsto L_2(\Omega)$ and $A_3 := (A_1, A_2)^T$, $F_1 := (f_1^\varepsilon, f_2^\varepsilon, \ldots, f_M^\varepsilon) \in E^M$ is the vector of discrete output data depends on (x,y,t), $F_3 = (F_1, F_2)^T$, E^M is Euclidean space. The inverse problem 3 is ill-posed because A_3 is a compact operator [1]. The compactness of operators A_1 and A_2 is established in papers [2] and [11, 12, 9], respectively. We will find the solution $q(x,y)$ of inverse problems in the class of functions $q(x,y) = \sum\limits_{k=1}^{K} q_k(x) \sin(2\pi k y / L_y)$ which means that we regularize our inverse problems using cut Fourier series [1].

3 VARIATIONAL FORMULATION OF INVERSE PROBLEMS

Inverse problems $A_i q = F_i$ can be reduced to the minimization problems $\min\limits_{q \in H^2(\Omega)} J_i(q)$, $i = 1,2$ [1]. Here $J_i(q) = \|A_i q - F_i\|^2$ are cost functions, $i = 1,2$.

In this section we find gradients of cost functions $J_i(q)$, $i = 1,2$, and introduce a cost function $J_3^{(\beta)}(q)$ for the combined IP 3.

3.1 Inverse problem 1

The conditions of well-posedness of IP 1 in one-dimensional case are given in [13]. The algorithm of constructing function $q(x, y)$ in two-dimensional case based on truncated singular value decomposition is proposed in [6, 7, 14].

The cost function $J_1(q)$ for IP 1 has the form:

$$J_1(q) = \|A_1 q - F_1\|^2_{L^2(0,T)} := \sum_{m=1}^{M} \int\limits_{T_m^{(1)}}^{T_m^{(1)}} \int\limits_{x_m-\varepsilon}^{x_m+\varepsilon} \int\limits_{y_m-\varepsilon}^{y_m+\varepsilon} [\eta(x, y, t; q) - f_m^{\varepsilon}(x, y, t)]^2 \, dy \, dx \, dt.$$

Lemma 1 [2]. *The gradient of the cost function $J_1(q)$ has the form $J_1' q = \psi_{1_t}(x, y, 0)$. Here $\psi_1 \in C(\Omega_T; \mathbf{H}^2(\Omega))$ is the weak solution of the following problem:*

$$\begin{cases} \mathcal{L}\psi_1 = R_1(x, y, t), & (x, y) \in \Omega,\ t \in (0, T), \\ \psi_1(x, y, T) = 0,\ \psi_{1_t}(x, y, T) = 0, & (x, y) \in \Omega, \\ \psi_1|_{\partial\Omega_T} = 0, & t \in (0, T), \end{cases} \tag{4}$$

$R_1(x, y, t) = -2 \sum\limits_{m=1}^{M} \{[\eta(x, y, t) - f_m^{\varepsilon}(x, y, t)]\theta(x - x_m + \varepsilon)\theta(x_m + \varepsilon - x) \cdot \theta(y - y_m + \varepsilon)\theta(y_m + \varepsilon - y)\theta(t - T_m^{(2)})\theta(T_m^{(2)} - t)\}.$

3.2 Inverse problem 2

The cost function $J_2(q)$ for IP 2 has the form:

$$J_2(q) = \|A_2 q - F_2\|^2_{L^2(0,T)} := \int\limits_{l_x^{(1)}}^{l_x^{(2)}} \int\limits_{l_y^{(1)}}^{l_y^{(2)}} (\eta(x, y, T) - F_2(x, y))^2 \, dy \, dx.$$

Lemma 2 [8, 15]. *The gradient of the cost function $J_2(q)$ has the form $J_2' q = \psi_{2_t}(x, y, 0)$. Here $\psi_2 \in \mathbf{H}^2(\Omega)$ is the weak solution of the following problem:*

$$\begin{cases} \mathcal{L}\psi_2 = 0, & (x, y) \in \Omega, t \in (0, T); \\ \psi_2(x, y, T) = 0,\ \psi_{2_t}(x, y, T) = R_2(x, y), & (x, y) \in \Omega; \\ \psi_2|_{\partial\Omega_T} = 0, & t \in (0, T) \end{cases}$$

$R_2(x, y) = 2\,(\eta(x, y, T) - F_2(x, y))\,\theta(x - l_x^{(1)})\theta(l_x^{(2)} - x) \cdot \theta(y - l_y^{(1)})\theta(l_y^{(2)} - y).$

3.3 Inverse problem 3

We introduce the cost function $J_3^{(\beta)}(q)$ for IP 3 in the form: $J_3^{(\beta)}(q) = \beta J_1(q) + (1 - \beta)J_2(q)$, $\beta \in [0, 1]$. The gradient of a cost function $J_3^{(\beta)}(q)$ has the form: $J_3^{(\beta)\prime} q = \beta J_1' q + (1 - \beta)J_2' q$.

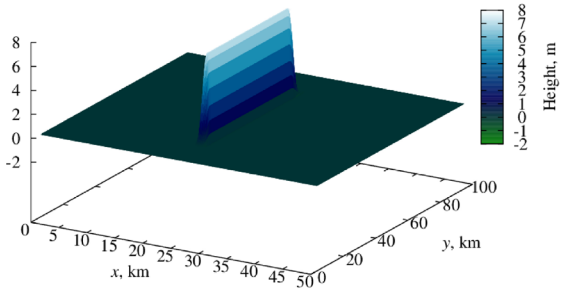


Figure 3: The exact solution $q_e(x, y)$ of inverse problems.

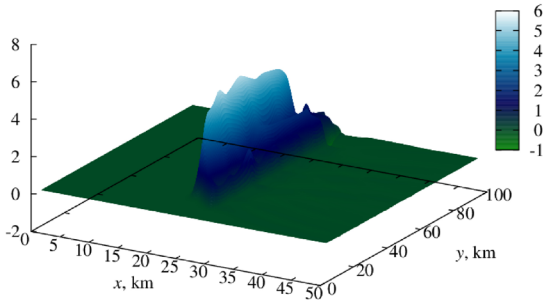


Figure 4: The reconstructed solution $q_5^{(1)}(x, y)$ of IP 1 from the random noisy output data with $\gamma = 3\%$.

4 RESULTS OF NUMERICAL CALCULATION

We apply the conjugate gradient method [16, 1] for solving IP 1, IP 2 and IP 3 numerically.

We choose the following parameters for numerical experiments: $L_x = 50$ km, $L_y = 100$ km, $T = 60$ min, $\varepsilon = 125$ m, $N_x = 750$, $N_y = 500$, $N_t = 600$. The bottom is assumed to be one-dimensional (see figure 2) with the highest $H_{\max} = 6$ km and lowest $H_{\min} = 5$ m average depth of the ocean. We choose an exact solution $q_e(x, y)$ of inverse problems with a wave height $A = 8$ m (see figure 3). We use the explicit finite-difference conservative scheme of the second order approximation [2] with Courant condition $h_t = 0,8 \cdot h_x h_y \left(h_x^2 + h_y^2\right)^{-1/2} / \|c\|_C$. We set data f_m^ε, $m = \overline{1,6}$, and F_2 with "white" noise 1-7%, *i.e.* $f_m^{\varepsilon,\gamma}(x, y, t) = f_m^\varepsilon(x, y, t) + \gamma \text{Random}(f_m^\varepsilon)\|f_m^\varepsilon\|$, $\gamma \in (0.01, 0.07)$. Noise data for IP 1 is generated from the discrete numerical solution of the direct problem in six points (x_m, y_m) equally-spaced on the interval $((40, 15); (47, 89))$. We choose an initial approximation $q_0 = H_{\max}$ which corresponds to an unperturbed sea surface.

We use the stopping condition $J_i(q_n) < \varepsilon_s$, $i = 1, 2$, where choosing of $\varepsilon_s > 0$ based on analysis of deficiency [2]. The behaviour of deficiency as a function of the iteration number n consists of three phases: the initial phase of rapid decrease but short duration, the second phase of slow decrease, and the third phase of almost constant behavior, after some iterations. The numerical results show the minimum value of difference between exact q_e and approximate q_n solutions is achieved between the second and third phases of the deficiency curve versus n.

Let us denote $q_n^{(i)}$, $i = 1, 2$, is n-th approximation of the solution of IP 1 and 2. The reconstructed solution $q_n^{(1)}$ of IP 1 from the random noisy output data $\gamma = 3\%$ is demonstrated on figure 4.

For solving numerically IP 2 we put $\omega = (0, 25) \times (0, 50)$ km. The reconstructed solution $q_n^{(2)}$ of IP 2 from the random noisy output data $\gamma = 3\%$ is demonstrated on figure 5.

The reconstructed solution $q_{n,\beta}^{(3)}$ of IP 3 from the random noisy output data with $\gamma = 3\%$, $\beta = 0.3$, is demonstrated on figure 6. Note, that the location of initial source as well as

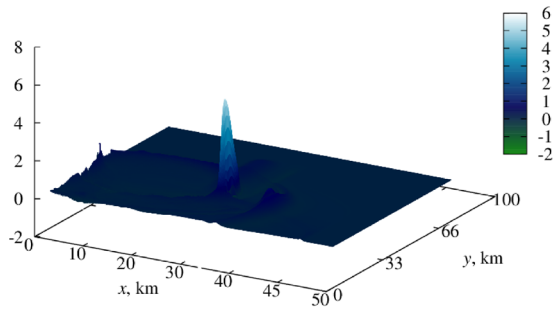


Figure 5: The reconstructed solution $q_{15}^{(2)}(x,y)$ of IP 2 from the random noisy output data with $\gamma = 3\%$.

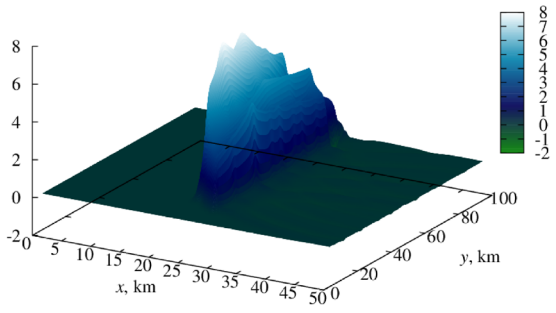


Figure 6: The reconstructed solution $q_{4,\beta}^{(3)}(x,y)$ of IP 3 from the random noisy output data with $\gamma = 3\%$.

its amplitude is reconstructed better than in case of IP 1 and IP 2. The parameter β in combined function $J_3^{(\beta)}(q)$ depends on sensitivity of the functional $J_1(q)$ and $J_2(q)$ (figure 7).

We compare relative accuracy error curves $E_i(n; q^{(i)}; \gamma) = \|q_e - q_n^{(i)}\|/\|q_e\|$, $i = 1, 2$, for IP 1, 2 and $E_3(n; q_\beta^{(3)}; \gamma)$ for IP 3. Figure 7 shows that using of combined underwater systems and satellite data allows one to increase the stability and efficiency of tsunami source reconstruction.

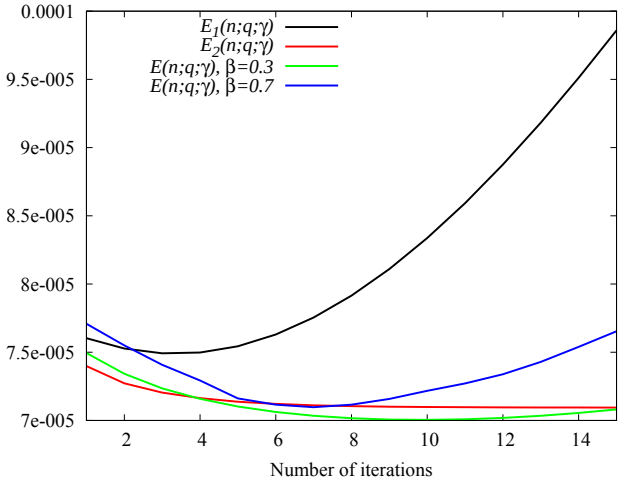


Figure 7: The relative accuracy errors $E_i(n; q^{(i)}; \gamma)$, $i = 1, 2, 3$, for IP 1, IP 2 and IP 3 for $\beta = 0.3$ and $\beta = 0.7$. Note, that curves E_3 are located below curves E_i, $i = 1, 2$, for $n = 7$ when $\beta = 0.7$ (blue line) and $n = 5, 6, \ldots, 14$ when $\beta = 0.3$ (green line).

Note, that after reconstruction $q(x, y)$ we can calculate the amplitude of the tsunami wave using Airy-Green formula in case of 1D bottom profile (figure 2) [17]. In case of 2D bottom profile and linear source $q(x, y) = g(y)\delta(x)$ we can solve 2D direct problem for the amplitude $S(z, y)$

$$\begin{cases} S_z + 0.5a_1 S_y + 0.5a_2 S = 0, & z > 0, y \in (-\infty, +\infty); \\ S(0, y) = g(y), & y \in (-\infty, +\infty) \end{cases}$$

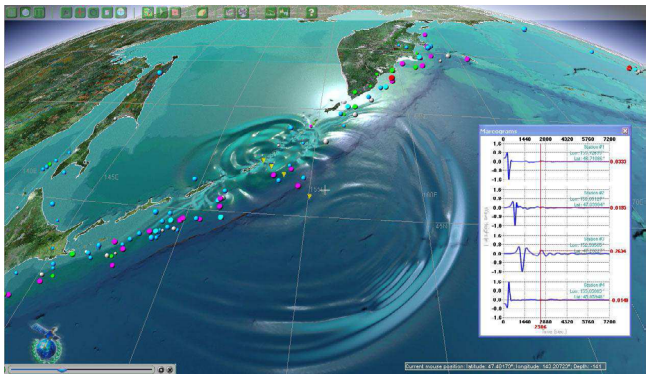


Figure 8: Numerical modelling of the Simushir tsunami 13.01.2007 using the built-in software.

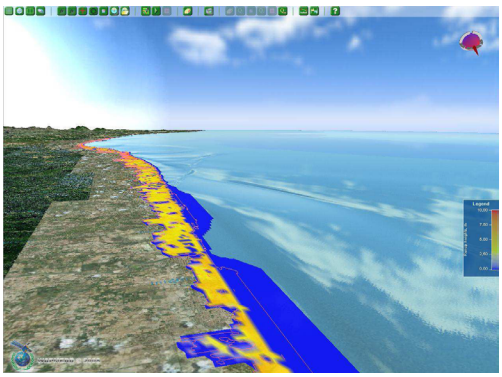


Figure 9: The flooding map of Nagapattinam, India, after tsunami run-up.

which coincides with Airy-Green formula in 1D case: $S(x) = \sqrt[4]{H(0)/H(x)}$. Here new variable $z = \tau(x, y)$ denotes the solution of eikonal equation $\tau_x^2 + \tau_y^2 = (gH(x, y))^{-1}$, a_1 and a_2 depend on $H(x, y)$ [17].

5 3D VISUALIZATION SYSTEM

Non-profit organization WAPMERR (World Agency of Planetary Monitoring and Earthquake Risk Reduction) in collaboration with GeoSystema Ltd. and ICM&MG SB RAS developed the Integrated Tsunami Research and Information System (ITRIS) to simulate tsunami waves and earthquakes, river course changes, coastal zone floods, and risk estimates for coastal constructions at wave run-ups and earthquakes [18]. The special scientific plug-in components are embedded in a specially developed GIS-type graphic shell for easy data retrieval, visualization and processing (see figure 8). A series of preliminary numerical experiments on the simplified three-dimensional models with invariable forcing is conducted based on the computational technology. These experiments give particular examples of fluid dynamics while interacting with external objects. The presented software can be used for analysis and research of various natural and man-made hazards. Figure 9 demonstrates the flooding map of artificial tsunami run-up near the Nagapattinam, India. For run-up modelling we solve nonlinear shallow water equations by the finite volume method. The main advantages of this method are using Total Variation Diminishing for stability control, high speed of calculations and adaptation for any topography of settlement area.

There are built-in catalogues and databases with set of interfaces for data managing. Fig. 10 present visualization of earthquake epicentres and tsunami locations around Japan.

The ICM&MG SB RAS develops mathematical and computational methods for Arctic region. The coupled model consisting of blocks for blocks of model of dynamics of the ocean, model of dynamics of the atmosphere and their interaction block is adapted for the computing platform of the supercomputer. A new method of observational data

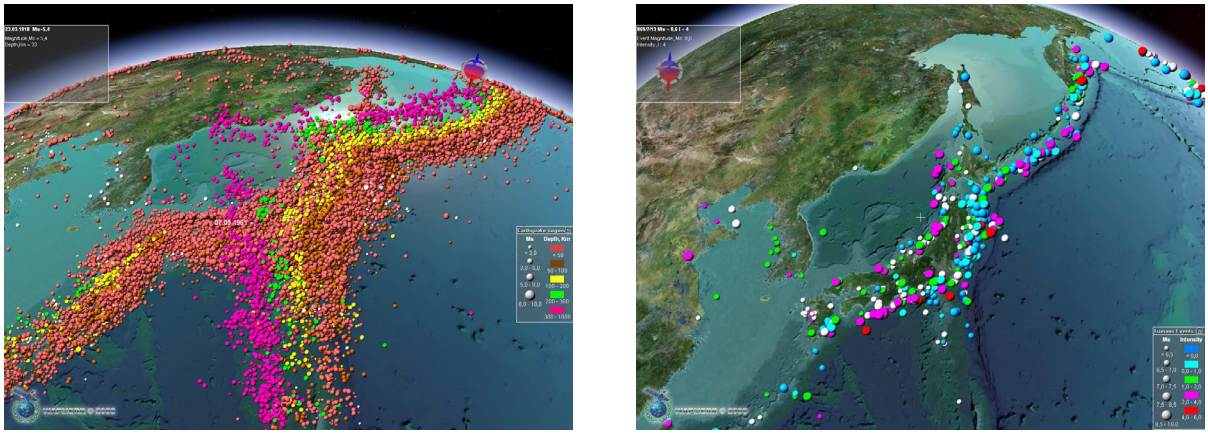


Figure 10: Visualization of the available seismic data (left) and tsunami locations (right) around Japan.

assimilation is elaborated, which is based on properties of diffusive random processes. Data on the Arctic region for the subsequent assimilation in the model are collected and programs for selection and control of these data are created.

ACKNOWLEDGMENTS

This work is supported by the Ministry of Education and Science of the Russian Federation and by the Russian Foundation for Basic Research (project No. 15-01-09230).

REFERENCES

[1] Kabanikhin, S.I. *Inverse and Ill-Posed Problems: Theory and Applications*. Berlin: de Gruyter, 2011.

[2] Kabanikhin, S., Hasanov, A., Marinin, I., Krivorotko, O. and Khidasheli, D. A variational approach to reconstruction of an initial tsunami source perturbation. *Appl. Numer. Math.* (2014) **83**:22-37.

[3] Vreugdenhil, C. V. *Numerical Methods for Shallow-Water Flow*. The Netherlands: Springer, 1994.

[4] Pelinovsky, E. *Waves in Geophysical Fluids. Tsunamis, Rogue Waves, Internal Waves and Internal Tides*, ed. J Grue and K Trulsen, Norway: SpringerWien-NewYork, 2006.

[5] Dutykh, D., Mitsotakis, D., Chubarov, L.B. and Shokin, Y.I. On the contribution of the horisontal sea-bed displacements into the tsunami generation process. *Ocean Modelling*. (2012) **56**:43-56.

[6] Voronina, T.A. and Tcheverda, V.A. Reconstruction of tsunami initial form via level oscillation. *Bull. Novosib. Comput. Cent., Ser. Math. Model. Geophys.* (1998) **4**:127-136.

[7] Voronina, T.A. Reconstruction of initial tsunami waveforms by a truncated SVD method. *J. Inverse Ill-Posed Probl.* (2011) **19**:615-629.

[8] Kabanikhin, S.I., Bektemesov, M.A., Nurseitov, D.B., Krivorotko, O.I. and Alimova, A.N. An optimization method in the Dirichlet problem for the wave equation. *J. Inverse Ill-Posed Probl.* (2012) **20**:193-211.

[9] Kabanikhin, S.I. and Krivorotko, O.I. Combined inverse tsunami problem. *Bull. Novosib. Comput. Cent., Ser. Math. Model. Geophys.* (2013) **16**:45-58.

[10] Kabanikhin, S. and Krivorotko, O. Optimization approach to combined inverse tsunami problem. *Proc. Inverse Problems – from Theory to Applications, Bristol, UK, 26-28 August, 2014*:102-107.

[11] Ivanov, V.K., Vasin, V.V. and Tanana, V.P. *Theory of Linear Ill-posed Problems and its Applications*. Boston: VSP, Vol. 36, 2002.

[12] Anikonov, Yu.E., Bubnov, B.A. and Erokhin, G.N. *Inverse and Ill-posed Sources Problems*. The Netherlands: VSP, Vol. 9, 1997.

[13] Romanov, V.G. and Moshkalev, P.S. The one-dimensional inverse problem of finding a tsunami source. *Journal of Applied and Industrial Mathematics* (2011) **14**:87-99.

[14] Voronina, T.A. and Voronin, V.V. Properties of the inverse problem operator for reconstructing the tsunami source. *Bull. Novosib. Comput. Cent., Ser. Math. Model. Geophys.* (2014) **17**:73-84.

[15] Kabanikhin, S.I. and Krivorotko, O.I. The numerical method of solving Dirichlet problem for the wave equation. *Journal of Applied and Industrial Mathematics* (2012) **15**:90-101.

[16] Hanke, M. *Conjugate Gradient Type Methods for Ill-Posed Problems*. England: Longman Scientific & Technical, 1995.

[17] Kabanikhin, S.I. and Krivorotko, O.I. A numerical method for determining the amplitude of a wave edge in shallow water approximation. *Appl. Comput. Math.* (2013) **12**:91-96.

[18] Kabanikhin, S.I., Krivorotko, O.I. and Marinin, I.V. *3D GIS for Research of Natural and Technogenic Emergency Situations*. Moscow: Palmarium Academic Publishing, 2013.

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis(Eds)

MODELLING PRESSURE CYCLE AND INTERACTION WITH REED VALVES IN A RECIPROCATING COMPRESSOR

BENOIT LEMOINE*, LOIC LE MARREC†, AVRAHAM HIRSCHBERG‡

*†Institut de Recherche Mathématique de Rennes (IRMAR)
Université de Rennes 1
Campus Beaulieu, 35042 Rennes, France
e-mail: benoit.lemoine@univ-rennes1.fr, loic.lemarrec@univ-rennes1.fr
web page: http://irmar.univ-rennes1.fr/

‡Technische Universiteit Eindhoven (TUe)
P.O. Box 513, 56000 MB Eindhoven, The Netherlands
e-mail: a.hirschberg@tue.nl
web page: http://www.tue.nl

Key words: Reciprocating Compressor, Reed Valve, Pressure behaviour, Oil Stiction, Gas Inertia

Abstract. Automatic reed valves (suction and discharge) in a reciprocating compressor are noise sources due to free vibrations and structure impacts on limiters and valve seat during a pressure cycle. Understanding the noise source generation and propagation needs a well-modelled pressure cycle in the compressor. Modelling the pressure behaviour in a cylinder requires a robust thermodynamical/mechanical model in which both physics interact. A Piao-McLinden model is proposed here to simulate the behaviour of a real gas like refrigerant. The motion of the valves is formulated as a one dimensional damped mass-spring system characterized by means of a Rayleigh method. Each model is compared to experimental data before simulating the whole coupled system: compressor gas/reed valve motion.

1 INTRODUCTION

An automotive air-conditioning (AC) reciprocating compressor consists of multiple out-of-phase pistons that compress a gas provided by a suction chamber before ejecting to a discharge chamber. To do so, there exist reed valves that automatically open and close due to material characteristics (mass, stiffness, damping) as well as external parameters necessary for better compressor yield. In some conditions, the reed valve appears to be a noise source that propagates into the whole AC system and radiates to the cabin.

To understand the sound generation of the reeds, it is aimed in this paper at modelling their motion regarding the thermodynamic cycle of refrigerant 134a that is used as an input for the dynamics model of the reed. Experimental data are used as reference for the simulations.
The thermodynamics is described by the equations of Piao [1] and McLinden [2] (section 2), delivering the main thermodynamic properties of the gas (figure 1-left). The reed valves are then modeled in terms of a one-dimensional damped mass-spring systems (figure 1-right), but accounting for the 2D-beam deflection by means of a Rayleigh method (section 3). Stiction effect due to the presence of a thin lubricating oil film between the seat and the valve plate is of major importance. A simple model allows to take this phenomenon into account. Moreover, the effect of gas flow inertia that crosses a valve during its aperture is discussed. Few details of the numerical implementation are eventually introduced in section 4 before showing results for the whole interacting model during a compressor cycle (section 5).

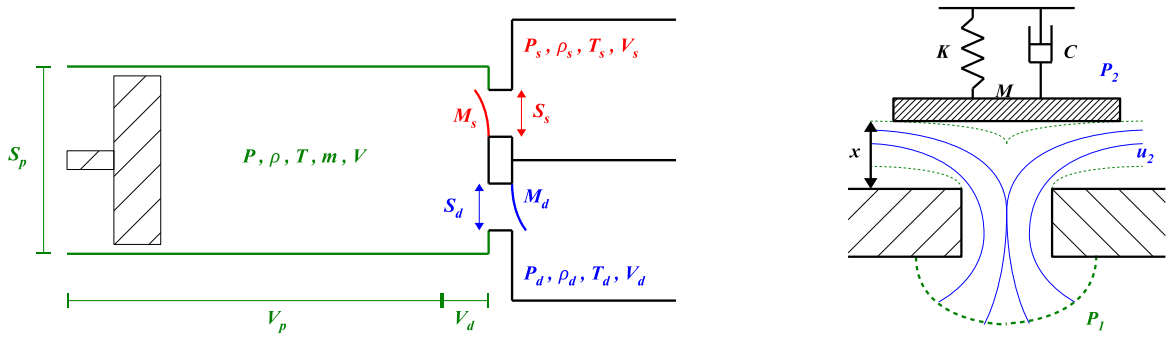


Figure 1: Left: sketch of one of the compressor cylinders with suction (s) and discharge (d) chambers (definition of the notations). Right: one-dimensional damped mass-spring system for the dynamics modelling.

2 Thermodynamical model

2.1 Piao-McLinden model

The thermodynamical model of the refrigerant R-134a is based on the equation of state of Piao *et al.* [1] and the heat capacity formulation of McLinden *et al.* [2]. Basically, it deals with the calculation of the Helmholtz function according to the equation of state of the pressure in vapor phase as well as the heat capacity at constant pressure, in order to determine all thermodynamic properties of the gas.
The equation of state for the pressure P is defined as a density power series that reads:

$$P(\rho,T) = P^* \left[\frac{T_r \rho_r}{Z^*} + \frac{b}{T_r^5}\rho_r^2 + \sum_{i=1}^{8} \rho_r^{i+1} \sum_{j=1}^{4} \frac{a_{ij}}{T_r^{j-1}} \right] \tag{1}$$

where coefficients a_{ij} and b are given in table 1 in appendix, superscript * means *critical value*, $\rho_r = \rho/\rho^*$ and $T_r = T/T^*$ are *reduced* density and temperature respectively. $Z^* = P^*/(r\rho^* T^*)$ is the compressibility factor of the gas in which r is the gas specific constant.
The specific Helmholtz function $f(\rho, T)$ is then calculated since $df = P/(\rho^* \rho_r^2) d\rho_r$, using Eq.1 and the formulation for the heat capacity at constant pressure proposed by McLinden *et al.* [2].
The first law of thermodynamics is then stated for a single compressor cylinder of volume V and the variation of temperature w.r.t. time is obtained:

$$\frac{dT}{dt} = \frac{1/V}{\rho\left(\frac{\partial h}{\partial T}\right)_\rho - \left(\frac{\partial P}{\partial T}\right)_\rho} \left\{ \frac{dQ}{dt} + \Phi - h\frac{dm}{dt} - V\frac{d\rho}{dt}\left[\rho\left(\frac{\partial h}{\partial \rho}\right)_T - \left(\frac{\partial P}{\partial \rho}\right)_T\right]\right\} \tag{2}$$

where $h(\rho, T)$ is the specific enthalpy of the gas in the compressor cylinder, dQ/dt is the heat transfer by conduction through the cylinder walls, m is the mass of gas in the cylinder volume V, and $\Phi = q_s h_s + q_d h_d$ is the total heat flux crossing the suction and discharge valves determined from the mass flows q_s and q_d and the specific enthalpy of each chamber.

2.2 Validation

A validation of the thermodynamical model alone has been conducted by considering the compression stage in the whole compressor cycle. In this way, Eq.2 is simplified since m becomes constant and $\Phi = 0$ (neither entering nor ejecting flows). Initial values for the thermodynamic variables (P, ρ, T, h) are taken as in suction phase (P_s, ρ_s, T_s, h_s). The calculated pressure prediction is eventually compared to measurements in figure 2, for two compressor rotation speeds in RPM (*Rotation Per Minute*). Results in pressure, for given cylinder volume $V(t)$ as function of the time t, show very good agreement between model and experiments. Similar results are found for expansion stage if initial conditions match with discharge characteristics.

3 Valve motion

Since both suction and discharge valves have a similar behaviour, only the modelling of discharge valve is introduced here. Equivalent formulations for suction valve are then implemented in the whole model simulations.
A valve consists of a reed that automatically opens and closes during a pressure cycle due to its mechanical characteristics (mass, stiffness, intrinsic damping) on one hand, some external paramaters like the presence of structural boundaries (up and down limiters) or lubricating oil on the other hand. The reed characteristics are accounted for by means of a Rayleigh method — equality of kinetic energies with effective and real mass — to be implemented into a simplified one-dimensional damped mass-spring system as shown in figure 1, where x represents the displacement of the valve plate.

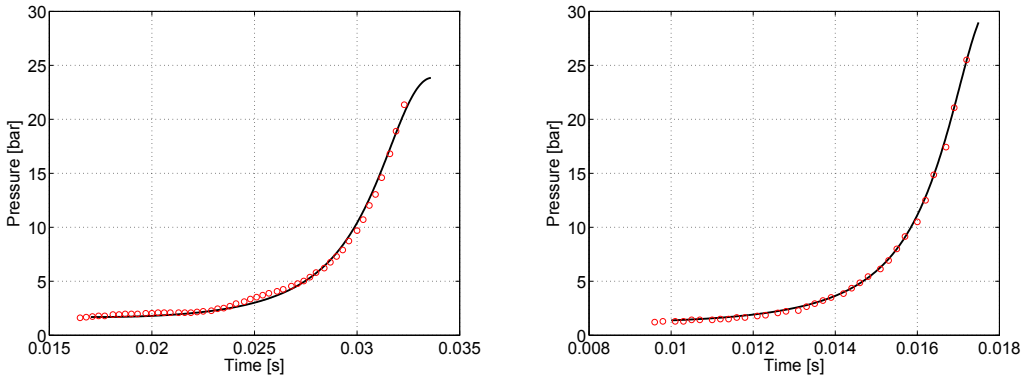


Figure 2: Pressure behaviour during compression stage at 1700 RPM (left) and 3200 RPM (right). Comparison between model prediction (line) and experiments (symbols).

3.1 Formulation

The dynamics equation of the one-dimensional system satisfied by displacement x reads:

$$\ddot{x} + \frac{\omega_0}{Q_0}\dot{x} + \omega_0^2 x = \frac{F_p}{M} + \frac{F_{stic}}{M} \tag{3}$$

where $\omega_0 = \sqrt{K/M}$ is the angular frequency with $K \approx 8 \cdot 10^3\ N.m^{-1}$ the spring stiffness and $M \approx 1.5 \cdot 10^{-4}\ kg$ the effective mass — in terms of the Rayleigh method — in order to consider the deflection of a real beam, Q_0 is a quality factor related to the damping coefficient $C = \omega_0 M/Q_0$, and F_p and F_{stic} are the pressure and stiction forces respectively. Besides, solving this equation suggests some complications due to the presence of structural obstacles and lubricating oil in a real compressor.

3.2 Structural boundaries

A first correction on displacement is necessary because of structural boundaries such as the valve seat and the limiter of height x_L. Since the resolution of the whole system {Thermodynamics+Mechanics} is performed step-by-step, this is simply done by imposing the opposite value of the valve plate velocity at previous step, corrected by a coefficient of restitution $0 < \eta < 1$ representing the absorbed momentum when the valve impacts a boundary. η is commonly defined as the square root of the before/after impact kinetic energy ratio. In other words, for calculated displacement x_i at time-step i:

$$\text{if} \begin{cases} x_i < 0 & \Rightarrow \quad x_i = 0 \quad , \quad \dot{x}_i = -\eta \dot{x}_{i-1} \\ x_i > x_L & \Rightarrow \quad x_i = x_L \quad , \quad \dot{x}_i = -\eta \dot{x}_{i-1} \end{cases}$$

According to Cross [3], it is taken here a coefficient of restitution for steel/steel impact of $\eta = 0.85$.

3.3 Gas inertia

The inertia of the fluid that crosses a valve during its aperture generates a pressure force which aims at keeping the valve opened instead of closing quickly as it is shown in figure 4 after the first maximum of the curves, whereas it is smoother in the experiments. A simple model for the correction term on pressure is performed here, from the unsteady Bernoulli's equation expressed here from the cylinder (point 1) to the discharge chamber (point 2) as in figure 1-right:

$$\rho_2 \frac{d\varphi}{dt} + \rho_2 \frac{u_2^2}{2} + P_2 = P_1 \tag{4}$$

with the potential $\varphi = \int_1^2 u\,ds \approx L_{eff}\,u_2$ in which u_2 is the gas jet speed and $L_{eff} \approx 1.2{\cdot}10^{-2}\ m$ is an effective length that is as the same order as the sum of the valve seat orifice and thickness. u_1, the gas flow speed at point 1 is neglected because the cylinder cross-section S_1 is much higher than the jet cross-section S_2. Thus, the correction corresponds to the time-derivative $\rho_2 L_{eff} du_2/dt$. A similar correction can be also applied at the suction valve but it has not been implemented here since the inertia correction is directly related to pressure difference between both chambers and the suction valve involves a much lower difference.

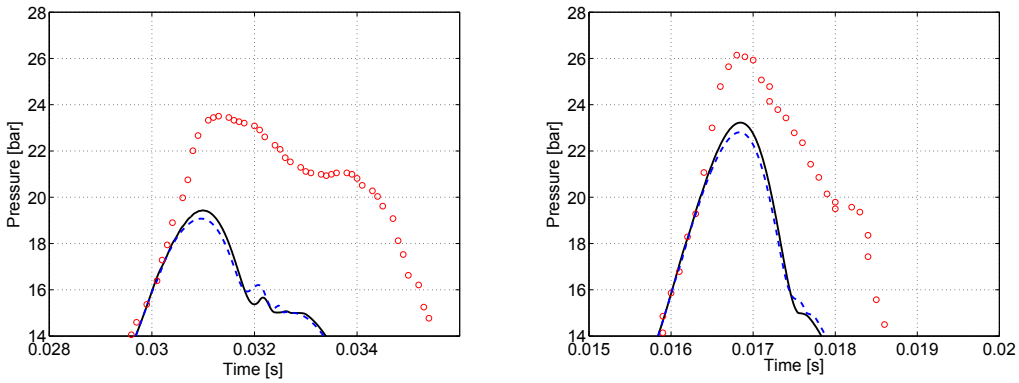


Figure 3: Inertia effect at 1700 RPM (left) and 3200 RPM (right). Experiments (symbols) *vs* model without inertia (—) and with inertia (−−)

Figure 3 shows that inertia effect is not so important whatever the compressor rotation speed. It apperas the maximum of pressure reached at the opening of the valve is slightly reduced and furthermore the decreasing slope during the valve aperture is attenuated. The time the valve closes is *a priori* not influenced, justified since the pressure difference is supposed to be equal or close to zero upon closing.

3.4 Stiction effect

Stiction appears due to the presence of a thin oil film between the valve seat and the valve plate. This phenomenon is shown to have an important effect [4, 5, 6, 7] on the opening of the valve by generating by capillarity a low pressure area in the oil which keeps the valve closed. This induces an overpressure in the cylinder required to lift the valve plate and consequently a delay of the opening time.
Here, stiction is modeled as a constant pressure force applied on a ring surface of 1 mm wide around the valve orifice and keeping the valve closed. While the calculated displacement x (Eq.3) is smaller than a hundredth of the limiter position x_L — same order as the oil film thickness of Khalifa and Liu [5] — then the stiction force is applied. Otherwise, when x becomes higher, the stiction force vanishes. Such a simple model seems to be sufficient to simulate the effect of stiction observed in the measurements (figure 4). It can be noted that the oil stiction force seems to increase with the rotation speed of the compressor which agrees with Bauer's observation [4]. Moreover, the maximum amplitude of pressure is much higher than the experiments at the opening which is counterbalanced by an important deacrease, but good agreement is observed at the closing.

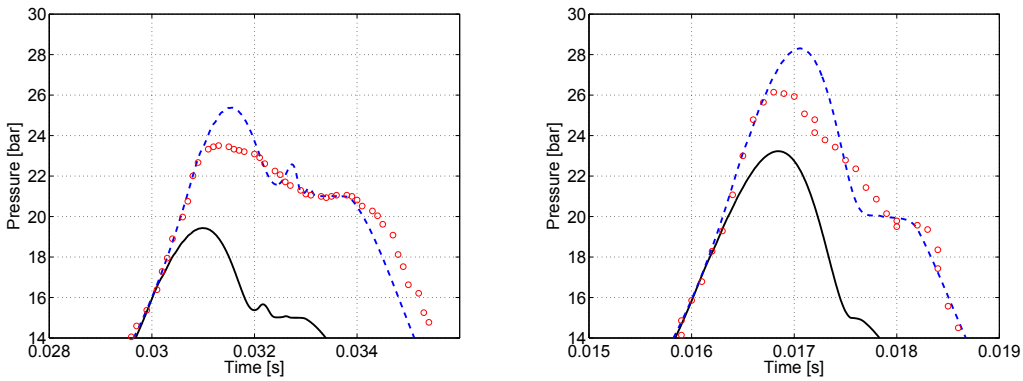


Figure 4: Stiction effect (no inertia) at 1700 RPM (left) and 3200 RPM (right). Experiments (symbols) *vs* model without stiction (—) and model with stiction (−−)

Of course, improvements can be brought by modelling this force by means of the Reynolds lubrication theory [8], as done by Pizarro *et al.* [6] who splitted the stiction force up as a combination of a viscous effect force (oil deformation), a capillary force (meniscus curvature) and an interfacial tension force. The main difference with the paper of Khalifa and Liu [5] is the latter consider an infinite oil volume whereas Pizarro *et al.* assume a finite oil volume that must be estimated.

4 Numerical implementation

Simulations have been conducted using the MATLAB R2013a calculation software, in a step-by-step resolution because of the couplings between equation variables. First of all,

the temperature variation is calculated from Eq. 2 in order to obtain the thermodynamic properties (T, ρ, P, m). Then the flow mass q is evaluated to be used as a mechanical model input. Displacement and velocity of the valve plate are then simulated according to all aforementioned conditions (stiction, inertia, limiters). The mass flow q is finally corrected so as to account for the real height x of the valve plate for the next step.
To solve the dynamics equation of valve motion (Eq.3), a classical 4^{th}-order Runge-Kutta scheme solver (RK4) is used at each time-step. Both suction and discharge valve motions are calculated independently. Time origin is taken when the cylinder volume is equal to the dead volume, so that thermodynamic properties are equal to the discharge values and the discharge valve plate is supposed opened (suction valve closed).
Due to analytical formulations, the whole calculation performs within 2 seconds for about 5000 time-steps per pressure cycle. The convergence is observed by the second cycle, because of discrepancies between assumed initial conditions and reality (especially the unknown exact position of discharge valve plate).

5 General results for a pressure cycle

Figure 5 shows the importance to model the aforementioned effects by comparing the simplest model with only structural boundaries accounted for (neither oil stiction nor gas inertia forces) and the model with all of the three conditions: limiters+stiction+inertia. The experimental data are plotted here as the reference for the models.

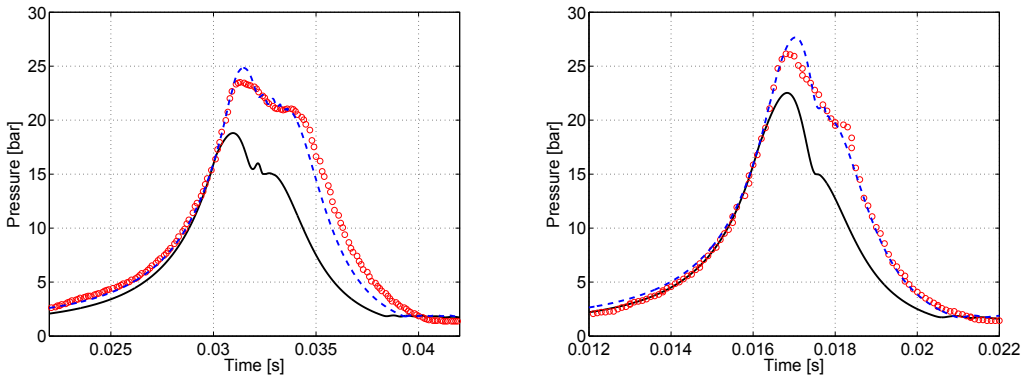


Figure 5: Comparisons between simulations (lines) and experiments (symbols) at 1700 RPM (left) and 3200 RPM (right) for a whole pressure cycle. (—) simplest model without limiters restitution, stiction nor inertia. (−−) model with limiters restitution, oil stiction and gas inertia.

Oil stiction appears to have the most significant impact on the pressure behaviour in the cylinder since it imposes a high increase by sticking the valve plate onto the seat. The inertia of the gas creates a pressure drop when it is at its maximum in the compressor cycle. This leads to a smoother slope in decreasing the pressure till the closing of the valve.

However, as observed in the previous sections, the maximum pressure is overestimated in the simulations when the valve opens, and the valve is predicted to close before the measurement while, especially at lowest compressor rotation speeds. This can be explained as the oil stiction is taken into account only for the opening but it also plays a role when the valve closes.

6 Conclusion

In automotive reciprocating compressor, all suction and discharge valves open and close successively during a pressure cycle. This generates acoustic interactions (amplification or inhibition) that appear to have an effect on the noise that propagates through the air-conditioning system and radiates to the cabin. Thus, modelling pressure and valve motion behaviours allow to discriminate the noise sources produced by the impacts of the reeds onto the limiters and the seat as well as by the vibrations of the valve plates (eigenfrequencies).
In this paper, a coupled system involving a thermodynamical model and a mechanical model has been implemented. Simulations give good agreements with experimental data by accounting for oil stiction and gas inertia phenomena. However, the latter is shown not to be of high importance compared to the other. The stiction model can also be improved according to recent studies [5, 6].
Moreover, only stiction between the valve plate and the seat at the opening has been considered here, but there likely exists oil stiction between the valve plate and the limiter. Besides, stiction should play a role at the closing, especially by inducing an accelerated decrease when oil film makes contact with the plate (oil suction) and a later opening after the plate bounded on the seat. Then, parametric sensitivity investigations must be done so as to optimize the involved variables. Since the mechanical model is based on a one-dimensional damped mass-spring system, considering a realistic two dimensional non-homogeneous beam should actually improve the predictions. Finally, the acoustic interactions (resonances) between all out-of-phase valve cycles in the suction and discharge chambers would be studied and compared to the measurements.

Acknowledgements

This work has been supported by Région Bretagne and Sanden Manufacturing Europe (SME). The authors are very grateful to the SME team for the fruitful discussions and experimental data they provided.

REFERENCES

[1] Piao, C.C., Noguchi, M., Sato, H., Watanabe, K. An improved equation of state for R-134a. *ASHRAE Trans., JSME* (1994):358–366.

[2] McLinden, M.O., Gallagher, J.S., Weber, L.A., Morrison, G., Ward, D., Goodwin,

A.R.H., Moldover, M.R., Schmidt, J.W., Chae, H.B., Bruno, T.J., Ely, J.F., Huber, M.L. Measurement and formulation of the thermodynamic properties of refrigerants 134a (1,1,1,2-tetrafluoroethane) and 123 (1,1-dichloro-2,2,2-trifluoroethane). *ASHRAE Trans.* (1989) **95**:263–283.

[3] Cross, R. The bounce of a ball. *Am. J. Phys.* (1999) **67**:222–227.

[4] Bauer, F. The influence of liquids on compressor valves. *Proc. Int. Comp. Conf. at Purdue* (1990) **2**:647–653.

[5] Khalifa, H.E., Liu, X. Analysis of stiction effect on the dynamics of compressor suction valve. *Proc. Int. Comp. Conf. at Purdue* (1998) **1**:87–92.

[6] Pizarro-Recabarren, R.A., Barbosa, J.R., Deschamps, C.J. Modeling the stiction effect in automatic compressor valves. *Int. J. Refrig.* (2013) **36**:1916–1924.

[7] Resch, M., Scheidl, R. A model for fluid stiction of quickly separating circular plates. *Proc. Inst. Mech. Eng., Part C: J. Mech. Engng Sc.* (2013) **228**:1540–1556.

[8] Booker, J.F. Squeeze films and bearing dynamics. In: *Booser, E.R. (Ed.) CRC Handbook of Lubrication (Theory and Practice of Tribology), Theory and Design*, CRC Press, Boca Raton, Vol. II., (1983):121–137.

Appendix

Table 1: Values of coefficients a_{ij} and b in the equation of state of Piao [1]

	a_{i1}	a_{i2}	a_{i3}	a_{i4}
a_{1j}	10.200664	-33.117138	27.006031	-9.3442006
a_{2j}	-32.533383	84.321026	-53.638675	5.1293252
a_{3j}	78.123563	-186.15733	105.08815	0
a_{4j}	-68.597038	163.7985	-89.655243	0
a_{5j}	28.411548	-68.88115	34.766061	0
a_{6j}	-4.9177332	13.40429	-5.7021512	0
a_{7j}	0	-0.61139649	0	0
a_{8j}	0.057702463	-0.077630849	0.069571809	0
b	0.043959113			

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis (Eds)

MULTIPOINT GLOBAL OPTIMAL SHAPE DESIGN BY MORPHING

A. NASTASE*

* Aerodynamics of Flight, RWTH, Aachen University
Templergraben 55, 52062 Aachen, Germany
e-mail: nastase@lafaero.rwth-aachen.de

Key words: Multipoint Global Optimal Design, Extended Variational Problems with Free Boundaries, Hybrid Solutions for Navier-Stokes PDEs, Meshless Solutions, Movable Leading Edge Flaps.

Abstract. *The determination of the global optimized (GO) shape of a flying configuration (FC), which is of minimum drag at two cruising Mach numbers, can be obtained by morphing. Two movable leading edge flaps are used for this purpose. They are retracted at higher Mach number and stretched by the lower one. Two consecutive enlarged variational problems with free boundaries occur and are solved by using the own iterative optimum-optimorum theory as optimization strategy. It uses only in its first step of iteration hyperbolic potential solutions and determine the inviscid GO shape of FC as surrogate model. Up the second step of iteration own hybrid solutions for Navier-Stokes layer are used and the total drag is the new functional.*

1 INTRODUCTION

The aim of this study is to determine the aerodynamic, global optimized (GO) shape of a flying configuration (FC), which is flying at two supersonic cruising Mach numbers. Due to the fact that the GO shape of a FC, with respect to minimum drag at cruise, changes its optimal planform very much, if the supersonic cruising Mach number is varied, this aim can be correct realized only by morphing. The morphing is obtained by using movable leading edge flaps. The surfaces of the wing, of the fuselage and of the flaps are supposed to be expressed or approximated in form of different superpositions of homogeneous polynoms in two variables with free coefficients. These coefficients, together with the similarity parameters of the planforms of these components of FC, are the free parameters of optimization. The determination of the GO shape of FC at two supersonic cruising Mach numbers leads to the solving of two consecutive extended variational problems with free boundaries. The first one consists in the determination of the GO shape of the unmovable integrated wing-fuselage part of FC with retracted flaps, which is of minimum drag at higher cruising Mach number. The second one consists in the determination of the shape of the stretched leading edge flaps, in such a manner that the entire FC with stretched flaps is of minimum drag, at the second lower cruising Mach number.

An own developed mathematical strategy called iterative optimum-optimorum (OO) theory is used for the solving of these both consecutive variational problems. In the first step

of the iterative OO theory, the surrogate GO shape is determined by using the three-dimensional hyperbolic potential solutions of the author as start solutions, as in [1] and the OO theory as strategy of optimization of the inviscid drag functional. It follows a computational checking of the inviscid GO shape of the surrogate model by using hybrid solutions for the Navier-Stokes layer (NSL), proposed by the author, as in [1,2] . The friction drag coefficient is computed and a weak interaction aerodynamics-structure can be performed. Up the second step of iteration the new drag functional is the total drag and additional constraints due to the structure requests can occur. The following premises for the optimization are taken into consideration:
- the FCs have sharp leading edges in order to avoid the bow shock and to fly with characteristic surface;
- the junction lines wing-fuselage and wing-flaps begin at the apex of the wing, in order to avoid the sonic boom interference;
- the surface of FC must be integrated (namely the wing and the fuselage have the same tangent plane along the junction lines between the wing and the fuselage) in order to avoid the negative effects of corners;
- the leading edges conturnements of their subsonic leading edges of wing with flaps in retracted position and of subsonic flaps in stretched position must be reduced. For these purposes the Kutta condition on leading edges is used.

2 THREE-DIMENSIONAL HYPERBOLIC START SOLUTIONS FOR INVISCID GLOBAL OPTIMAL SHAPE DESIGN OF FLYING CONFIGURATION

Let us firstly consider an integrated wing-fuselage FC with arbitrary camber, twist and thickness distributions, which is flying at the higher cruising Mach number M_∞ with the flaps in retracted position. The FC with retracted flaps is considered like an integrated wing-fuselage FC with arbitrary camber, twist and thickness distributions and is flying at the higher cruising Mach number M_∞. Dimensionless coordinates are used for the computation of the distributions of velocity's components:

$$\widetilde{x}_1 = \frac{x_1}{h_1} \quad , \quad \widetilde{x}_2 = \frac{x_2}{\ell_1} \quad , \quad \widetilde{x}_3 = \frac{x_3}{h_1} \; . \qquad (\; \widetilde{y} = \frac{\widetilde{x}_2}{\widetilde{x}_1}) \tag{1}$$

Further, integrated wing-fuselage FCs are considered, namely, for which the mean surface is continuous and the thickness distributions on the wing and on the fuselage are different and three-dimensional hyperbolic potential solutions, previously given by the author as in [1], are used as start solutions for the determination of the inviscid GO shape of FC with retracted flaps, at higher cruising Mach number M_∞ and also for the determination of the inviscid GO shape of the stretched flaps of the FC, at lower cruising Mach number M'_∞ .

The downwashes on the thin and thick-symmetrical components of the integrated FCs, with flaps in retracted position w and w^*, w'^* (on the wing and on the fuselage of FCs), are expressed in form of superposition of homogeneous polynoms with arbitrary coefficients:

$$w \equiv \widetilde{w} = \sum_{m=1}^{N} \widetilde{x}_1^{m-1} \sum_{k=0}^{m-1} \widetilde{w}_{m-k-1,k} \left|\widetilde{y}\right|^k \quad ,$$

$$w^* \equiv \widetilde{w}^* = \sum_{m=1}^{N} \widetilde{x}_1^{m-1} \sum_{k=0}^{m-1} \widetilde{w}^*_{m-k-1,k} \left|\widetilde{y}\right|^k \quad , \qquad w'^* \equiv \overline{w}^* = \sum_{m=1}^{N} \widetilde{x}_1^{m-1} \sum_{k=0}^{m-1} \overline{w}^*_{m-k-1,k} \left|\widetilde{y}\right|^k \quad . \qquad \text{(2a-c)}$$

The coefficients of the downwashes and the similarity parameter $\nu = B\ \ell \quad \overline{\nu} = Bc$ ($\ell = \ell_1 / h_1$, $B = \sqrt{M_\infty^2 - 1}$) , of the planform of the wing are the free parameters of the optimization and $\ell,\ \ell_1,\ h_1$ are the dimensionless span, the half-span and the depth of the planform of the delta wing. The quotient $\bar{k} \equiv \overline{\nu} / \nu = c / \ell$ of the similarity parameters of the wing and of the fuselage, which depend on the purpose of the FC, is supposed constant. If the principle of minimal singularities (which fulfill the jumps of the velocity's components) and the hydrodynamic analogy of Carafoli are used, the following expressions for the axial disturbances of the thin and of the thick-symmetrical components of the integrated wing-fuselage FC with subsonic leading edges are obtained, as in [1]:

$$u \equiv \ell\ \widetilde{u} = \ \ell\ \sum_{n=1}^{N} \widetilde{x}_1^{n-1} \left\{ \sum_{q=0}^{E\left(\frac{n}{2}\right)} \frac{\widetilde{A}_{n,2q} \widetilde{y}^{2q}}{\sqrt{1-\widetilde{y}^2}} + \sum_{q=1}^{E\left(\frac{n-1}{2}\right)} \widetilde{C}_{n,2q} \widetilde{y}^{2q} \cosh^{-1} \sqrt{\frac{1}{\widetilde{y}^2}} \right\} \quad , \qquad \text{(3a)}$$

$$u^* \equiv \ell\ \widetilde{u}^* = \ell\ \sum_{n=1}^{N} \widetilde{x}_1^{n-1} \left\{ \sum_{q=0}^{n-1} \widetilde{H}^*_{nq} \widetilde{y}^q \left(\cosh^{-1} M_1 + (-1)^q \cosh^{-1} M_2 \right) \right.$$

$$+ \sum_{q=0}^{n-1} \widetilde{G}^*_{nq}\ \widetilde{y}^q \left(\cosh^{-1} S_1 + (-1)^q \cosh^{-1} S_2 \right)$$

$$\left. + \sum_{q=1}^{E\left(\frac{n-1}{2}\right)} \widetilde{C}^*_{n,2q}\ \widetilde{y}^{2q} \cosh^{-1} \sqrt{\frac{1}{\nu^2 \widetilde{y}^2}} + \sum_{q=0}^{E\left(\frac{n-2}{2}\right)} \widetilde{D}^*_{n,2q}\ \widetilde{y}^{2q} \sqrt{1-\nu^2 \widetilde{y}^2} \right\} \quad . \qquad \text{(3b)}$$

$$\left(M_{1,2} = \sqrt{\frac{(1+\nu)(1 \mp \nu\ \widetilde{y})}{2\nu(1 \mp \widetilde{y})}} \right)$$

The lift and the pitching moment coefficients of the integrated wing-fuselage FCs with stretched flaps, computed with the hyperbolic potential theory, are the following:

$$C_\ell \equiv 8\ell \int_{\widetilde{O}\widetilde{A}_1\widetilde{C}_1} \widetilde{u}\ \widetilde{x}_1 d\widetilde{x}_1 d\widetilde{y} \quad , \qquad C_m \equiv -8\ell \int_{\widetilde{O}\widetilde{A}_1\widetilde{C}_1} \widetilde{u}\ \widetilde{x}_1^2 d\widetilde{x}_1 d\widetilde{y} \quad . \qquad \text{(4a,b)}$$

The inviscid drag coefficients of the thin, thick-symmetrical and thick, lifting FCs with retracted flaps are quadratical forms with respect to the downwashes coefficients :

$$C_d \equiv \ell \widetilde{C}_d = 8\ell \int\limits_{\widetilde{O}\widetilde{A}_1\widetilde{C}} \widetilde{u}\ \widetilde{w}\ \widetilde{x}_1 d\widetilde{x}_1 d\widetilde{y}\ ,$$

$$C_d^* \equiv \ell \widetilde{C}_d^* = 8\ell \left[\int\limits_{\widetilde{O}\widetilde{C}_1\widetilde{C}} \widetilde{u}^* \overline{w}^* \widetilde{x}_1 d\widetilde{x}_1 d\widetilde{y} + \int\limits_{\widetilde{O}\widetilde{A}_1\widetilde{C}_1} \widetilde{u}^* \widetilde{w}^* \widetilde{x}_1 d\widetilde{x}_1 d\widetilde{y} \right],$$

$$C_d^{(i)} \equiv \ell \widetilde{C}_d^{(i)} = \ell \left(\widetilde{C}_d + \widetilde{C}_d^* \right). \qquad \text{(5a-c)}$$

Let us now consider the FCs with stretched flaps. The downwashes w and w^*, $\overline{w}^*$ on the thin and thick-symmetrical components of the integrated FC remain unchanged because the surface of the wing part of FC remains unchanged. The downwashes w'' and w''^* on the thin and thick-symmetrical stretched flaps are also supposed to be approximated in form of superposition of homogeneous polynoms in two variables with free coefficients, namely:

$$w'' \equiv \widetilde{w}' = \sum_{m=1}^{N} \widetilde{x}_1^{m-1} \sum_{k=0}^{m-1} \widetilde{w}'_{m-k-1,k} \left|\widetilde{y}\right|^k \quad , \quad w''^* \equiv \widetilde{w}'^* = \sum_{m=1}^{N} \widetilde{x}_1^{m-1} \sum_{k=0}^{m-1} \widetilde{w}'^*_{m-k-1,k} \left|\widetilde{y}\right|^k \quad . \qquad \text{(6a,b)}$$

The corresponding axial disturbance velocities of the thin and thick-symmetrical components of the integrated FCs with stretched flaps at the second lower cruising Mach number M'_∞ are, as follows:

$$u = L \sum_{n=1}^{N} \widetilde{x}_1^{n-1} \left[\sum_{q=0}^{n-1} \widetilde{\widetilde{A}}_{nq} \widetilde{y}^q \left(\cosh^{-1} N_1'' + (-1)^q \cosh^{-1} N_2'' \right) + \sum_{q=0}^{E\left(\frac{n}{2}\right)} \frac{\widetilde{\widetilde{F}}_{n,2q}\ \widetilde{y}^{2q}}{\sqrt{1-\widetilde{y}^2}} \right.$$

$$\left. + \sum_{q=1}^{E\left(\frac{n-1}{2}\right)} \widetilde{\widetilde{C}}_{n,2q} \widetilde{y}^{2q} \cosh^{-1} \sqrt{\frac{1}{\widetilde{y}^2}} \right] . \qquad \text{(7a)}$$

$$\left(\quad N'_{1,2} = \sqrt{\frac{(1+k')\ (1 \mp \widetilde{y})}{2\ (k' \mp \widetilde{y})}} \quad \right)$$

$$u^* = L \sum_{n=1}^{N} \widetilde{x}_1^{n-1} \Bigg[\sum_{q=0}^{n-1} \widetilde{\widetilde{H}}^*_{nq}\, \widetilde{y}^q \left(\cosh^{-1} M_1' + (-1)^q \cosh^{-1} M_2' \right)$$

$$+ \sum_{q=0}^{E\left(\frac{n}{2}\right)} \widetilde{\widetilde{D}}^*_{n,2q}\, \widetilde{y}^{2q} \sqrt{1-\widetilde{\nu}'^2\, \widetilde{y}^2} + \sum_{q=1}^{E\left(\frac{n-1}{2}\right)} \widetilde{\widetilde{C}}^*_{n,2q}\, \widetilde{y}^{2q} \cosh^{-1} \sqrt{\frac{1}{\widetilde{\nu}'^2\, \widetilde{y}^2}}$$

$$+ \sum_{q=0}^{n-1} \widetilde{\widetilde{F}}^*_{nq}\, \widetilde{y}^q \left(\cosh^{-1} N_1'' + (-1)^q \cosh^{-1} N_2'' \right)$$

$$+ \sum_{q=0}^{n-1} \widetilde{\widetilde{G}}^*_{nq}\, \widetilde{y}^q \left(\cosh^{-1} S_1' + (-1)^q \cosh^{-1} S_2' \right) \Bigg] \quad . \qquad (7b)$$

$$\left(\; M'_{1,2} = \sqrt{\frac{(1+\nu')\ (1 \mp \widetilde{\nu}'\widetilde{y})}{2(\nu' \mp \widetilde{\nu}'\widetilde{y})}} \; , \qquad N''_{1,2} = \sqrt{\frac{\left(1+\widetilde{\nu}'\right)\left(1 \mp \widetilde{\nu}'\widetilde{y}\right)}{2\nu'\left(1 \mp \widetilde{y}\right)}} \; , \qquad S'_{1,2} = \sqrt{\frac{\left(1+\overline{\nu}'\right)\left(1 \mp \widetilde{\nu}'\widetilde{y}\right)}{2(\overline{\nu}' \mp \widetilde{\nu}'\widetilde{y})}} \; \right)$$

Hereby are: L, $\widetilde{\nu}' = B' L$, $\nu' = B' \ell$, $\overline{\nu}' = B' c$, the dimensionless span and the similarity parameters of the FC with stretched flaps, of its wing and of its integrated fuselage, at lower cruising Mach number M'_∞, the quotients $\widetilde{k} = \ell / L$, $\overline{k} = c / L$, $B' = \sqrt{M'^2_\infty - 1}$ and here $\widetilde{y} = y / L$. The lift and the pitching moment coefficients of the integrated wing-fuselage FC with stretched flaps are linear and homogeneous functions with respect to the coefficients of the downwashes:

$$C_\ell \equiv 8\ell \int_{\widetilde{O}\widetilde{C}\widetilde{F}_1} \widetilde{u}\ \widetilde{x}_1\, d\widetilde{x}_1\, d\widetilde{y} = \ell \sum_{n=1}^{N} \sum_{j=0}^{n-1} \left(\widetilde{\Lambda}_{nj} \widetilde{w}_{n-j-1,j} + \widetilde{\widetilde{\Lambda}}_{nj} \widetilde{\widetilde{w}}_{n-j-1,j} \right) \quad ,$$

$$C_m \equiv 8\ell \int_{\widetilde{O}\widetilde{C}\widetilde{F}_1} \widetilde{u}\ \widetilde{x}_1^2\, d\widetilde{x}_1\, d\widetilde{y} = \ell \sum_{n=1}^{N} \sum_{j=0}^{n-1} \left(\widetilde{\Gamma}_{nj} \widetilde{w}_{n-j-1,j} + \widetilde{\widetilde{\Gamma}}_{nj} \widetilde{\widetilde{w}}_{n-j-1,j} \right) \quad . \qquad (8a,b)$$

The inviscid drag coefficients of the thin, thick-symmetrical and thick, lifting FC are:

$$C_d \equiv \ell \widetilde{C}_d = 8\ell \left[\int_{\widetilde{O}\widetilde{C}_1\widetilde{A}_1} \widetilde{u}\widetilde{w}\widetilde{x}_1\, d\widetilde{x}_1\, d\widetilde{y} + \int_{\widetilde{O}\widetilde{A}_1\widetilde{F}_1} \widetilde{u}\widetilde{\widetilde{w}}\widetilde{x}_1\, d\widetilde{x}_1\, d\widetilde{y} \right]$$

$$C_d^* \equiv \ell \widetilde{C}_d^* = 8\ell \left[\int_{\widetilde{O}\widetilde{C}_1\widetilde{A}_1} \widetilde{u}^*\, \widetilde{w}^*\, \widetilde{x}_1\, d\widetilde{x}_1\, d\widetilde{y} + \int_{\widetilde{O}\widetilde{C}\widetilde{C}_1} \widetilde{u}^*\, \overline{w}^*\, \widetilde{x}_1\, d\widetilde{x}_1\, d\widetilde{y} + \int_{\widetilde{O}\widetilde{A}_1\widetilde{F}_1} \widetilde{u}^*\, \widetilde{\widetilde{w}}^*\, \widetilde{x}_1\, d\widetilde{x}_1\, d\widetilde{y} \right]$$

$$C_d^{(i)} \equiv \ell \widetilde{C}_d^{(i)} = \ell \left(\widetilde{C}_d + \widetilde{C}_d^* \right). \qquad (9a\text{-}c)$$

The optimal planform of the GO shape of FC presents an important change with respect to the chosen cruising Mach number . That is the reason that the multipoint design is realized by morphing.

3 INVISCID GLOBAL OPTIMAL SHAPE DESIGN OF FLYING CONFIGURATION WITH RETRACTED AND WITH STRETCHED FLAPS

The determination of the inviscid GO shape of the FC with movable leading edge flaps leads to two consecutive enlarged variational problems with free boundaries .

The first one concerns the determination of the GO shape of the integrated wing-fuselage configuration with retracted flaps, which is of minimum drag at higher cruising Mach number M_∞. The free parameters of the optimization are the coefficients of the downwashes w , w^* and w'^* and also the similarity parameters ν and $\bar{\nu}$ of the planforms of the wing and of the fuselage. The quotient of these similarity parameters, which depends on the purpose of the FC, is supposed to be constant.

The constraints of the inviscid GO shape's design are the following: the given lift, pitching moment and the Kutta condition along the subsonic leading edges of the thin FC component (in order to cancel the induced drag at cruise and to suppress the transversal conturnement of the flow around the subsonic leading edges, in order to increase the lift) and the given relative volumes of the wing and of the fuselage zone, the cancellation of thickness along the leading edges and the new introduced integration conditions along the junction lines between the wing and fuselage zone of the thick-symmetrical FC component (in order to avoid the detachment of the flow along these lines).

According to the optimum-optimorum theory, the GO shape of the FC is searched among the elitary FCs with the same area of their planforms which belong to the same class of FCs. The class is defined by the common properties of the elitary FCs which belong to this class. The similarity parameter ν of the planform of FC is sequentially varied and a lower limit-line of the inviscid drag functional of elitary FCs, as function of this similarity parameter ν, is obtained. For FCs with subsonic leading edges is: $0 < \nu < 1$. The position of the minimum of this limit-line gives the optimal value of the similarity parameter $\nu = \nu_{opt}$ and the corresponding elitary FC is, at the same time, the global optimized FC of the class.

The author has used its OO theory for the determination of the inviscid GO shapes of three models, namely, Adela (a wing alone) and Fadet I and Fadet II (two fully-integrated wing-fuselage FCs), which are of minimum drag , respectively, at cruising Mach numbers $M_\infty = 2;\ 2.2;\ 3$. In the (Fig. 1) is represented the GO shape of the model Fadet II.

The theoretical predicted lift, pitching moment and the pressure coefficients of the upper side of six delta FC's models (namely the wedged delta, the double wedged delta, the wedged delta wing fitted with central conical fuselage, the global optimized delta wing Adela, the

global optimized and fully-integrated wing-fuselage models Fadet I and Fadet II) were measured in the trisonic wind tunnel of DLR-Köln, in the frame of research projects of the author, sponsored by the DFG.

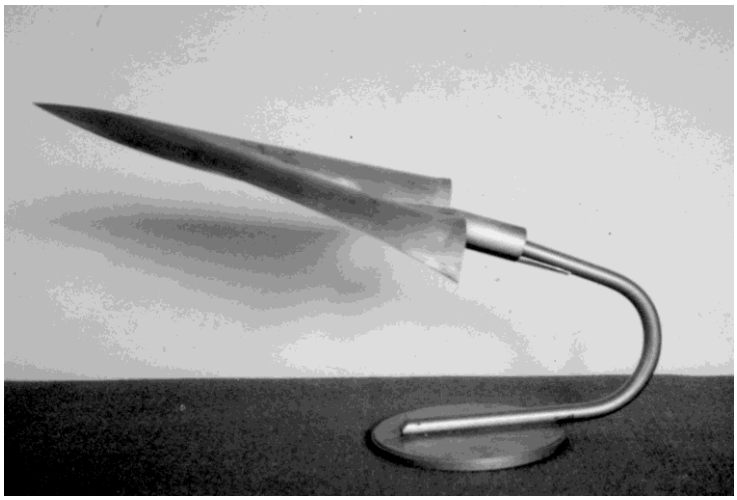
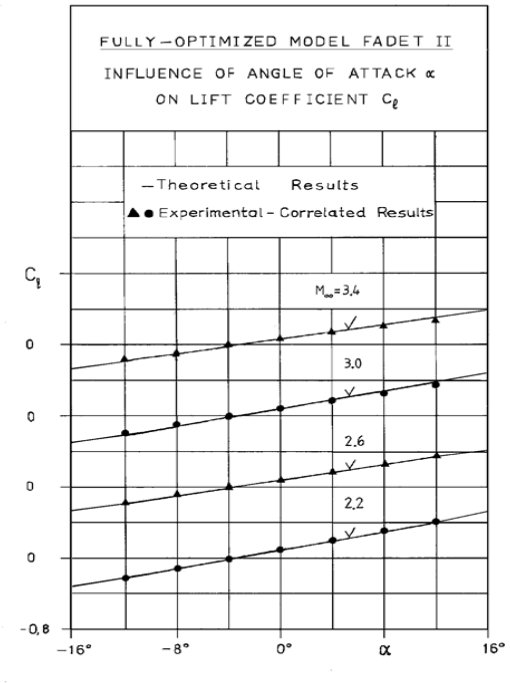
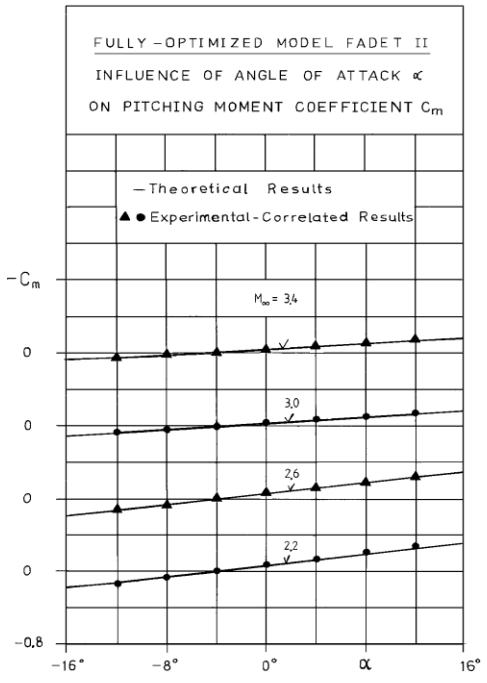
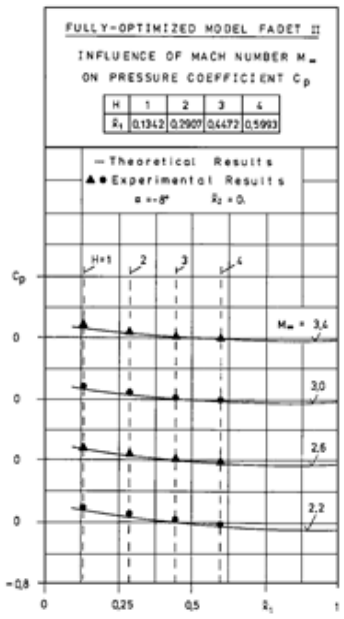


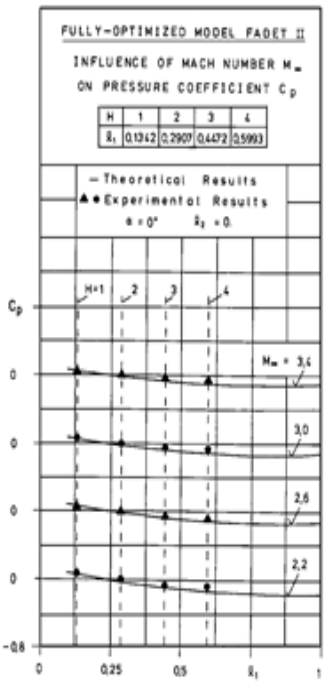
Figure 1: The Global Optimized and Fully-Integrated Shape of the Model Fadet II at Cruising Mach Number $M_\infty = 3$

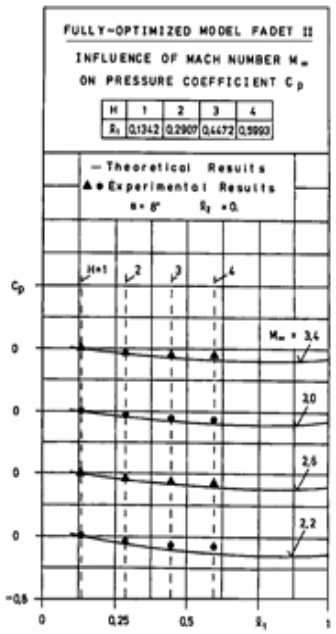




Figures 2a,b: The Variation of Lift and Pitching Moment Coefficients of Model Fadet II







Figures 3a-c: The Variations of Pressure Coefficient on the Upper Side of Model Fadet II, in its Central Longitudinal Cut, for the Angles of Attack $\alpha = -8^{\circ};\ 0^{\circ};\ 8^{\circ}$

The comparisons of theoretical and experimental-correlated values of the lift and pitching moment coefficients of all these models were in very good agreements with the experimental results. In the (Figs. 2a,b), are presented these agreements between the theoretical and the experimental correlated values of the lift and pitching moment coefficients.

In (Figs. 3a-c) are represented the agreements between the theoretical and the experimental interpolated values of pressure coefficients on the upper side of the longitudinal central cuts of the model Fadet II at the angles of attack $\alpha = -8^{\circ};\ 0^{\circ};\ 8^{\circ}$.

The second new extended variational problem concerns the determination of the GO shape of the stretched flaps, in order that entire FC to be of minimum inviscid drag at the second, lower cruising Mach number $M_{\infty}^{'}$. The central part of the FC remains unchanged and the free parameters of optimization are now the coefficients of the downwashes $\widetilde{\widetilde{w}}$ and $\widetilde{\widetilde{w}}^{*}$ on the thin and thick-symmetrical components of the flaps and the similarity parameter $\nu' = B^{'}L$ of the planform of the entire FC with stretched flaps.

The constraints for the thin component are: the given lift and pitching moment coefficients and the fulfilling of Kutta condition on the subsonic leading edges of the flaps and, for the thick-symmetrical component, the given relative volume of the flaps, null-thickness along the leading edges of the flaps (sharp leading edges) and full-integration along the junction lines wing-flaps. The optimal values of the free parameters $\widetilde{\widetilde{w}}_{\theta\sigma}$, $\widetilde{\widetilde{w}}^{*}_{\theta\sigma}$ and ν' are obtained by cancellation of their coefficients in the first variation of the inviscid Hamilton's operator.

For the computation of the total drag, including friction, hybrid NSL's solutions are further proposed.

4 HYBRID SOLUTIONS FOR THE NAVIER-STOKES LAYER

The proposed hybrid solutions for the NSL use the hyperbolic potential solutions given here of the same FC, given here as outer flow at the NSL's edge and to reinforce the NSL's solutions. The structure of the velocity's components of the NSL are expressed as products between the corresponding potential velocity's components with polynoms with arbitrary coefficients, versus a spectral variable. These coefficients are used to satisfy the NSL's PDEs, in an arbitrary chosen number of points. Let us firstly introduce a spectral variable:

$$\eta = \frac{x_3 - Z(x_1,\ x_2)}{\delta(x_1,\ x_2)} \qquad (0 \le \eta \le 1) \tag{10}$$

The proposed forms for the hybrid numerical solutions of the velocity's components are , as in [1,2], the following:

$$u_{\delta} = u_e \sum_{i=1}^{N} u_i\ \eta^i \quad , \quad v_{\delta} = v_e \sum_{i=1}^{N} v_i\ \eta^i \quad , \quad w_{\delta} = w_e \sum_{i=1}^{N} w_i\ \eta^i \ . \tag{11a-c}$$

The here introduced logarithmic density function $R = \ln \rho$ and the absolute temperature T are the following:

$$R = R_w + \left(R_e - R_w \right) \sum_{i=1}^{N} r_i \, \eta^i \quad , \qquad T = T_w + \left(T_e - T_w \right) \sum_{i=1}^{N} t_i \, \eta^i \quad . \tag{12a,b}$$

The pressure p is computed by using the physical equation of perfect gas and, for the viscosity μ , an exponential law is used:

$$p = R_g \, \rho \, T = R_g \, e^R \, T \quad , \qquad \mu = \mu_\infty \left(\frac{T}{T_\infty} \right)^{n_1} \quad . \tag{13a,b}$$

The free coefficients u_i, v_i, w_i, r_i and t_i are used to satisfy the NSL's PDEs in some chosen points. If the hybrid forms for the velocity's components (11a-c) are introduced in the continuity's PDE and the collocation method is used, the coefficients r_i are determined only as functions of the coefficients of the velocity's components, by solving a linear algebraic system and the coefficients t_i satisfy the PDE of absolute temperature and are also obtained only as functions of the coefficients of the velocity's components by solving of a transcendental algebraic system. A splitting of the NSL's PDEs is obtained and the physical entities are expressed only as function of the spectral coefficients of the velocities components and can be easy updated in an iterative process. A speed up of computation time is obtained. The coefficients of velocity's components are determined by using the impulse PDEs, which are iteratively solved.

The hybrid solutions for the NSL presented here are reinforced numerical solutions, which present important analytical properties, namely: they have correct last behaviors, they have correct jumps due to the singularities located only along the singular lines (like junction lines wing-fuselage, subsonic leading edges of the wing with retracted flaps and junction lines wing-fuselage, junction lines wing-flaps and subsonic leading edges of the flaps of the FC with stretched flaps) obtained according to the principle of minimal singularities which fulfill the jumps and the singularities are balanced, they are accurate because the partial-derivatives of velocity's components can be exactly computed, they are split due to the use of the logarithmic density function and therefore they produce a speed up of the computation time, they fulfill automatically the non-slip condition on the FCs surface, they are matched with the outer potential flow and for moderate perturbations, they are reduced to the potential solutions at the NSL's edge, if some boundary conditions are satisfied and they do not need interface. Additionally, for hyperbolic PDEs the boundary condition on its characteristic surface is automatically fulfilled . The hybrid solutions of the NSL's PDEs are used for the computation of the friction drag coefficient of the FC. The skin friction coefficient at the wall is:

$$\tau_{x_1}^{(w)} \equiv \tau_{x_1} \Big|_{\eta = 0} = \mu_f \frac{\partial \, u_\delta}{\partial \, \eta} \Big|_{\eta = 0} = \mu_f u_1 u_e \quad . \tag{14}$$

The friction drag coefficient and the total drag of the FC, with arbitrary camber, twist and thickness distributions are:

$$C_d^{(f)} = 8 \, \nu_f u_1 \int_{\widetilde{O}\widetilde{A}_1\widetilde{C}} u_e \widetilde{x}_1 \, d\widetilde{x}_1 \, d\widetilde{y} \quad , \qquad C_d^{(t)} = C_d^{(f)} + C_d^{(i)} \quad . \tag{15a,b}$$

These hybrid NSL's solutions are also used for the viscous design of the GO shape of FC.

5 ITERATIVE OPTIMUM-OPTIMORUM THEORY

The viscous iterative OO theory of the author is proposed for the viscous determination of the GO shape of the FC with flaps in retracted position in order to present a total minimum drag at the first higher cruising Mach number and also for the viscous determination of the GO shape of the flaps of the FC with stretched flaps, in order to obtain a minimum total drag for the entire wing-fuselage FC with the flaps in stretched position, at the second lower Mach number. The viscous iterative OO theory uses the inviscid hyperbolic potential solutions as start solutions and the inviscid GO shape of these FCs as surrogate models, **only** in its first step of iteration. An intermediate computational checking of this inviscid GO shape of the FC is made with own hybrid solvers, for the three-dimensional compressible NSL. The friction drag coefficient $C_d^{(f)}$ of the FC is computed and the inviscid GO shape is checked also for the structure point of view. A weak interaction aerodynamics-structure is proposed. Additional or modified constraints, introduced in order to control the camber, twist and thickness distributions of the GO shape, for structure reasons, are here proposed. In the second step of optimization, the predicted inviscid GO shape of the FC is corrected by including these additional constraints in the variational problem and of the friction drag coefficient in the drag functional. The chart flow of the iterative OO theory is given in the (Fig. 4).

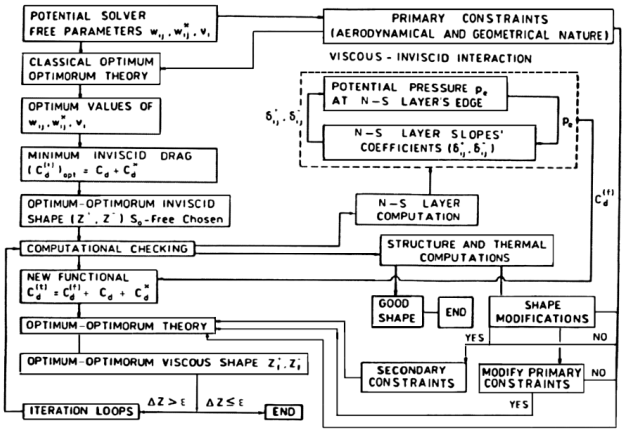


Figure 4: The Iterative Optimum-Optimorum Theory

6 CONCLUSIONS

- The author proposes analytical three-dimensional hyperbolic potential solutions for the computation of the axial disturbance velocity over the wing-fuselage FC fitted with flaps in retracted and in stretched positions, which are useful for the computation of its lift, pitching moment and pressure coefficients of the FC. These solutions are given in integrated forms, are in good agreement with experimental results and are used as start solutions for the determination of the inviscid GO shapes of FC with retracted and with stretched flaps. These potential solutions are the start solutions for the determination of the inviscid GO shape of the FC.

- The hybrid numerical solutions for the three-dimensional full PDEs of NSL, proposed here, use these hyperbolic potential solutions twice, namely: as outer flow at the NSL's edge and for the analytical hybridization of the numerical solutions. The velocity's components are expressed as products of potential solutions and polynoms with arbitrary coefficients, which are used to satisfy the NSL's PDEs in some chosen points. By using a logarithmic density a splitting of NSL's PDEs is realized.
- The hybrid numerical solutions of the NSL, proposed here, are accurate because they are meshless, their derivatives can be easy and exact computed, they automatically fulfill the non-slip condition and have important analytical properties given above, due to the analytical hybridization and due to the splitting. Artificial viscosity and correction coefficients are not used.
- The evolutive iterative optimum-optimorum theory is a special deterministic strategy, developed by the author, for the solving of the enlarged variational problem with free boundaries. The GO shape of FC is searched inside of a given class of elitary FCs. Therefore it has almost all attributes of genetic algorithms like: migrations (in the drag functional and in the constraints in the early steps of its iteration), crossover (by the construction of its NSL start solutions up the second step of its iterations), mutation (in the start solutions and in the constraints), multiple selections (inside of a class and among different classes of elitary FCs).
- It allows the performing of the multipoint global optimal design by morphing. Movable leading edge flaps are used for this purpose and the both enlarged variational problems with free boundaries, which occur, are solved.
- It allows the performing of the multidisciplinary global optimization. A weak interaction aerodynamics-structure is proposed, via additional or modified constraints requested for the structure purpose.
- It is flexible because it can use different start solutions, drag functionals and constraints, which can be changed in the early steps of iterations.
- It is economic and competitive, due to the splitting of the NSL's PDEs and of analytical hybridization a speed up of computation time occurs.

The morphing of FCs, by using spanwise movable leading edge flaps, is especially useful for supersonic aircraft of small size, like business and rescue jets, which must be able to take off and to land on shorter runways and for UAVs.

REFERENCES

[1] A. Nastase. *Computation of supersonic flow over flying configurations*. Oxford, Elsevier, UK, 2008.

[2] A. Nastase. Hybrid Navier-Stokes solutions for aerodynamical global optimal shape's design , *Proc. of International Conference EngOpt.,* Paper 750, Rio de Janeiro, Brazil, 2008.

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis(Eds)

RESISTANCE REDUCTION OF A MILITARY SHIP BY VARIABLE-ACCURACY METAMODEL-BASED MULTIDISCIPLINARY ROBUST DESIGN OPTIMIZATION

C. LEOTARDI*, E.F. CAMPANA* AND M. DIEZ*

*National Research Council-Marine Technology Research Institute (CNR-INSEAN)
Via di Vallerano 139, 00128 Rome, Italy
e-mail: c.leotardi@insean.it, (matteo.diez, emiliofortunato.campana)@cnr.it
web page: http://www.insean.cnr.it

Key words: Simulation-based design (SBD), multidisciplinary robust design optimization (MRDO), uncertainty quantification (UQ), metamodel-based optimization

Abstract. A method for simulation-based multidisciplinary robust design optimization (MRDO) affected by uncertainty is presented, based on variable-accuracy metamodelling. The approach encompasses a variable level of refinement of the design of experiments (DoE) used for the metamodel training, a variable accuracy for the uncertainty quantification (UQ), and a variable level of coupling between disciplines for the multidisciplinary analysis (MDA). The results of the present method are compared with a standard MRDO, used as a benchmark and solved by fully coupled MDA and fully accurate UQ, without metamodels. The hull-form optimization of the DTMB 5415 subject to stochastic speed is presented. A two-way steady coupled system is considered, based on hydrodynamics and rigid-body equation of motion. The objective function is the expected value of the total resistance, and the design variables pertain to the modification of the hull form. The effectiveness and the efficiency of the present method are evaluated in terms of optimal design performances and number of simulations required to achieve the optimal design.

1 INTRODUCTION

The design of complex engineering system requires simulation-based analysis, addressing the interaction of mutually coupled disciplines. Real world applications are affected by uncertainty and require uncertainty quantification (UQ) and multidisciplinary robust design optimization (MRDO) formulations. Simulation-based design (SBD) for shape optimization has been used in diverse engineering fields, including naval applications [1]. SBD has been widely extended to multidisciplinary design optimization (MDO) problems, including ship design [2]. The assessment of uncertainty in SBD has been presented in [3], whereas a MRDO application addressing operational uncertainties has been shown

in [4]. In order to reduce the MDO and MRDO computational costs, metamodels have been widely applied in several engineering fields. In naval applications, a metamodels-based UQ may be found in [5, 6] and a dynamic radial basis function metamodel for UQ applications in ship hydrodynamics has been presented in [7]. The numerical solution of the MRDO represents a challenge from both the algorithmic and computational viewpoints, especially if computationally expensive simulations are required. Simulation-based multidisciplinary analysis (MDA), UQ, and design optimization need to be effective and efficient, in order to define an optimal solution at a reasonable computational cost.

The objective of the present work is the development and validation of a variable-accuracy method for simulation-based MRDO. Specifically, the optimal solution is identified by variable-accuracy, metamodel-based design optimization. The focus is on two-way steady problems and the method encompasses (a) a variable level of refinement of the design of experiments (DoE) used for the metamodel training, (b) a variable accuracy in the UQ analysis and (c) a variable level of coupling between disciplines in MDA [8].

The SBD application pertains to the hull-form optimization of the DTMB 5415 model, an open-to-public early concept of the DDG-51, a USS Arleigh Burke-class destroyer, widely used for both experimental [9] and numerical investigations [2]. Herein, the SBD optimization is aimed at the reduction of the expected value of the total resistance in calm water, considering stochastic speed. The two-way MDA is defined by the steady hydrodynamics provided by a linear potential flow solver and the rigid body equation of motion. The convergence of MDA is achieved iteratively, for each value of the stochastic speed. Monte Carlo method coupled with Latin Hypercube Sampling (LHS) [8] is used for UQ. The optimization is performed using a single objective deterministic particle swarm optimization (PSO) algorithm [10], using a thin plate spline (TPS) metamodel built on subsequent DoEs, obtained with variable UQ accuracy and MDA coupling. The results are compared to a benchmark solution, obtained by optimization without metamodel and a high level of UQ accuracy and MDA coupling.

2 PROBLEM FORMULATION

The single-objective MDO problem is formulated as

$$\text{minimize } f(\mathbf{x}, \mathbf{a}), \qquad \mathbf{x} \in X \subseteq \mathbb{R}^{N_{\text{DV}}} \tag{1}$$

whereas the MRDO extension to problems affected by uncertainty reads

$$\text{minimize } \mu(f) = \int_Y f(\mathbf{x}, \mathbf{y}, \mathbf{a}) p(\mathbf{y}) d\mathbf{y}, \qquad \mathbf{x} \in X \subseteq \mathbb{R}^{N_{\text{DV}}} \tag{2}$$

where $\mathbf{x}$ collects N_{DV} deterministic design variables, f is the deterministic objective function, μ in Eq. 2 is the expected value of f and $p(\mathbf{y})$ is the probability density function of the stochastic environmental and operating conditions, collected in $\mathbf{y}$. Box and functional constraints may apply, if required.

The function f depends on several interconnected disciplines. The input of the i-th discipline Δ_i is defined by the set of design variables, $\mathbf{x} = [\mathbf{x}_i^T, \mathbf{x}_{\text{S}}^T]^T$, the set of output

parameters provided by other disciplines Δ_j $(i \neq j)$, $\{\mathbf{a}_j\}_{i \neq j}$, and, for Eq. 2, the set of uncertain parameters $\mathbf{y} = [\mathbf{y}_i^T, \mathbf{y}_\mathrm{S}^T]^T$. Variables indicated by $\mathbf{x}_\mathrm{S}$ are shared by all (or part of) the disciplines, whereas the corresponding $\mathbf{x}_i$ are assumed to be local to the i-th discipline Δ_i. Similarly, the vector $\mathbf{y}_\mathrm{S}$ is shared by all (or part of) the disciplines involved and $\mathbf{y}_i$ is local to the i-th discipline Δ_i only (see, e.g., [4]).

In MDO, once the multidisciplinary equilibrium, $\mathbf{a} = [\mathbf{a}_1^T, \ldots, \mathbf{a}_{N_\Delta}^T]^T$, is achieved (generally by iterative procedures), the deterministic objective function $f = f(\mathbf{x}, \mathbf{a})$ is evaluated and an optimization algorithm is put on top of MDA.

In MRDO, the multidisciplinary equilibrium is conditional to $\mathbf{y} \in Y$ and UQ is needed on top of MDA (as shown in the top box of Fig. 1). Once the multidisciplinary equilibrium is achieved, the deterministic objective function $f = f(\mathbf{x}, \mathbf{y}, \mathbf{a})$ is evaluated, the stochastic objective function $\mu(f)$ is assessed by UQ, and finally an optimization algorithm is put on top of UQ (see Fig. 1).

3 VARIABLE-ACCURACY METHOD FOR MRDO

A metamodel is interposed between UQ and the optimizer, as shown in Fig. 1. The variable-accuracy metamodel-based MRDO is based on subsequent optimization stages, characterized by: (a) a refinement of the DoE used for the metamodel training, (b) a variable accuracy in the UQ analysis and (c) a variable level of coupling in MDA.

At the first stage, the training points are distributed in the whole design domain and the corresponding objective function values are obtained considering both a low level of accuracy in UQ and a weak coupling between disciplines. After the first optimization stage, a refined subdomain centered in the current optimum is defined. A new training set is used, with the corresponding objective function values obtained increasing both the accuracy of UQ and the coupling in MDA. The procedure is iterated for an appropriate number of stages, achieving fully accurate UQ and fully converged MDA, at the last optimization stage. The pseudo-code of the methodology is presented in Fig. 2, where N_OS is the number of optimization stages and N_TP is the number of training points per optimization stage.

4 SBD FRAMEWORK FOR MRDO

The SBD optimization framework encompasses three essential and interconnected elements: (a) the analysis tools, (b) the optimization algorithm, (c) the tool for the design modifications. The present toolbox includes a steady potential flow code, coupled with rigid-body equation of motion, a UQ tool based on MC-LHS simulation, a TPS metamodel, a deterministic version of the PSO algorithm, and a tool for geometry modifications based on orthogonal basis functions.

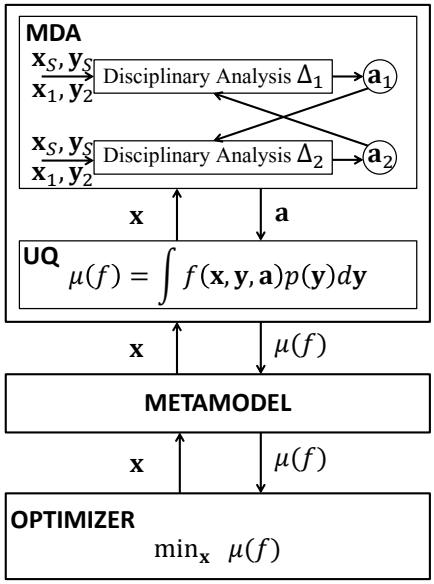


Figure 1: Metamodel-based MRDO method.

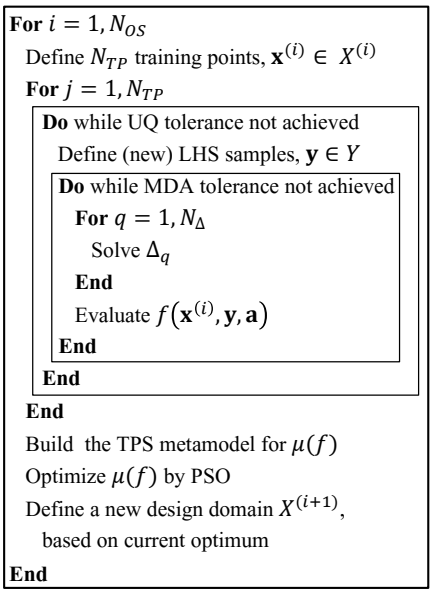


Figure 2: Pseudo-code of the variable-accuracy metamodel-based MRDO method.

4.1 Tools for multidisciplinary analysis

The hydrodynamics is solved using the code WARP (WAve Resistance Program), developed at CNR-INSEAN. Wave resistance computations are based on linear potential flow theory. For details of equations, numerical implementation and validation see [11]. The wave resistance is evaluated with the transverse wave cut method [12], whereas the frictional resistance is estimated using a flat-plate approximation, based on the local Reynolds number [13]. The inputs ($\mathbf{x}_1$ and $\mathbf{y}_1$) of the hydrodynamic solver are the design variables and the speed, respectively, along with sinkage (σ) and trim (τ) values (collected in $\mathbf{a_2}$). The corresponding outputs (collected in $\mathbf{a}_1$) are the hydrostatic and hydrodynamic forces and moments.

The steady rigid-body equation of motion is evaluated for the 2DOF (sinkage and trim) problem and includes the equilibrium of the vertical forces F_z, and y-moments M_y. The incremental sinkage ($\Delta\sigma$) at the center of gravity (CG) and trim ($\Delta\tau$) are evaluated by the linearized equations, $\Delta\sigma = F_z/\rho g S_{\mathrm{WL}}$ (where ρ is the water density, g is the gravity acceleration and S_{WL} is the waterline surface area), and $\Delta\tau = M_y/\rho g I_{\mathrm{WL}}$ (where I_{WL} is the waterline-area moment of inertia about the y-axis, passing at CG). The inputs are the hydrostatic and hydrodynamic forces and moments ($\mathbf{a}_1$), and the weight force, whereas the outputs ($\mathbf{a}_2$) are formed by sinkage and trim values.

The MDA achieves the equilibrium solution for $\mathbf{a}_1$ and $\mathbf{a}_2$ by an iterative procedure.

4.2 UQ method

The integral in Eq. 2 is approximated as $\mu(f) = 1/N_{\mathrm{UQ}} \sum_{i=1}^{N_{\mathrm{UQ}}} f(\mathbf{x}, \mathbf{y}_i)$, using the MC method. LHS is applied by dividing the uncertain parameter domain with $N_{\mathrm{UQ}} = 2^k + 1$ ($k \in \mathbb{N}$) evenly spaced bins [8].

4.3 TPS metamodel

A TPS metamodel is used for the design optimization. Given a set of training points $\{\mathbf{x}_i\}, i = 1, \ldots, N_{\mathrm{TP}}$, and the corresponding objective function values $f(\mathbf{x}_i)$, the objective function $f(\mathbf{x})$ is approximated, using radial-basis functions (RBF), as per $\hat{f}(\mathbf{x}) = \sum_{i=1}^{N_{\mathrm{TP}}} d_i \, \varphi[r(\mathbf{x}, \mathbf{x}_i)]$, where $\varphi(r) = r^2 \log r$ is the RBF kernel, with $r = \|\mathbf{x} - \mathbf{x}_i\|$. The coefficients d_i are the solution of $\{\varphi(\mathbf{x}_i, \mathbf{x}_j)\}\{d_j\} = \{f(\mathbf{x}_i)\}$. Herein, the training set is distributed in the design domain using Hammersley sequence sampling (HSS) [8].

4.4 Optimization algorithm

The deterministic version of the PSO algorithm is used for the optimization [10]. The swarm dimension is set to $4 \cdot N_{\mathrm{DV}}$, the swarm is initialized using HSS over the design variables domain and its boundaries. The PSO coefficients are set as $\chi = 0.721$, $c_1 = c_2 = 1.655$, [10]. The number of function evaluations is set to $N_{\mathrm{PSO}} = 512 \cdot N_{\mathrm{DV}}$.

4.5 Shape modification method

Shape modifications are represented in terms of orthogonal basis functions $\boldsymbol{\psi}_j$ ($j = 1, ..., N_{\mathrm{DV}}$), defined over surface-body patches as

$$\boldsymbol{\psi}_j(\xi, \eta) := \alpha_j \sin\left(\frac{p_j \pi \xi}{A_j} + \phi_j\right) \sin\left(\frac{q_j \pi \eta}{B_j} + \chi_j\right) \mathbf{e}_{k(j)} \qquad (\xi, \eta) \in [0; A] \times [0; B] \tag{3}$$

where α_j is the j-th (dimensional) design variable; p_j and q_j define the order of the basis function in ξ and η direction respectively; ϕ_j and χ_j are the corresponding spatial phases; A_j and B_j are the patch extension in ξ and η respectively, and $\mathbf{e}_{k(j)}$ is a unit vector. Modifications may be applied in x, y or z direction ($k(j) = 1, 2, 3$ respectively).

5 OPTIMIZATION PROBLEMS

The objective function for the problem in Eq. 1 is the total resistance ($f = R_T$) of the DTMB 5415 [9] in calm water at 24 [kn], whereas the objective function for the problem in Eq. 2 is the expected value of the total resistance in calm water, evaluated over a stochastic speed y, with $y \in [18; 30]$ [kn] following a uniform probability density function.

Two (normalized) design variables, $x_1 = \alpha_1/2$ and $x_2 = \alpha_2$, are used. The shape modifications are obtained as per Eq. 3 with $j = 1, 2$ and $k = 2$. The associated parameters are $p_1 = 2.0$, $\phi_1 = 0.0$, $q_1 = 1.0$, $\chi_1 = 0.0$, $\alpha_1 \in [-2.0; 2.0]$ and $x_1 \in [-1.0; 1.0]$; $p_2 = 1.0$, $\phi_2 = 0.0$, $q_2 = 2.0$, $\chi_2 = 0.0$, $\alpha_2 \in [-1.0; 1.0]$ and $x_2 \in [-1.0; 1.0]$. Geometric constraints include fixed length between perpendiculars (L_{BP}) and displacement.

Potential flow calculations are performed using a 150x30 panel grid for the hull surface. The computational domain for the free surface is defined within 1 hull length upstream, 3 lengths downstream and 1.5 lengths aside, discretized with 30x44, 90x44 and 30x44 surface panels, respectively. The grid convergence analysis is provided in [14].

6 NUMERICAL RESULTS

Benchmark solutions for both deterministic MDO and stochastic MRDO are presented in the following, and compared to the variable-accuracy metamodel-based results.

6.1 Benchmark solution for MDO

Convergence of iterative MDA is conducted for both design variables and stochastic parameter at domain center ($x_1 = x_2 = 0.0$, $y = 24$ [kn], corresponding to $Fr = y/\sqrt{gL_{\mathrm{BP}}} = 0.330$). Fig. 3 shows the solution change for total resistance, pitch and vertical force coefficients, sinkage and trim (respectively $C_t = 2F_x/\rho y^2 S_{\mathrm{WL}}$, $C_{My} = 2M_y^h/\rho y^2 S_{\mathrm{WL}}$, $C_{Fz} = 2F_z^h/\rho y^2 S_{\mathrm{WL}}$, σ and τ, where F_x is the total resistance, M_y^h is the hydrodynamic moment, F_z^h is the hydrodynamic force) versus MDA iterations.

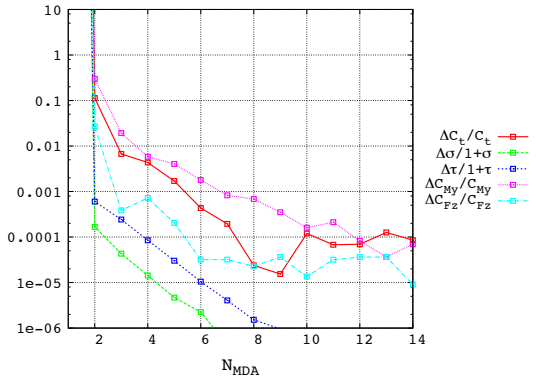


Figure 3: MDA convergence ($x_1 = x_2 = 0$, $y = 0.330$).

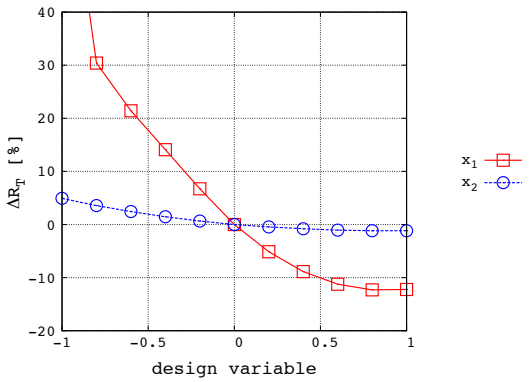


Figure 4: Sensitivity analysis for calm-water total resistance at $Fr = 0.330$.

A benchmark tolerance of 10^{-4} is set for all parameters and the maximum number of MDA iteration $N_{\mathrm{MDA}}^{(\max)}$ is set equal to 14. The corresponding maximum number of discipline evaluations is defined as $N_{\mathrm{S}}^{(\max)} = N_{\mathrm{PSO}} \cdot N_{\mathrm{MDA}}^{(\max)}$ and summarized in Tab. 1.

The sensitivity analysis for the resulting R_T is shown in Fig. 4, versus x_1 and x_2. The benchmark MDO results are included in Tab. 2, in terms of optimal design variables, objective function value and number of discipline evaluations.

Table 1: MDO and MRDO maximum number of simulations.

Problem	ID	Evaluations
MDO	N_{PSO}	$512 \cdot N_{DV}$
	$N_{\text{MDA}}^{(\max)}$	14
	$N_{\text{S}}^{(\max)}$	$7{,}168 \cdot N_{DV}$

Problem	ID	Evaluations
MRDO	N_{PSO}	$512 \cdot N_{DV}$
	$N_{\text{UQ}}^{(\max)}$	33
	$N_{\text{MDA}}^{(\max)}$	14
	$N_{\text{S}}^{(\max)}$	$236{,}554 \cdot N_{DV}$

Table 2: MDO and MRDO results.

Problem	Design var. 1	Design var. 2	Obj. [kN]	$\left\lvert\frac{Obj.^{(\text{B})} - Obj.}{Obj.^{(\text{B})}}\right\rvert$ %	N_{S}/N_{DV}
MDO (benchmark)	0.99	0.99	573.64	-	5,024
Variable-coupling MDO	0.96	0.99	571.30	0.41	674
MRDO (benchmark)	1.00	0.53	712.34	-	167,587
Variable-accuracy MRDO	0.98	0.36	712.82	0.07	18,876

6.2 Variable-accuracy MDO

Three subsequent optimization stages are considered, the MDA coupling is increased at each stage, whereas the number of training points N_{TP} is fixed for each stage. The tolerance for the solution change of C_t, C_{My}, C_{Fz}, σ and τ is set to 10^{-2}, 10^{-3} and 10^{-4} for first, second and third optimization stage, respectively.

In order to define N_{TP}: (1) a target number of total evaluations $N_{\text{S}}^{(\text{T})} = 0.1\, N_{\text{S}}^{(\max)}$ is assumed; (2) at the j-th stage, the number of evaluations is calculated as $N_S^{(j)} = N_{\text{TP}} \cdot N_{\text{MDA}}^{(j)}$; (3) a target number of MDA iterations is defined according to Fig. 3, providing $N_{\text{MDA}}^{(T)} = 4,\ 7,\ 14$ for first, second and third optimization stage. Accordingly, N_{TP} equals 29.

Table 3: Metamodel-based MDO and MRDO evaluations.

Problem	ID	Stage 1	Stage 2	Stage 3	$N_{\text{S}}/N_{\text{DV}}$
Variable-coupling MDO	N_{TP}	29	29	29	
	$N_{\text{S}}^{(j)}$	87	210	377	674
Variable-accuracy MRDO	N_{TP}	39	39	39	
	$N_{\text{S}}^{(j)}$	1,209	4,680	12,987	18,876

Variable-coupling metamodel-based MDO results are included in Tab. 2 and are found in close agreement with the benchmark values. The number of evaluations for each optimization stage is summarized in Tab. 3. The number of evaluations associated to the variable coupling metamodel-based MDO is about 13.4% of the corresponding benchmark value in Tab. 2. Figure 5 shows the subsequent metamodel-based optimization stages, with the corresponding optima.

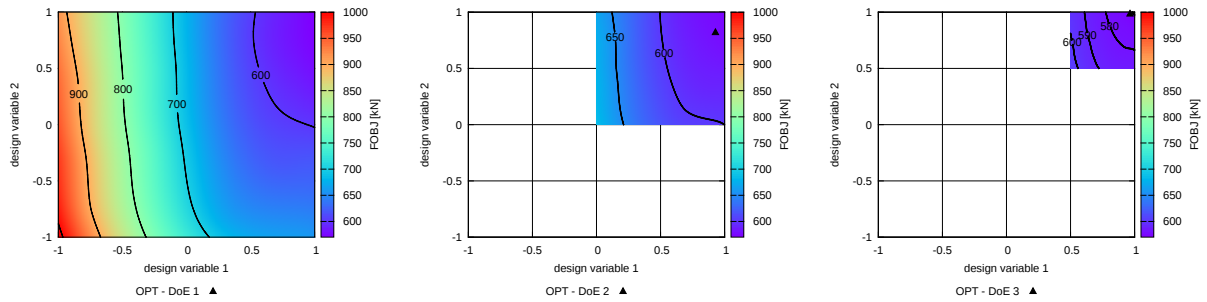


Figure 5: Metamodel-based MDO with subsequent refinement of design variable domain.

6.3 Benchmark solution for MRDO

Convergence of UQ is studied versus the number of samples $N_{\mathrm{UQ}} = 2^k + 1,\ k \in \mathbb{N}$, for $x_1 = x_2 = 0.0$ and a benchmark tolerance for MDA equal to 10^{-4}. Fig. 6 shows the solution change of the expected value of the total resistance $(\Delta\mu/\mu)$ versus N_{UQ}.

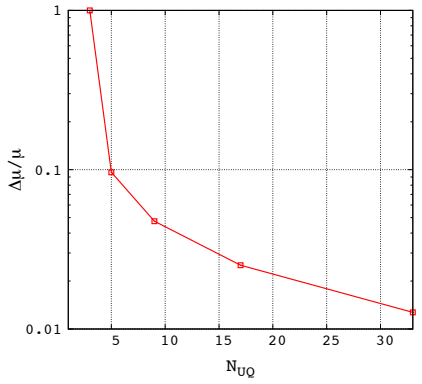


Figure 6: UQ convergence ($x_1 = x_2 = 0$, $y \in [0.250, 0.410]$).

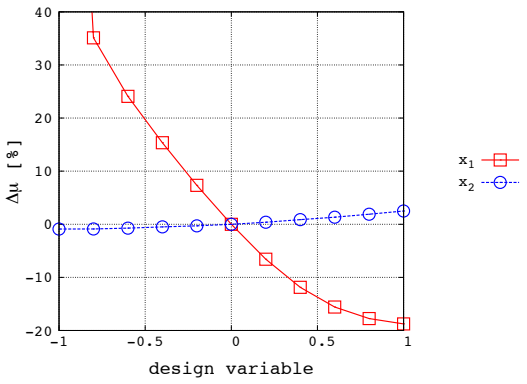


Figure 7: Sensitivity analysis for calm water expected value of total resistance.

A benchmark tolerance for solution change is set to $2{\cdot}10^{-2}$ and a number of $N_{\mathrm{UQ}}^{(\max)} = 33$ is set. The maximum number of evaluations is defined as $N_{\mathrm{S}}^{(\max)} = N_{\mathrm{PSO}} \cdot N_{\mathrm{UQ}}^{(\max)} \cdot N_{\mathrm{MDA}}^{(\max)}$

and summarized in Tab. 1.

The sensitivity analysis for each design variable is shown in Fig. 7, whereas the benchmark MRDO results are included in Tab. 2.

6.4 Variable-accuracy MRDO

Three subsequent optimizations stages are considered, the UQ accuracy, together with the MDA coupling, is increased at each stage, whereas the number of training points N_{TP} is fixed for each optimization stage. The tolerance for the solution change of the expected value of the total resistance is set to $5 \cdot 10^{-2}$, $3 \cdot 10^{-2}$ and $2 \cdot 10^{-2}$, for the first, second and third optimization stage, respectively. The MDA tolerances are the same as for the MDO problem.

In order to define N_{TP}: (1) a target number of total evaluations $N_{\mathrm{S}}^{(\mathrm{T})} = 0.1\, N_{\mathrm{S}}^{(\mathrm{max})}$ is assumed; (2) the number of evaluations for the j-th optimization stage is expressed as $N_S^{(j)} = N_{\mathrm{TP}} \cdot N_{\mathrm{UQ}}^{(j)} \cdot N_{\mathrm{MDA}}^{(j)}$; (3) a target number of UQ samples is selected according to Fig. 6), which provides $N_{\mathrm{UQ}}^{(T)} = 9,\ 17,\ 33$ for the first, second and third optimization stage; (4) the target number of MDA iterations is the same as the MDO problem. Accordingly, N_{TP} is found equal 39.

Metamodel-based MRDO results are included in Tab. 2 and are found in close agreement with benchmark values. The number of evaluations for each optimization stage are summarized in Tab. 3. The number of simulations associated to the variable-accuracy metamodel-based MRDO is about 11.3% of the corresponding benchmark values in Tab. 2. Figure 8 shows the subsequent metamodel-based optimization stages with the corresponding optima.

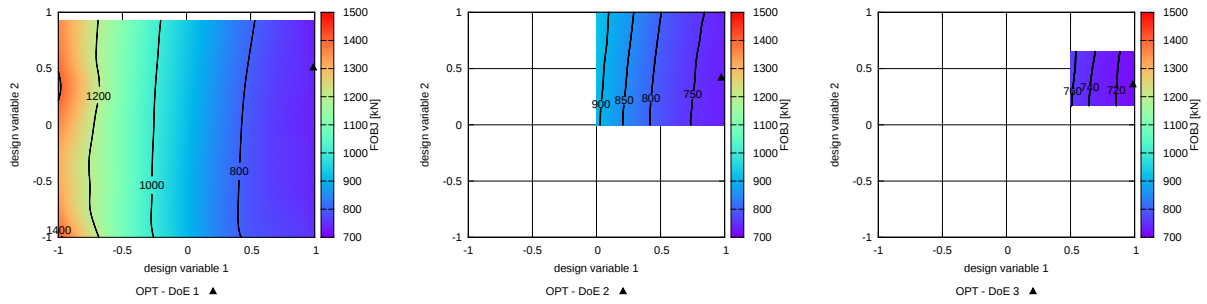
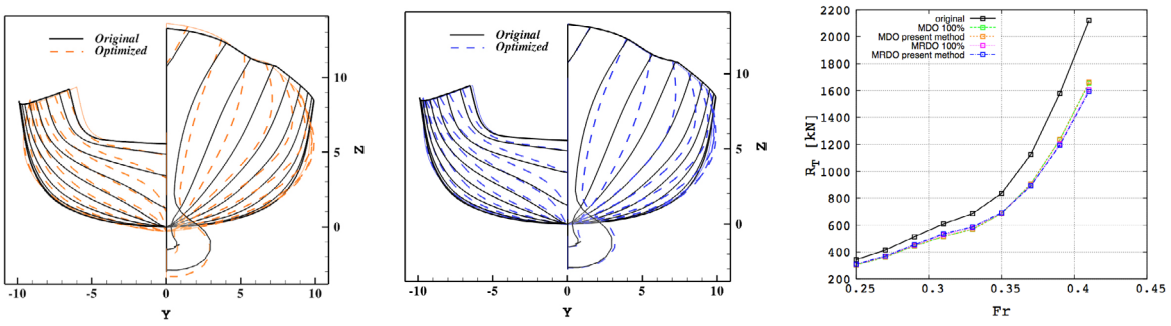


Figure 8: Metamodel-based MRDO with subsequent refinement of design variable domain.

Figure 9 (a) and (b) show the MDO and MRDO optimal shapes, compared to the original. It is worth noting that the optimal MRDO and MDO configurations fall in different points of the design variable domain. Nevertheless, the corresponding shapes are similar, as also found in earlier research for a catamaran configuration [5].

Finally, a parametric analysis of the total resistance versus Fr is presented in Fig. 9 (c), for the optimal and original configurations. The performances of benchmark and variable-



(a) (b) (c)

Figure 9: (a) Optimal MDO hull shape *vs* original, (b) Optimal MRDO hull shape *vs* original, (c) Parametric analysis of the total resistance *vs* Fr number.

accuracy solutions are found in close agreement for both MDO and MRDO, and present an overall improvement in the whole speed range. Moreover, MDO and MRDO solutions perform similarly; however the MRDO solution presents an overall best performance, whereas the performance of the MDO solution is found slightly better for deterministic $Fr = 0.330$.

7 CONCLUSIONS

A methodology for simulation-based multidisciplinary design optimization for problems affected by uncertainty has been presented. The approach encompasses a variable-accuracy metamodel-based design optimization, with (a) a variable level of refinement of the DoE used for the metamodel training, (b) a variable accuracy in the UQ analysis, and (c) a variable level of coupling in MDA.

The methodology has been applied to the resistance optimization of the the DTMB 5415 model, within a stochastic speed range. A potential flow solver has been coupled with the rigid body equation of motion (steady), the UQ has been performed by MC-LHS, a TPS metamodel has been used for the design optimization, and a PSO algorithm has been used for the optimization. Two design variables have been used, for the hull form modification.

Both deterministic MDO and stochastic MRDO have been solved. MDO and MRDO variable-accuracy metamodel-based solutions have been found in very close agreement with the corresponding benchmark solutions (obtained by fully convergent UQ and MDA, without metamodels). The present method allowed for a reduction of the computational cost by 13.4% and 11.3% for MDO and MRDO, respectively (Tab. 3).

The parametric analysis, conducted for the optimal designs over the speed range, revealed that the MRDO solution presents an overall best performance, whereas the MDO solution shows the best deterministic performance. As in earlier research, MDO and MRDO solutions have been found similar.

Future work will focus on the use of different surrogate techniques. Specifically, the

current *a priori* definition of the training points for the metamodel will be replaced by dynamic metamodelling techniques (e.g. [7]). The sensitivity of the results to diverse metamodels will be also investigated. The extensions to higher-dimensional problems (with more than two design variables) will be also addressed.

Acknowledgements

The present research is supported by the Italian Flagship Project RITMARE, coordinated by the Italian National Research Council and funded by the Italian Ministry of Education, Research Program 2011-2013, and by the US Navy Office of Naval Research, NICOP grant N62909-11-1-7011, under the administration of Dr Ki-Han Kim and Dr Woei-Min Lin.

REFERENCES

[1] Diez, M., Campana, E.F. and Stern, F. Design-space dimensionality reduction in shape optimization by Karhunen-Loève expansion. *Computer Methods in Applied Mechanics and Engineering* (2015), **283**(0):1525 -1544.

[2] Campana, E.F., Peri, D., Tahara, Y. and Stern, F. Shape optimization in ship hydrodynamics using computational fluid dynamics. *Computer Methods in Applied Mechanics and Engineering* (2006) **196**(1-3):634-651.

[3] Du, X. and Chen, W. Efficient uncertainty analysis methods for multidisciplinary robust design. *AIAA Journal* (2002) **40**:545-552.

[4] Diez, M., Peri, D., Fasano, G. and Campana, E.F. Hydroelastic optimization of a keel fin of a sailing boat: a multidisciplinary robust formulation for ship design. *Structural and Multidisciplinary Optimization* (2012) doi:10.1007/s00158-012-0783-7.

[5] Diez, M., He, W., Campana, E.F. and Stern, F. Uncertainty quantification of Delft catamaran resistance, sinkage and trim for variable Froude number and geometry using metamodels, quadrature and Karhunen-Loeve expansion. *J Mar Sci Technol* (2013) doi 10.1007/s00773-013-0235-0.

[6] He, W., Diez, M., Campana, E.F., Stern, F. and Zou, Z. A one-dimensional polynomial chaos method in CFD-Based uncertainty quantification for ship hydrodynamic performance. *Journal of Hydrodynamics* (2013) **25**(5):655-662.

[7] Volpi, S., Diez, M., Gaul, N.J., Song, H., Iemma, U., Choi, K.K., Campana, E.F. and Stern, F. Development and validation of a dynamic metamodel based on stochastic radial basis functions and uncertainty quantification. *Structural Multidisciplinary Optimization* (2014) doi 10.1007/s00158-014-1128-5.

[8] Leotardi, C., Diez, M., Serani, A., Iemma, U. and Campana E.F. A framework for efficient simulation-based multidisciplinary robust design optimization with application to a keel fin of a racing sailboat. *Proceedings of the 1st International Conference in Engeneering and Applied Sciences Optimization* (2014) pp. 1177-1193, ISBN: 978-960-99994-5-8.

[9] Stern, F., Longo, J., Penna, R., Olivieri, A., Ratcliffe, T. and Coleman, H. International Collaboration on Benchmark CFD Validation Data for Surface Combatant DTMB Model 5415. *Proceedings of the Twenty-Third Symposium on Naval Hydrodynamics* (2000) Val de Reuil, France, September 17-22.

[10] Serani, A., Diez, M., Leotardi, C., Peri, D., Fasano, G., Iemma, U. and Campana E.F. On the use of synchronous and asynchronous single-objective deterministic Particle Swarm Optimization in ship design problems. *Proceedings of the 1st International Conference in Engeneering and Applied Sciences Optimization* (2014) pp. 1218-1240, ISBN: 978-960-99994-5-8.

[11] Bassanini, P., Bulgarelli, U., Campana, E.F. and Lalli, F. The wave resistance problem in a boundary integral formulation. *Surv. Math. Ind.* (1994) **4**:151–194.

[12] Telste, J.G. and Reed, A.M. Calculation of Transom Stern Flows. *Proceedings of the Sixth International Conference on Numerical Ship Hydrodynamics* (1994) 23–69.

[13] Schlichting, H. and Gersten, K. *Boundary-Layer Theory*, Springer-Verlag, Berlin 2000.

[14] Serani, A., Diez, M. and Campana E.F. Single- and multi-objective design optimization study for DTMB 5415, based on low fidelity solvers. *INSEAN Technical Report 2015-TR-002* (2015).

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis (Eds)

STRUCTURAL DESIGN USING A PARALLEL SEQUENTIAL APPROXIMATE OPTIMIZATION

DONGHUI WANNG*, ZEPING WU*, FAN HU* AND ZHENYU JIANG*

* College of Aerospace Science and Engineering
National University of Defense Technology
109 Deya Road, 410073 Changsha, China
e-mail:lightblue117@gmail.com

Key words: Sequential approximate optimization, Structural design optimization, Parallel sampling strategy.

Abstract. In this paper, a parallel sampling strategy dedicated to SAO is proposed to enhance the competence of exploration and convergence simultaneously in the optimization process. In the parallel sampling procedure, new sampling points are identified in the pareto front of the multi-objective optimization problem, which is solved with the Normal constraint (NC) method. The objectives of the optimization problem are the two indices to represent the competence of exploration and convergence, and the sampled points will be evaluated by the true model in a parallel way. Furthermore, the proposed methodology is evaluated on two benchmark tests. Compared to other optimization algorithms, the PSAO algorithm yields equivalent or better objective values while the number of optimization iterations required to find the same global optima is reduced by multiple orders of magnitude, which substantially reduce the computing costs in the tested structural design optimization tasks, highlighting the applicability of the PSAO structural design optimization problems.

1 INTRODUCTION

Structural design optimization remains an important and challenging topic in the engineering design of lighter, more effective structures [1]. Design optimization aims to determine the optimal shape of a structure by maximizing or minimizing a given criterion, such as stiffness or weight, subject to stress or displacement constraints. Over the past decade, a number of optimization algorithms have been extensively used in structural optimization tasks, such as gradient-based algorithms, evolutionary algorithms (EAs) and approximation-based optimization algorithms [2]. Of course, advantages and disadvantages are associated with any optimization technique.

Several examples of gradient-based optimization applied to structural design problems exist in the literature [3, 4]. However, gradient-based optimization techniques are extremely sensitive to the initial guess and prone to trapping in local optima.

In the past few decades, structural design optimization problems have been increasingly solved by EAs such as the genetic algorithm [5], simulated annealing [6], particle swarm optimization [7] and the artificial bee colony algorithm [8]. EAs present several advantages over gradient-based methods [9]. Drawbacks of these methods are the huge number of function

evaluations required.

In approximation-based optimization techniques, objective functions are expressed as low order polynomial approximations to explicit functions. The accuracy of these techniques is acceptable and the computational cost is much reduced. Commonly applied approximation techniques include the response surface method [10, 11], neural network [12, 13], polynomial regression models [14], Kriging methods [15] and the radial basis function (RBF) [16]. However, these techniques introduce error into the meta-model, which reduces their reliability.

The SAO algorithm has been recognized as one of the most attractive approaches for engineering optimization [17]. The success of an SAO algorithm depends chiefly on the approximation technique and the sampling strategy. Therefore, this paper focuses on improving the sampling strategy to the extent that SAO becomes applicable to structural design optimization problems. The paper is structured as follows: Section 2 introduces the general formulation of the SAO approach and proposes a parallel sampling strategy. Finally, the procedure of the PSAO is presented. Section 3 presents two structural optimization case studies, which are used to demonstrate the efficacy of the methodology. Concluding remarks are presented in Section 4.

2 PARALLEL SEQUENTIAL APPROXIMATION OPTIMIZATION

2.1 General framework of the sequential approximation optimization

Let us consider a general structural optimization problem with constraints:

$$\begin{aligned} &\text{find } \boldsymbol{X} \\ &\min \ f(\boldsymbol{X}) \\ &s.t. \ \ g_i(\boldsymbol{X}) \le 0 \quad i = 1,2,\cdots,l \\ &\quad h_j(\boldsymbol{X}) = 0 \quad j = 1,2,\cdots,k \\ &\quad \boldsymbol{X}^L \le \boldsymbol{X} \le \boldsymbol{X}^U \end{aligned} \tag{1}$$

For most structural design problems, the objective function and constraints are implicit functions of design variables, usually obtained by finite element analysis (FEA). Since the computational cost for FEA may be high, the number of analyses carried out during the optimization has the main impact on the efficiency of the algorithm. This has initiated the development of optimization techniques that are suitable for structural design problems [18].

In the classical approximation-based optimization procedure summarized in [19], the accuracy of the surrogate model could be degraded by an ill-chosen initial sample, leading to a deceptively positioned optimum. Here we assume that our optimum design is the best result of the true function, not that of the surrogate. Results from the surrogate are therefore evaluated by comparison with the true function evaluations. Additional calls to the true function are used to both validate the surrogate and enhance its accuracy. Thus, as shown in Figure 1, the SAO approach selects the new points at which the true function is called. Applying a series of new infill points based on some infill criteria (also known as a sampling strategy), the objective function is sampled using a constantly changing surrogate model [20].

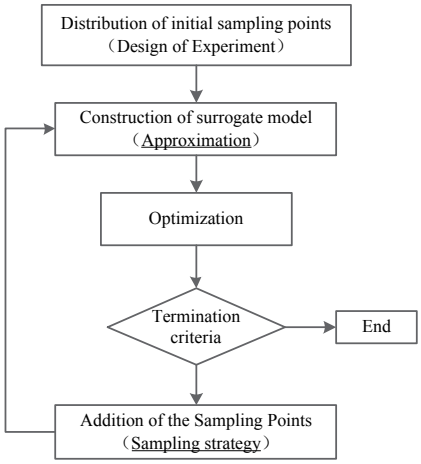


Figure 1: General framework of the sequential approximation optimization

The SAO-based structural optimization can be mathematically expressed as:

$$\begin{aligned} &\text{for} \quad n = 1,2,\cdots \\ &\text{find} \quad \boldsymbol{x} \\ &\min \ f^{(n)}(\boldsymbol{x}) \\ &s.t. \ \ g_i^{(n)}(\boldsymbol{x}) \le 0 \quad i = 1,2,\cdots,l \\ &\qquad h_j^{(n)}(\boldsymbol{x}) = 0 \quad j = 1,2,\cdots,k \\ &\qquad S^{(n)}(\boldsymbol{x}) \le 0 \\ &\qquad \boldsymbol{x}^L \le \boldsymbol{x} \le \boldsymbol{x}^U \end{aligned} \tag{2}$$

the m-dimensional design variable $\boldsymbol{X}$ is scaled into an m-dimensional unit hypercube $\boldsymbol{x}$ by

$$x_i = \frac{X_i - X_i^L}{X_i^U - X_i^L} \quad i = 1,2,\cdots,m \tag{3}$$

Here $f^{(n)}(\boldsymbol{x})$ is the nth approximation of the objective function, $g_i^{(n)}(\boldsymbol{x})$ and $h_j^{(n)}(\boldsymbol{x})$ are the nth approximations of the constraints. In addition, $S^{(n)}(\boldsymbol{x})$ is a constraint deduced from the sampling strategy in the nth iteration.

2.2 Parallel sampling strategy

Since its proposal, a number of sampling strategies have been applied to the SAO algorithm [21, 22]. These strategies can be roughly divided into three categories: exploitation, exploration and balanced exploitation/exploration. The sampling strategies are usually performed by infilling the new point sequentially, which imposes restrictions on the capability of exploration and convergence. In this section, a parallel sampling strategy is proposed.

The simplest exploitation sampling strategy for SAO is to find the optimum of the surrogate model $s(\boldsymbol{x})$, while exploration maximizes the minimal Euclidean distance between sampling points $d(\boldsymbol{x})$, given by

$$d(\boldsymbol{x}) = \min(\sqrt{(\boldsymbol{x} - \boldsymbol{x}_i^{(n)})^T (\boldsymbol{x} - \boldsymbol{x}_i^{(n)})}) \quad (i = 1, 2, \cdots, N^{(n)}) \tag{4}$$

where $N^{(n)}$ is the number of sampling points before the nth sampling. Here we adopt an adaptive sampling strategy by solving the multi-objective optimization problem

$$\begin{aligned} max: \ & s^{(n)}(\boldsymbol{x}),\ d(\boldsymbol{x}) && \boldsymbol{x}_{\min} \le \boldsymbol{x} \le \boldsymbol{x}_{\max} \\ s.t. \ & g_i^{(n)}(\boldsymbol{x}) \le 0 && i = 1, 2, \cdots, l \\ & h_j^{(n)}(\boldsymbol{x}) = 0 && j = 1, 2, \cdots, k \end{aligned} \tag{5}$$

where l is the number of inequality constraints and m is the number of equality constraints of original optimization problems, $s^{(n)}(\boldsymbol{x})$ is the meta-model constructed before the nth sequential sampling. Then the optimal solution of (1) together with the real response evaluated by the original model will be regarded as new sampling points to update the surrogate model. Solving problem (5), the Pareto front is easy to be obtained by Multi-objective Optimization Evolutionary Algorithms (MOEAs), such as NSGA or NSGA-II.

Given the number of parallel sampling points, denoted by k , the procedure of parallel sampling is shown as follows:

Step 1: solve the multi-objective optimization problem (5) to get the Pareto front;

Step 2: select k points on the Pareto front evenly as the sampling points, which is illustrated in Figure 2;

Step 3: Once these k sampling points are identified, those selected points are evaluated in parallel using the true model.

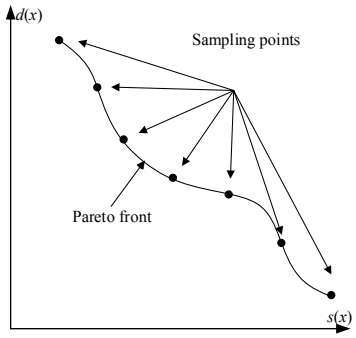


Figure 2: The illustration on the sampling points on the Pareto front

2.3 Procedure of the parallel sequential approximation optimization

Figure 3 presents a detailed flowchart of the PSAO algorithm, which is roughly divided into four blocks: the Initial stage, Approximation stage, Termination criteria and Sampling stage. Each block is elaborated below:

(1) Initial stage

In this stage, the m-dimensional design variable is scaled into an m-dimensional unit hypercube, which is then sampled by the Optimal Latin Hypercube Design (OLHD) method. The number of sampling points N is generally estimated from the following rule:

$$N = \begin{cases} 5m - 10m & m \le 10 \\ 100 & m > 10 \end{cases} \tag{6}$$

Finally, the objective function and the constraints of the sampling points are evaluated from the true model, and an initial sample set is generated.

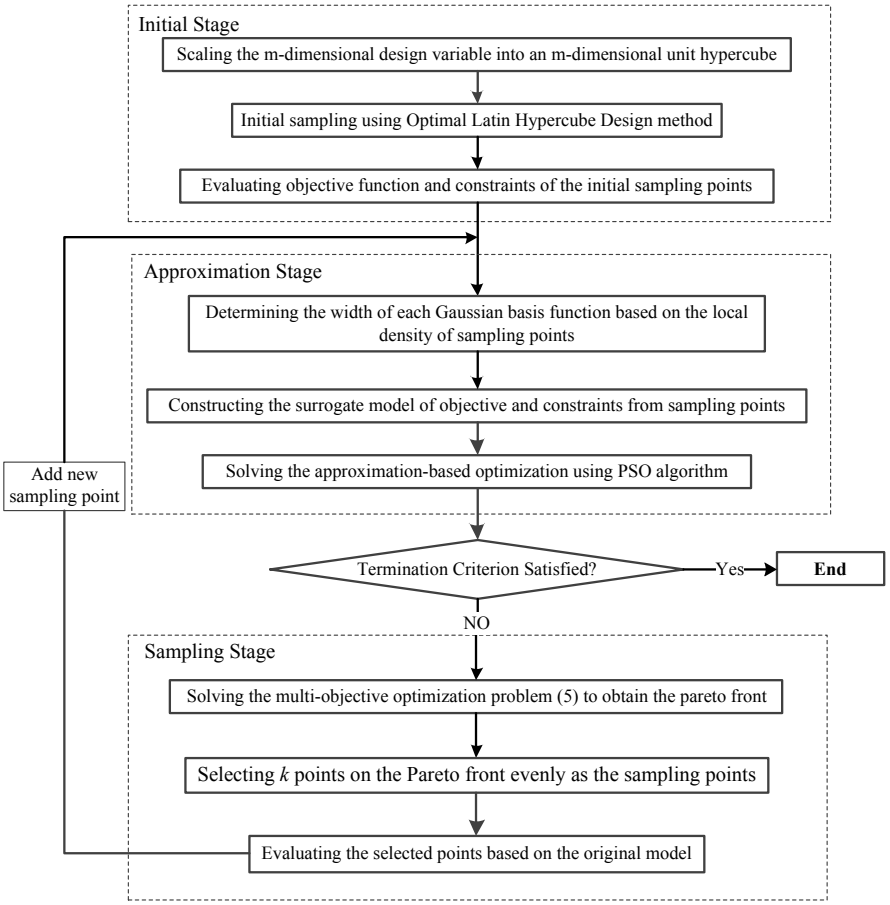


Figure 3: Procedure of the PSAO

(2) Approximation stage

Based on the sample set, a surrogate model of the objective function and its constraints is constructed using the enhanced approximate technique [23], whose kernel widths are obtained based on the local density of sampling points. The proposed point-density based estimate of the width of basis function demonstrates good performance for both uniform and non-uniform sampling points. Finally, the surrogate-based optimization problem is solved by the PSO algorithm [7].

(3) Termination criteria

The PSAO is terminated under the following criteria:

(i) If the relative distance between the optimal solutions of two successive iterations is below 1%, then evaluate criterion (ii). Otherwise, advance the SAO to the Sampling stage;

(ii) If the relative error between the optimal objective functions under the constraints imposed by a penalty is less than 1%, then evaluate criterion (iii). Otherwise, advance the PSAO to the Sampling stage;

(iii) If the relative error between the objective functions of the surrogate model and the true model in the current iteration is less than 1%, then convergence is reached and the PSAO algorithm is terminated. Otherwise, advance the PSAO to the Sampling stage.

(4) Parallel sampling stage

The detail of the parallel sampling strategy has been descripted in section 2.2. When sampling points are selected, the sampling points together with their responses based on the original model are added to the sample set to update the surrogate model in the next iteration.

3 STRUCTURAL DESIGN OPTIMIZATION CASE STUDIES

In this section, two case studies from simple to complex are taken from the engineering practices to investigate the general-purpose application and advantages of the proposed PSAO. The effectiveness and robustness of the proposed PSAO algorithm is validated in comparison with other optimization techniques in this section as well.

3.1 Test case 1

The first test case is a 72-member space truss with numerous design variables and constraints. The geometry and material properties, as well as the node and member numbering system, are shown in Figure 4. The optimization objective is to minimize the structural weight. The design variables are the cross-sectional areas of the truss members, grouped as shown in Figure 4. This grouping reduces the number of design variables to 16 member groups. The area is allowed to vary between 0.1 and 2.5 in2. The structure is subject to two loading conditions, as detailed in Table 2. The maximum allowable stress (tension or compression) is 25.0 ksi per member group, while the maximum allowable planar displacement of each node, in either the x or y direction, is ±0.25 in for both loading cases.

Table 1: 72-bar truss member area groups.

Area group	Truss members	Area group	Truss members
A1	1, 2, 3, 4	A9	37, 38, 39, 40
A2	5, 6, 7, 8, 9, 10, 11, 12	A10	41, 42, 43, 44, 45, 46, 47, 48
A3	13, 14, 15, 16	A11	49, 50, 51, 52
A4	17, 18	A12	53, 54
A5	19, 20, 21, 22	A13	55, 56, 57, 58
A6	23, 24, 25, 26, 27, 28, 29, 30	A14	59, 60, 61, 62, 63, 64, 65, 66
A7	31, 32, 33, 34	A15	67, 68, 69, 70
A8	35, 36	A16	71, 72

Table 2: 72-bar truss loading cases

Load case	Node	Fx [kips]	Fy [kips]	Fz [kips]
1	1	5.0	5.0	-5.0

2	1	0.0	0.0	-5.0
	2	0.0	0.0	-5.0 -5.0
	3	0.0	0.0	-5.0 -5.0
	4	0.0	0.0	-5.0 -5.0

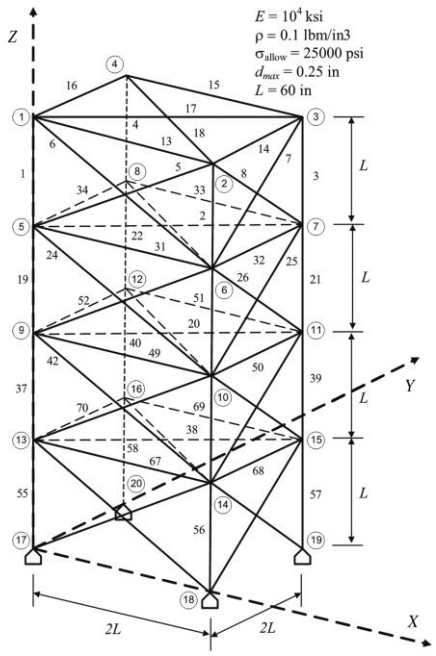


Figure 4: 72-Bar truss geometry

The optimization problem is solved by the proposed PSAO algorithm for 100 initial sampling points. In each iteration, 3 sampling points are selected in parallel. Figure 5 illustrates the relative distance and the relative error during the iteration history. The termination criterion is satisfied after 102 iterations and the evolution of the objective function is shown in Figure 6. The PSAO results are compared against those of recent publications; namely, the original SAO [9], penalty based PSO [24], and ant colony algorithms [25]. Table 3 summarizes the previously published results for the 72-bar truss problem using the different optimizers. Due to the parallel sampling strategy, the capacity of exploration and convergence of the original SAO is greatly enhanced. Hence, the number of iteration is reduced. It is noteworthy that the number of function evaluations for PSAO (406) is more than that for the original SAO (252). However, in PSAO, the sampling points are evaluated by the true model in parallel in each iteration, the actual computation time is equal to 202 function evaluations. Moreover, the SAO algorithm (PSAO and original SAO) shows a much better performance than EAs, and the number of function evaluations required to find the optimal solution is reduced from order 10^4 to order 10^2, indicating a substantial reduction in computing costs.

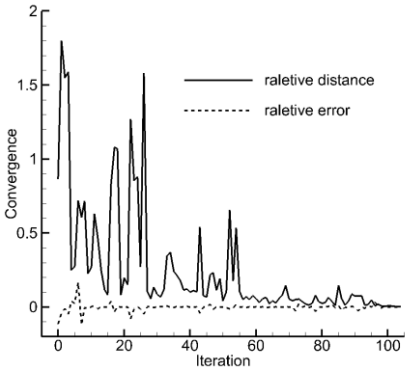


Figure 5: The best solution convergence history

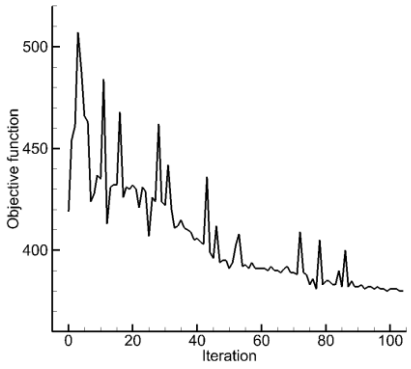


Figure 6: The objective function history

Table 3: Optimization results for the 72-bar truss

Design Variables	PSAO	SAO [9]	Perez and Behdinan(PSO) [24]	Camp and Bichon (ACO)[25]
A1 [in]	**0.157**	0.157	0.162	0.156
A2 [in]	**0.546**	0.549	0.509	0.550
A3 [in]	**0.405**	0.406	0.497	0.390
A4 [in]	**0.566**	0.555	0.562	0.592
A5 [in]	**0.520**	0.513	0.514	0.561
A6 [in]	**0.518**	0.529	0.546	0.492
A7 [in]	**0.100**	0.100	0.100	0.100
A8 [in]	**0.100**	0.100	0.110	0.107
A9 [in]	**1.258**	1.252	1.308	1.303
A10 [in]	**0.513**	0.524	0.519	0.511
A11 [in]	**0.100**	0.100	0.100	0.101
A12 [in]	**0.100**	0.100	0.100	0.100
A13 [in]	**1.898**	1.832	1.743	1.948
A14 [in]	**0.513**	0.512	0.519	0.508
A15 [in]	**0.100**	0.100	0.100	0.101
A16 [in]	**0.100**	0.100	0.100	0.102
Max. stress [psi]	**24999.67**	24943.87	24485.67	24939.59
Max. disp.[in]	**0.2500**	0.24992	0.2497	0.2500
Weight [lb]	**379.61**	379.90	381.91	380.24
Function	**202***	252	>20000	18500

3.2 Test case 2

This section presents the design of a moderate-dimensional case study of a bracket structure based on the proposed structural design framework. As depicted in Figure 7 eleven parameters that significantly affect the performance of the bracket are selected as design variables. Table 4 presents the feasible range of the design variables. The CAE model of the bracket is shown in Figure 8. Young's modulus and Poisson's ratio are settled as 200 GPa and 0.3. The bracket is

subjected to a stretching force and a bending moment induced by the forces P1=4.7 kN and P2 = 4.2 kN loaded at the center of such screws as L1, L2, L3 and L4. In addition, the bracket is fixed at the screws including R1, R2, R3 and R4. The optimization objective is to minimize the volume while the maximum stress is constrained to be less than 200 MPa and the total displacement less than 2cm.

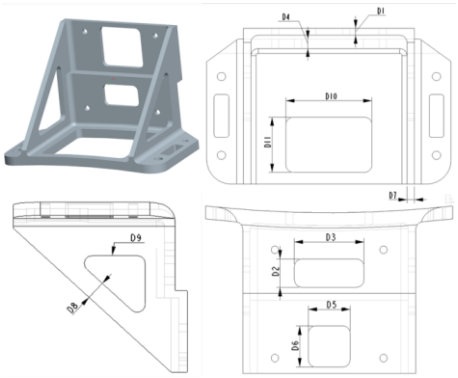


Figure 7: Description of the selected design variables

Table 4: Design space

Design variable (cm)	Lower bound	Upper bound
D1	1.0	3.5
D2	12.0	40.0
D3	10.0	25.0
D4	2.0	6.0
D5	12.0	40.0
D6	10.0	20.0
D7	1.5	4.0
D8	5.0	13.0
D9	25.0	35.0
D10	35.0	55.0
D11	8.0	50.0

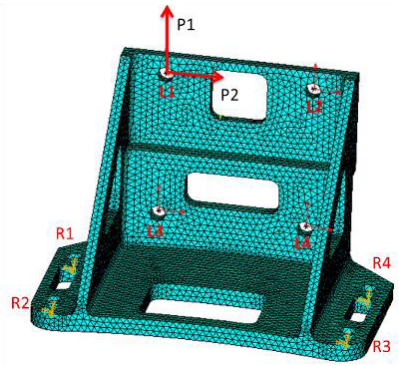


Figure 8: Details of the CAE model

The proposed PASO algorithm is used to solve this optimization problem and the number of initial sampling points is settled as 100. In every iteration, 4 sampling points are selected in parallel. In this test case, the proposed approach is also benchmarked against other optimization techniques as shown in Table 5. The best result from the zero-order optimization method locates on the point of 34675.67 cm^3 after 186 function evaluations. Pure GA with a population of 50 individuals and PSO with a swarm of 40 individuals are also used to solve the optimization problem, and the optimal volumes of 31493.33 cm^3 and 31510.26 cm^3 are found after 120 iterations (6000 FEA evaluations) and 140 iterations (5600 FEA evaluations), respectively. The original SAO requires 196 iterations to find the same level of optimal result, i.e. 296 FEA evaluations in sum [23]. Comparatively, the proposed PASO runs 152 iterations to get the same result as shown in Figure 9 and Figure 10. The total number of FEA iterations is 708. Because of the parallel mechanism, the equivalent number of function evaluations is 252. Based on the proposed approach, the computing costs of the structural optimization are immensely reduced.

Table 5: Comparison of the design optimization results

Optimization method	Volume (cm^3)	Displacement (cm)	Stress (MPa)	FEA evaluations
CAE built-in optimization	34675.67	1.86	196.5	186
GA	31493.33	1.94	198.3	6000
PSO	31510.26	1.96	195.4	5600
SAO	31481.48	1.96	197.6	296
PSAO	**31502.22**	**1.96**	**196.5**	**252***

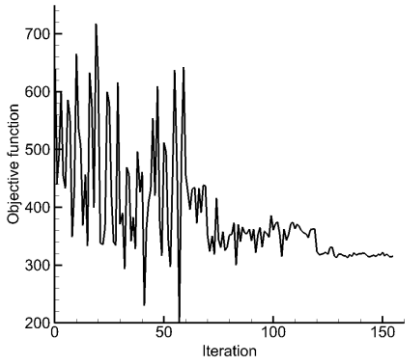


Figure 9: The convergence history

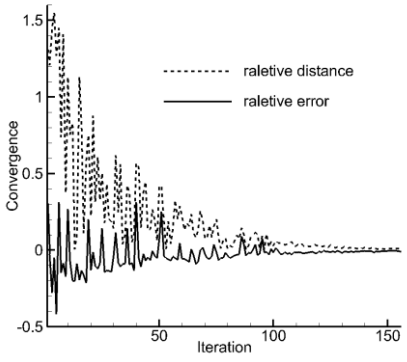


Figure 10: The iteration history of the objective function

4 CONCLUSIONS

This research has presented a parallel sequential approximate optimization (PSAO) algorithm that is suitable for structural design optimization tasks. This approach intends to reduce the computational costs normally associated with structural design problems.

We introduce a parallel sampling strategy that balances exploration and exploitation, allowing high-efficiency searching of the global optimum during the optimization process. The

parallel sampling strategy substantially reduces the computation costs required to find the optimal solutions. The feasibility, convenience and efficacy of the proposed structural design optimization algorithm have been investigated through two cased studies. Furthermore, the effectiveness and computational efficiency of the PSAO methodology are benchmarked against other optimization approaches as well. The time consumed to find the optimum is reduced from the order of 10^3 to 10^2 by the PSAO. Accordingly, the proposed PSAO proves to be an adequate strategy for effectively and efficiently handling structural design optimization problems.

REFERENCES

[1] Yildiz, A.R. Comparison of evolutionary-based optimization algorithms for structural design optimization. *Eng. Appl. Artif. Intell* (2013) **26**: 327-333.

[2] Bonte, M.H.A., Boogaard, A.H., and Huétink J. An optimization strategy for industrial metal forming processes. S*truct. Multidiscip. Optim* (2008) **35**: 571-586.

[3] Edke, M.S. and Chang, K.H. Shape optimization of heavy load carrying components for structural performance and manufacturing cost. *Struct. Multidiscip. Optim* (2006) **31**: 344-354.

[4] Dai, L., Guan, Z.Q., Chen, B.S. and Zhang, H.W. An open platform of shape design optimization for shell structure. *Struct. Multidiscip. Optim* (2008) **35**: 609-622.

[5] Goldberg, D. *Genetic algorithms in search, optimization, and machine learning*. New York, NY: Addison-Wesley, (1989).

[6] Kirkpatrick, S., Gelatt, C., and Vecchi, M. Optimization by simulated annealing. *Science* (1983) **220**: 671-680

[7] Kennedy, J and Eberhart, R. *Particle swarm optimization*. Piscataway: IEEE international conference on neural networks, (1995)..

[8] Karaboga, D and Basturk, B. A powerful and efficient algorithm for numerical function optimization: artificial bee colony (ABC) algorithm. *J. Global. Optim* (2003) **39**: 459-471.

[9] Wang, D., Wu, Z., Fei, Y. and Zhang, W. Structural design employing a sequential approximation optimization approach. *Comput. Struct* (2014) **134**: 75-87.

[10] Barthelemy, J.F.M. and Haftka, R.T. Approximation concepts for optimum structural design-a review. *Struct. Optim* (1993) **5**: 129-144.

[11] Haftka, R.T and Scott, E.P. Optimization and Experiments-A Survey. *Theor. Appl. Mech* (1996) **25**: 303-321.

[12] Nikolaidis, E., Long, L. and Ling, Q. *Neural Networks and Response surface polynomials for Design of Vehicle Joints*. AIAA-98-1777, (1998).

[13] Berke, L. and Hajela, P. Application of artificial neural nets in structural mechanics, *Struct. Optim* (1992) **4**: 90-98.

[14] Giunta, A., Watson, L.T. and Koehler, J. *A comparison of approximation modeling techniques: polynomial versus interpolating models*. St.Louis: 7th AIAA/USAF/NASA/ ISSMO Symposium on Multidisciplinary Analysis & Optimization, (1998).

[15] Simpson, T.W., Mauery, T.M., Korte, J.J. and Mistree, F. *Comparison of response surface and Kriging models for multidisciplinary design optimization.* St. Louis: 7th AIAA/ USAF/NASA/ISSMO Symposium on Multidisciplinary Analysis & Optimization, (1998).

[16] Broomhead, D.S. and Loewe, D. Multivariate functional interpolation and adaptive networks. *Adv. Complex. Syst* (1988) **2**: 321-355.

[17] Bonte, M.H.A. and Fourmentm, L. Optimization of forging processes using finite element simulations - a comparison of sequential approximate optimization and other algorithm. *Struct. Multidiscip. Optim* (2010) **42**: 797-810.
[18] Cheng, G., Xu, L. and Jiang. L. A sequential approximate programming strategy for reliability-based structural optimization. *Comput. Struct* (2006) **84**: 1353-1367.
[19] Kitayama, S., Arakawa, M. and Yamazaki, K. Sequential approximate optimization using radial basis function network for engineering optimization. *Optim. Eng* (2011) **12**: 535-557.
[20] Forrester, A.I.J and Keane, A.J. Recent advances in surrogate-based optimization. *Prog. Aerosp. Sci* (2009) **45**: 50-79.
[21] Hastie, T., Tibshirani, R. and Friedman, J. *The elements of statistical learning*. Springer, Vol. I, (2001).
[22] Jones, D.R. A taxonomy of global optimization methods based on response surfaces. *J. Glob. Optim* (2001) **21**: 345-383.
[23] Wang, D., Hu, F., Ma, Z., Wu, Z. and Zhang, W. A CAD/CAE integrated framework for structural design optimization using sequential approximation optimization. *Adv. Eng. Softw* (2014) **44**:7656-7668.
[24] Perez, R. and Behdinan, K. Particle swarm approach for structural design. *Comput. Struct* (2007) **85**:1579-1588.
[25] Camp, C. and Bichon, J. Design of space trusses using ant colony optimization. *J. Struct. Eng* (2004) **130**: 741-751.

3D-SIMULATION OF ELECTROMAGNETIC AND TEMPERATURE FIELDS IN THE CONTINUOUS INDUCTION HEATERS

VICTOR B. DEMIDOVICH*, FEDOR V. CHMILENKO*, VLADISLAV V. ANDRUSHKEVICH* AND IRINA I. RASTVOROVA*

*Sankt-Petersburg Electrotechnical University (LETI)
5 Prof.Popov str, St.Petersburg, 197376, Russian Federatio
e-mail: vbdemidovich@mail.ru http://www.eltech.ru

Key words: Electromagnetic and Temperature Fields, Induction Heating Applications, Computing Methods, FDM, FEM and IEM.

Abstract. The quasi-3-D model was developed for continuous induction heating of billets with arbitrary cross section. This computer model is intended for evaluation of electrical and thermal both stationary and non-stationary processes of stage and continuous working regime an induction heater with magnetic and non-magnetic loading with any of cross section shape. The combination of the most effective numerical methods for modeling induction heating process was used in this software: Finite Difference Method (FDM), Finite Element Method (FEM), Boundary Element Method (BEM), Integral Equation Method (IEM) and their combination.

1 INTRODUCTION

One of the basic areas, where the induction heating is widely applied, is the through induction heating of metals before plastic deformation, forging, rolling, pressing, etc. The shape of the cross section of the heating billets can be relatively simple, as for cylindrical or rectangular billets, and more complex, as for example, for flange beams. The purpose of heating is usually achievement of the given temperature with the certain allowable non-uniformity in volume of the product. For the complete description of the temperature field in a heated up body it is necessary to know as temperature distribution in the cross section, as along the length of the billet. Usually continuous in-line induction heaters are used for heating of metal with high production rate. The power of these installations can reach 10-15 MW and accurate predictions of the temperature field's distributions in the billets are very important at the design stage of installation as well for the control of the induction heaters.

The development of the theory of induction heating is indissolubly connected to development and application of modern methods of mathematical modeling. Obtaining results with sufficient accuracy, describing parameters of induction system - one of the important ways to increase productivity, improvement of constructional and working parameters of the equipment of induction systems. The sharp growth of productivity of computer facilities gives an opportunity to use in mathematical models the most effective numerical methods, such as Finite Difference Method (FDM), Finite Element Method (FEM), Boundary Element Method

(BEM), Integral Equation Method (IEM) and their combination. Now these methods are widely used at modeling electromagnetic and thermal processes.

Analysis has shown that the greatest dissemination has received so-called electrothermal models, which mirrored most essential features of the induction heating - coupling of electromagnetic and thermal fields. 2-D electrothermal models for calculation of the induction heating of cylindrical and plane bodies have designed at the end of 70th. Because of considerable time costs on preparing and input initial data, computation time and verification of the results, total cost of the design work was very high.

Essentially the situation has changed with appearance of PCs. Their availability has reduced in necessity of development models with the "friendly" interface, which would allow the user, unfamiliar with the numerical methods, only on the basis of physical phenomena of the task easy to input data and receive appropriate results of computations. The work in this direction was carried out in the St.Petersburg Electrotechnical University (LETI). As the result of such work the complex of the programs has been designed [1,2], which are widely used by leading worldwide organizations, which design and manufacture induction equipment.

Increasing of the production capacities, requirements to production rate of the induction equipment and quality of the heated body has give rise to increasing of the requirements to the designing software. Especially it has a major value at designing of whole technological line, which can consist of several in-line inductors. One of the major factors in software designing is dimension of the solving task. In most cases it is enough to use 2-D setting of the task. For a fast estimation of thermal processes and deriving of integral parameters of the induction system it is possible to use even one-dimensional model, especially when the influence of the “temporal factor”, which usually very short, has vital importance. In this case one-dimensional model is predictable. For deriving the complete information about electromagnetic and thermal processes in induction system user should use 2-D or 3-D model. The most commercial packages, which inclusive the latest reaching in the field of mathematical simulation, are calculated for wide application in different areas of a science and engineering. Usage of non-highly tailored software encounters on serial of difficulties, concerned with complexity of the calculating model. User should has additional knowledge of the methodology of usage such software: defining geometry of the system, boundary conditions, applied sources, solvers, viewing results of computation. Chances of errors appearance at data input, which can be invisible for the user, in this case are relatively high. Total time of computation may be very form several hours to several days, even using modern high-end computers. Software should include different Databases of the material properties, which usually nonlinear, especially depending magnetic permeability on both temperature and intensity of the magnetic field. Computation should inform user about computation process and main characteristics of the induction system, give the possibility to have an influence on computation process.

The advantages of using this software in comparison with some commercial packages:

- Software is designed especially for induction heating applications and it realizes essential coupling of electromagnetic and temperature fields during computations that provides great accuracy
- Most important features of the induction heating installations are taken into account, including characteristics of power sources and possibility to change power, voltage,

current and frequency during the process as well as possibility to simulate stages of holding billets in the inductor and its transport.

- Developer's knowledge of the induction heating theory is included in the software that allows, for example, to eliminate the problem of meshing and essentially to ease preparation of input data and to decrease time for it.
- User has possibility on the base of the visual information to interrupt a computation, to change data and to continue computations.
- Software can be used in the control systems in the real time.
- It is easy to implement procedures of optimal design and optimal control of the induction heating installations.
- Special database of materials and refractories that used in the induction heating installations is developed.

2 DESCRIPTION OF THE SOFTWARE

A high-efficient quasi-3D model was developed for induction heating of billets with arbitrary cross sections. Special algorithm was developed that comprises two 2D models. The first one allows receiving distribution of electromagnetic and temperature fields in the each cross-section of the billet along the length. This model is based on the FEM in the Ritz approach. The second model allows predicting distribution of the electromagnetic field along the length of the billets. Each model can be used separately. At the same time the method of boundary impedance conditions is used for coupling these two models into one quasi-3D model.

This software is intended for electrical and thermal computation both stationary and non-stationary processes of periodic and continuous working regime an induction heater with magnetic and not magnetic loading with any of cross section shape. The calculation of induction heaters with multisectional and polyphase coils, with usual and autotransformer closing is provided. The program includes solution of the exterior electrical and internal electrothermal task. As a result of exterior electrical calculation at a given supply voltage the total true and reactive power, true power in load, electrical losses in coils, current of the coils, electrical efficiency of the coils and distribution of the intensity of magnetic field on surface of the load are defined. The internal electrothermal calculation gives distribution of the temperature in load at particular boundary conditions on a surface of a load, which are set, either from condition of free heat exchange with an environment, or in view of a refractory. Common thermal losses and temperature distribution on internal surface of the refractory simultaneously calculated. The external task is based on the basis of a boundary integral method and the computation of the distribution of an electromagnetic and temperature field in the workpiece (internal task) is made on the basis of a finite elements method. Pasting of external and internal electromagnetic tasks is carried out on elements of a surface of the workpiece by statement of impedance boundary conditions:

$$z_{0i}=\dot{E}_i / \dot{H}_{ti}=\frac{\rho_i}{\delta_{ei}}\left(VR_i+jVX_i\right), i \in G \tag{1}$$

where ρ_i - is specific resistance, δ_{ei} - penetration depth, G - surface perimeter of the workpiece.

The values of impedances z_0 or factors VX_i and VR_i can be found directly from the solving of an internal task.

The process of induction heating of conductive bodies in two-dimensional setting is reduced to solution of a quasi-stationary nonlinear differential partial equation concerning a magnetic intensity

$$\frac{\partial}{\partial x}\left(\rho\frac{\partial \dot{H}}{\partial x}\right)+\frac{\partial}{\partial y}\left(\rho\frac{\partial \dot{H}}{\partial y}\right)=j\omega\mu\mu_0\dot{H} \tag{2}$$

And non-stationary heat conduction equation:

$$C_v\frac{\partial T}{\partial t}=\frac{\partial}{\partial x}\left(\lambda\frac{\partial T}{\partial x}\right)+\frac{\partial}{\partial y}\left(\lambda\frac{\partial T}{\partial y}\right)+w \tag{3}$$

Due to symmetry of distribution of an electromagnetic and temperature field in cross section, only a quarter of the load is considered. The boundary conditions then record as follows:

$$\left.\begin{aligned}
&\left.\frac{\partial \dot{H}}{\partial y}\right|_{y=0}=\left.\frac{\partial \dot{H}}{\partial x}\right|_{x=0}=0;\\
&\left.\dot{H}\right|_{surface}=\left.\dot{H}\right|_{surface}=\dot{H}_m(x,y);\\
&\left.\frac{\partial T}{\partial y}\right|_{y=0}=\left.\frac{\partial T}{\partial x}\right|_{x=0}=0;\\
&-\lambda\left.\frac{\partial t}{\partial x}\right|_{surface}=-\lambda\left.\frac{\partial T}{\partial y}\right|_{surface}=\Delta p_0(T).
\end{aligned}\right\} \tag{4}$$

The specific power of thermal losses is defined as follows:

$$\Delta p_0=\varepsilon\sigma\left(T^4-T_e^4\right)+\alpha(T-T_e), \tag{5}$$

where ε - emissivity, σ - Stefan-Bolzman constant, α - heat transfer coefficient, T_e - temperature of an environment; here T and T_e - in degrees of Kelvin.

The internal electrothermal task is solved by a method of finite element method. Usage this method gives a possibility to define fields distribution in arbitrary cross section. For

solution of the thermal task with the help of FEM, it is necessary to use a functional, bound with the equation (3):

$$f = \int_V \frac{1}{2}\left[\lambda\left(\frac{\partial T}{\partial x}\right)^2 + \lambda\left(\frac{\partial T}{\partial y}\right)^2 - 2\left(w - C_v \frac{\partial T}{\partial t}\right)T \right] dV + \int_S \Delta p_0 T dS \tag{6}$$

For solution of the electrical task the functional for (2) was obtained:

$$f = \frac{1}{2}\int_V \left[\rho\left(\frac{\partial H}{\partial x}\right)^2 + \rho\left(\frac{\partial H}{\partial y}\right)^2 - 2 j\omega\mu H^2 \right] dV \tag{7}$$

Analysis of electromagnetic and temperature distributions in the billets with complicated cross-sections are given in the paper.

3 CASE STUDY. CONTINUOUS INDUCTION HEATING OF THE SLABS WITH ROUNDING CORNERS

The following example illustrates features and abilities of the software UNIVERSAL3D.

Titanium slabs with rounding corners go through line of three inductors (Figure1). Speed of the movement is 4 cm/sec. Length of each inductor is 50 cm with distance of 30 cm between them. The inductors have chamotte refractory with thickness 1 cm. Frequency of the current is 8000 Hz. Coils are connected in parallel and the output voltage of power source is constant during heating - 800 V. The initial temperature of the load is 20 °C.

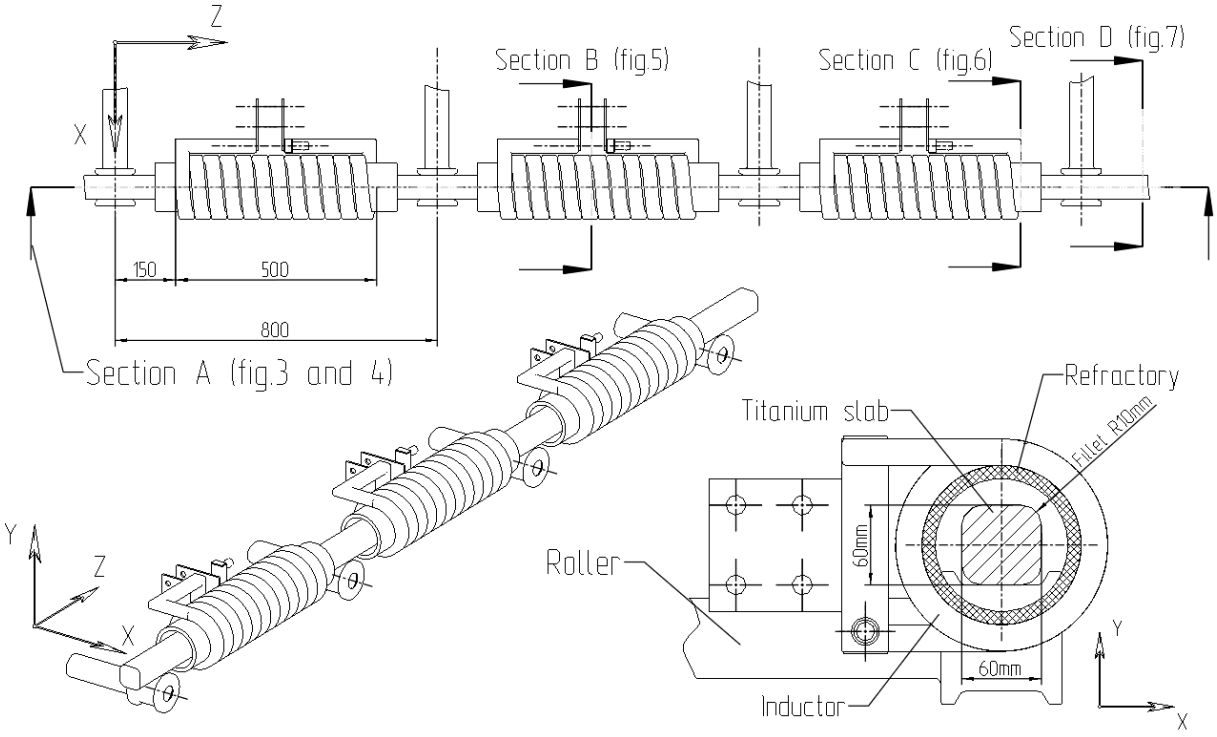


Figure 1. Sketch of the induction heating system

Figure 2. Window UNIVERSAL3D

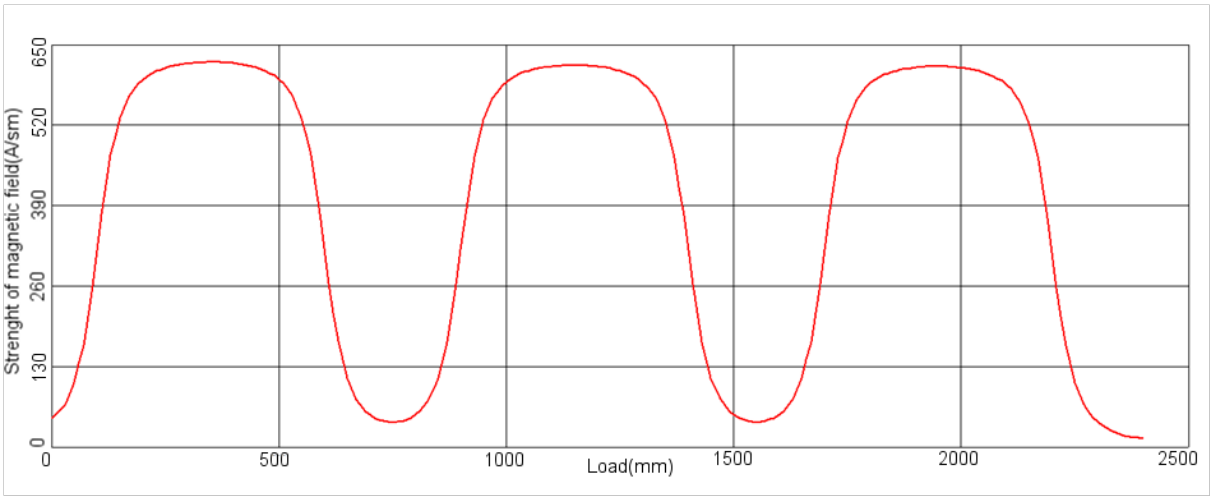


Figure 3. Strength of magnetic field at the surface of the load

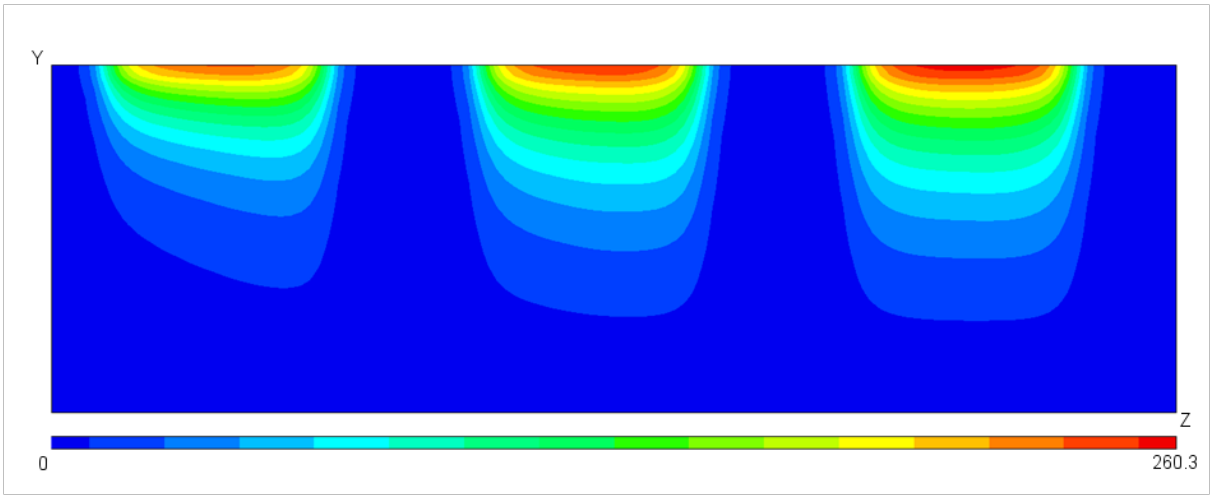


Figure 4. Distribution of the heat sources in the 1/2 section A of the load

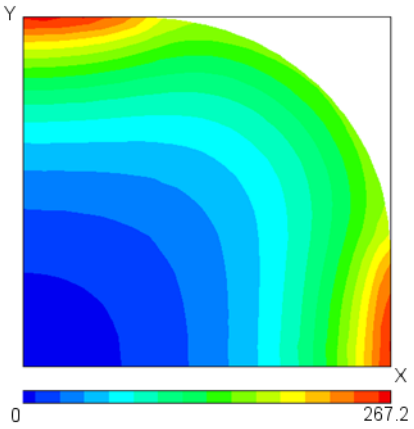


Figure 5. Distribution of the heat sources in the 1/4 section B of the load

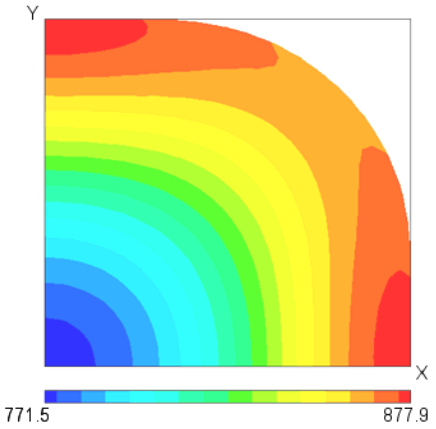


Figure 6. Distribution of the temperature field in the 1/4 section C of the load

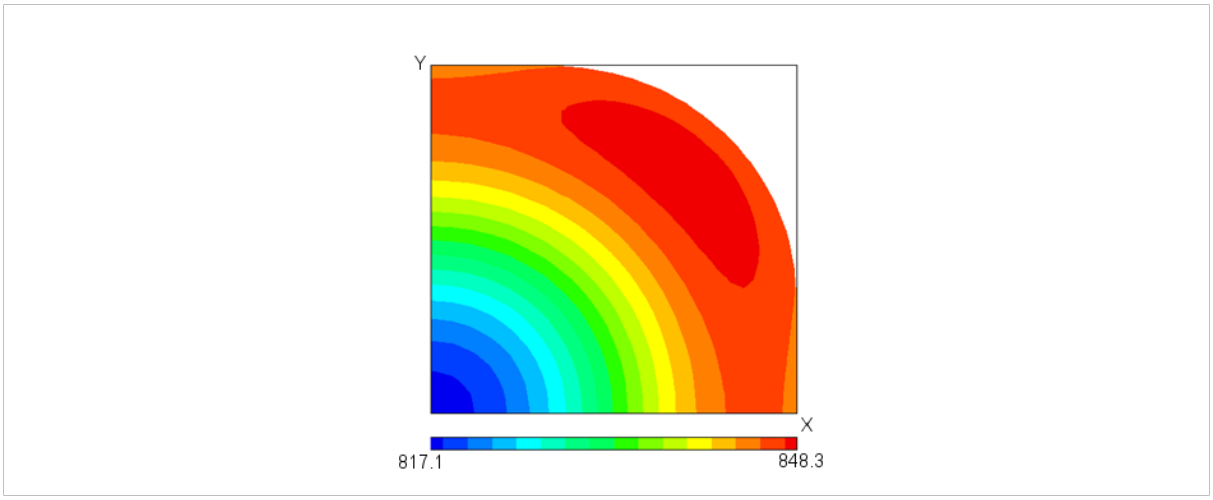


Figure 7. Distribution of the temperature field in the 1/4 section D of the load

ACKNOWLEDGEMENTS

This work has been supported by Ministry of Education and Science (project №2014/187)

REFERENCES

[1] Demidovitch V. (1996) Special Software IndHeat for Modelling Induction Heating Processes, *Proceedings of 3rd International Workshop on ELECTRIC AND MAGNETIC FIELDS EMF'96*, Liege, Belgium, pp. 273-278.

[2] Demidovich V. (2012) Computer simulation and optimal designing of energy-saving technologies of the induction heating of metals, Thermal Engineering. V. 59. Issue 14, pp. 1023–1034.

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis (Eds)

AN ELECTROELASTIC PROBLEM OF A SEMI-INFINITE BODY WITH D_∞ SYMMETRY SUBJECTED TO DISTRIBUTED SURFACE LOADING

MASAYUKI ISHIHARA*, YOSHIHIRO OOTAO* AND YOSHITAKA KAMEO*

* School of Engineering, Osaka Prefecture University
1-1 Gakuen-cho, Naka-ku, Sakai-shi, Osaka 599-8531, Japan
e-mail: ishihara@me.osakafu-u.ac.jp

Key words: Electroelastic Problem, Semi-infinite Body, Point Group D_∞.

Abstract. Electroelastic field in an semi-infinite body with D_∞ symmetry subjected to a locally uniform electric potential on its surface is investigated. By extending a potential function method for transversely isotropic bodies, the electroelastic field inside the body is formulated. Furthermore, numerical calculation is performed to investigate the field qualitatively and quantitatively.

1 INTRODUCTION

The concepts of carbon neutrality have attracted considerable attention recently because of an increasing demand for a reduction in environmental loads. From the viewpoint of engineering production, wooden materials are one of the most promising candidates for achieving carbon neutrality.

To ensure the quality of wooden materials, nondestructive evaluation techniques need to be developed. In particular, the detection of local defects such as cracks, knots, and pith are of great importance for ensuring structural integrity. Wood has been known as a piezoelectric material since the middle of the 20th century, when Fukada succeeded in experimentally verifying the direct and converse piezoelectric effects of wood[1]. These effects are expected to be employed for nondestructive evaluation techniques[2-4].

From a *mesoscopic* viewpoint, woods are considered to belongs to point group D_∞ [1], which is characterized by an ∞*-fold rotation axis* and a *two-fold rotation axis* perpendicular to it[5]. The nonzero components of the piezoelectric constant are d_{14} and $d_{25}(=-d_{14})$ only, allowing the ∞ -fold rotation axis be the third axis. In that case, the electric field perpendicular to the ∞ -fold rotation axis (third axis) induces shear strain in the plane perpendicular to the direction of the electric field.

The elastic problems of transversely isotropic bodies, which correspond to a special case in the absence of the piezoelectric effects in body with D_∞ symmetry, were extensively analyzed[6—8]. On the other hand, electroelastic problems of bodies with D_∞ symmetry were investigated experimentally[9-12]. However, for sound operation of nondestructive evaluation

techniques, not only the input/output relationship but also the electroelastic field *inside* the material must be elucidated.

In this paper, therefore, we analyze the electroelastic field in a body with D_∞ symmetry. As an example, we treat a semi-infinite body subjected to a locally uniform electric potential on its surface. First, the displacement and electric field are expressed in terms of the potential functions. The governing equations for these functions are obtained by the equilibrium equations of stresses and the Gauss law. By solving the governing equations, the electroelastic field quantities are formulated. Moreover, by performing numerical calculation, the stress and electric field are investigated qualitatively and quantitatively, which helps us to understand the electroelascit field inside a body with D_∞ symmetry.

2 THEORETICAL ANALYSIS

2.1 Problem

We consider a semi-infinite piezoelectric body belonging to point group D_∞, as shown in Fig. 1, where the z axis is parallel to the ∞-fold rotation axis of the body. The surface of the body is subjected to a locally uniform distribution of electric potential and free from traction. The displacements and electric potential are assumed to be zero at infinity. Thus, the boundary conditions are given as

$$\left.\begin{array}{ll} x=0: & \sigma_{xx}=0,\sigma_{zx}=0,\sigma_{xy}=0,\Phi=\Phi_0\cdot H\left(|\delta|-y\right)H\left(|\delta|-z\right); \\ \sqrt{x^2+y^2+z^2}\to\infty: & u_x\to 0,u_y\to 0,u_z\to 0,\Phi\to 0 \end{array}\right\}. \tag{1}$$

In Eq. (1), σ_{ij}, u_i, and Φ denote the stress, displacement, and electric potential, respectively, $H(\cdot)$ denotes the Heaviside step function, and δ denotes the half-length of the square where the uniform electric potential is applied.

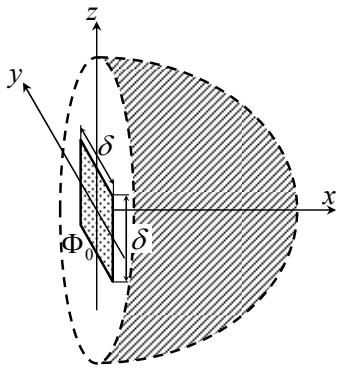


Figure 1: Analytical model

2.2 Governing equations

Let ε_{ij}, E_i, and D_i $(i,j=x,y)$ be the the strain, electric field, and electric displacement, respectively. The constitutive equations of the body are given as

$$\begin{Bmatrix}\sigma_{xx}\\ \sigma_{yy}\\ \sigma_{zz}\\ \sigma_{yz}\\ \sigma_{zx}\\ \sigma_{xy}\end{Bmatrix}=\begin{bmatrix} c_{11} & c_{12} & c_{13} & 0 & 0 & 0\\ & c_{11} & c_{13} & 0 & 0 & 0\\ & & c_{33} & 0 & 0 & 0\\ & & & c_{44} & 0 & 0\\ & & & & c_{44} & 0\\ \text{sym.} & & & & & \frac{c_{11}-c_{12}}{2}\end{bmatrix}\begin{Bmatrix}\varepsilon_{xx}\\ \varepsilon_{yy}\\ \varepsilon_{zz}\\ 2\varepsilon_{yz}\\ 2\varepsilon_{zx}\\ 2\varepsilon_{xy}\end{Bmatrix}-\begin{bmatrix} 0 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\\ e_{14} & 0 & 0\\ 0 & -e_{14} & 0\\ 0 & 0 & 0\end{bmatrix}\begin{Bmatrix}E_x\\ E_y\\ E_z\end{Bmatrix}, \tag{2}$$

$$\begin{Bmatrix}D_x\\ D_y\\ D_z\end{Bmatrix}=\begin{bmatrix} 0 & 0 & 0 & e_{14} & 0 & 0\\ 0 & 0 & 0 & 0 & -e_{14} & 0\\ 0 & 0 & 0 & 0 & 0 & 0\end{bmatrix}\begin{Bmatrix}\varepsilon_{xx}\\ \varepsilon_{yy}\\ \varepsilon_{zz}\\ 2\varepsilon_{yz}\\ 2\varepsilon_{zx}\\ 2\varepsilon_{xy}\end{Bmatrix}+\begin{bmatrix} \eta_{11} & 0 & 0\\ & \eta_{11} & 0\\ \text{sym.} & & \eta_{33}\end{bmatrix}\begin{Bmatrix}E_x\\ E_y\\ E_z\end{Bmatrix}, \tag{3}$$

where c_{ij} , η_{kl} , and e_{kj} denote the elastic stiffness constant, dielectric constant, and piezoelectric constant, respectively. The displacement-strain relations are given as

$$\varepsilon_{xx}=\frac{\partial u_x}{\partial x},\ \varepsilon_{yy}=\frac{\partial u_y}{\partial y},\ \varepsilon_{zz}=\frac{\partial u_z}{\partial z},\ 2\varepsilon_{yz}=\frac{\partial u_y}{\partial z}+\frac{\partial u_z}{\partial y},\ 2\varepsilon_{zx}=\frac{\partial u_z}{\partial x}+\frac{\partial u_x}{\partial z},\ 2\varepsilon_{xy}=\frac{\partial u_x}{\partial y}+\frac{\partial u_y}{\partial x}. \tag{4}$$

The equilibrium equations of stresses and the Gauss law are given, respectively, by

$$\frac{\partial \sigma_{xx}}{\partial x}+\frac{\partial \sigma_{xy}}{\partial y}+\frac{\partial \sigma_{zx}}{\partial z}=0,\quad \frac{\partial \sigma_{yy}}{\partial y}+\frac{\partial \sigma_{yz}}{\partial z}+\frac{\partial \sigma_{xy}}{\partial x}=0,\quad \frac{\partial \sigma_{zz}}{\partial z}+\frac{\partial \sigma_{zx}}{\partial x}+\frac{\partial \sigma_{yz}}{\partial y}=0, \tag{5}$$

$$\frac{\partial D_x}{\partial x}+\frac{\partial D_y}{\partial y}+\frac{\partial D_z}{\partial z}=0. \tag{6}$$

Referring to a solution technique for elastic problems of transversely isotropic bodies[6], the displacement potential functions φ_i and ϑ_i are introduced as

$$u_x=\sum_{i=1}^{2}\left(\frac{\partial \varphi_i}{\partial x}+\frac{\partial \vartheta_i}{\partial y}\right),\quad u_y=\sum_{i=1}^{2}\left(\frac{\partial \varphi_i}{\partial y}-\frac{\partial \vartheta_i}{\partial x}\right),\quad u_z=\frac{\partial}{\partial z}\sum_{i=1}^{2}k_i\varphi_i, \tag{7}$$

where

$$k_i=\frac{c_{11}\mu_i-c_{44}}{c_{13}+c_{44}}, \tag{8}$$

μ_1 and μ_2 are the roots of a quadratic equation for μ :

$$c_{11}c_{44}\mu^2-\left(c_{11}c_{33}-c_{13}^2-2c_{13}c_{44}\right)\mu+c_{33}c_{44}=0. \tag{9}$$

The components of the electric field are expressed by the electric potential function as

$$E_x = -\frac{\partial \Phi}{\partial x}, \quad E_y = -\frac{\partial \Phi}{\partial y}, \quad E_z = -\frac{\partial \Phi}{\partial z}. \tag{10}$$

Substituting Eqs. (4), (7), and (10) into Eqs. (2) and (3) and the results into Eqs. (5) and (6), we have

$$\left(\Delta_{\mathrm{p}} + \mu_i \frac{\partial^2}{\partial z^2}\right)\varphi_i = 0, \quad \left(\Delta_{\mathrm{p}} + \upsilon_i \frac{\partial^2}{\partial z^2}\right)\vartheta_i = 0 \quad \frac{e_{14}\mu_3}{c_{44}}\frac{\partial \Phi}{\partial z} = \sum_{i=1}^{2}\left(\Delta_{\mathrm{p}} + \mu_3 \frac{\partial^2}{\partial z^2}\right)\vartheta_i, \tag{11}$$

where υ_1 and υ_2 are the roots of a quadratic equation with respect to υ:

$$\upsilon^2 - \left[\mu_3\left(1 + k_{\mathrm{couple}}^2\right) + \eta\right]\upsilon + \mu_3\eta = 0 \tag{12}$$

and

$$\mu_3 = \frac{2c_{44}}{c_{11} - c_{12}}, \quad \eta = \frac{\eta_{33}}{\eta_{11}}, \quad k_{\mathrm{couple}}^2 = \frac{e_{14}^2}{c_{44}\eta_{11}}, \quad \Delta_{\mathrm{p}} \equiv \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}. \tag{13}$$

2.3 Electroelastic field quantities

By considering the symmetry of the electroelastic field and Eq. (1) and applying the Fourier transform techniques[13] to Eq. (11), the solutions to Eq. (11) are obtained as

$$\left.\begin{aligned} \varphi_i &= \int_0^\infty \int_0^\infty A_i(\alpha,\beta)\exp\left(-\gamma_{\mu i}x\right)\sin(\alpha y)\sin(\beta z)\mathrm{d}\alpha\mathrm{d}\beta, \\ \vartheta_i &= \int_0^\infty \int_0^\infty C_i(\alpha,\beta)\exp\left(-\gamma_{\upsilon i}x\right)\cos(\alpha y)\sin(\beta z)\mathrm{d}\alpha\mathrm{d}\beta, \\ \Phi &= \frac{c_{44}}{e_{14}}\frac{1}{\mu_3}\sum_{i=1}^{2}\left[\int_0^\infty \int_0^\infty (\mu_3 - \upsilon_i)\beta C_i(\alpha,\beta)\exp\left(-\gamma_{\upsilon i}x\right)\cos(\alpha y)\cos(\beta z)\mathrm{d}\alpha\mathrm{d}\beta\right] \end{aligned}\right\}, \tag{14}$$

where

$$\gamma_{\mu i} = \sqrt{\alpha^2 + \mu_i\beta^2}, \quad \gamma_{\upsilon i} = \sqrt{\alpha^2 + \upsilon_i\beta^2}; \tag{15}$$

$A_i(\alpha,\beta)$ and $C_i(\alpha,\beta)$ $(i = 1,2)$ are unknown constants to be determined by the boundary conditions described by Eq. (1). The distribution function for surface electric potential is expressed in the Fourier integral form[13] as

$$\left.\begin{aligned} H(|\delta| - y)H(|\delta| - z) &= \int_0^\infty \int_0^\infty f^*(\alpha,\beta)\cos(\alpha y)\cos(\beta z)\mathrm{d}\alpha\mathrm{d}\beta, \\ f^*(\alpha,\beta) &= \frac{4\delta^2}{\pi^2}\frac{\sin(\alpha\delta)}{\alpha\delta}\frac{\sin(\beta\delta)}{\beta\delta} \end{aligned}\right\}. \tag{16}$$

By substituting Eq. (14) into Eqs. (2)－(4), (7), and (10), the electroelastic field quantities are formulated, for example, as

$$
\left.\begin{aligned}
\sigma_{xx} &= \frac{c_{44}}{\mu_3}\sum_{i=1}^{2}\left\{\int_0^\infty\int_0^\infty\begin{bmatrix}\left(2\alpha^2+(1+k_i)\mu_3\beta^2\right)A_i(\alpha,\beta)\exp(-\gamma_{\mu i}x)\\ +2\gamma_{\upsilon i}\alpha C_i(\alpha,\beta)\exp(-\gamma_{\upsilon i}x)\end{bmatrix}\right.\right.\\
&\qquad\left.\left.\cdot\sin(\alpha y)\sin(\beta z)\mathrm{d}\alpha\mathrm{d}\beta\right\},\\
\sigma_{yz} &= \frac{c_{44}}{\mu_3}\sum_{i=1}^{2}\left\{\int_0^\infty\int_0^\infty\begin{bmatrix}\mu_3(1+k_i)\alpha\beta A_i(\alpha,\beta)\exp(-\gamma_{\mu i}x)\\ +\upsilon_i\gamma_{\upsilon i}\beta C_i(\alpha,\beta)\exp(-\gamma_{\upsilon i}x)\end{bmatrix}\cos(\alpha y)\cos(\beta z)\mathrm{d}\alpha\mathrm{d}\beta\right\},\\
\sigma_{zx} &= \frac{c_{44}}{\mu_3}\sum_{i=1}^{2}\left\{\int_0^\infty\int_0^\infty\begin{bmatrix}-\mu_3(1+k_i)\gamma_{\mu i}\beta A_i(\alpha,\beta)\exp(-\gamma_{\mu i}x)\\ -\upsilon_i\alpha\beta C_i(\alpha,\beta)\exp(-\gamma_{\upsilon i}x)\end{bmatrix}\sin(\alpha y)\cos(\beta z)\mathrm{d}\alpha\mathrm{d}\beta\right\},\\
\sigma_{xy} &= \frac{c_{44}}{\mu_3}\sum_{i=1}^{2}\left\{\int_0^\infty\int_0^\infty\begin{bmatrix}-2\gamma_{\mu i}\alpha A_i(\alpha,\beta)\exp(-\gamma_{\mu i}x)\\ -\left(2\alpha^2+\upsilon_i\beta^2\right)C_i(\alpha,\beta)\exp(-\gamma_{\upsilon i}x)\end{bmatrix}\cos(\alpha y)\sin(\beta z)\mathrm{d}\alpha\mathrm{d}\beta\right\},\\
E_x &= \frac{c_{44}}{e_{14}\mu_3}\sum_{i=1}^{2}\left[\int_0^\infty\int_0^\infty(\mu_3-\upsilon_i)\gamma_{\upsilon i}\beta C_i(\alpha,\beta)\exp(-\gamma_{\upsilon i}x)\cos(\alpha y)\cos(\beta z)\mathrm{d}\alpha\mathrm{d}\beta\right],\\
E_y &= \frac{c_{44}}{e_{14}\mu_3}\sum_{i=1}^{2}\left[\int_0^\infty\int_0^\infty(\mu_3-\upsilon_i)\alpha\beta C_i(\alpha,\beta)\exp(-\gamma_{\upsilon i}x)\sin(\alpha y)\cos(\beta z)\mathrm{d}\alpha\mathrm{d}\beta\right]
\end{aligned}\right\}. \quad (17)
$$

By substituting Eqs. (14), (16), and (17) into Eq. (1), $A_i(\alpha,\beta)$ and $C_i(\alpha,\beta)$ $(i=1,2)$ are obtained as

$$
\begin{Bmatrix}A_1(\alpha,\beta)\\ A_2(\alpha,\beta)\\ C_1(\alpha,\beta)\\ C_2(\alpha,\beta)\end{Bmatrix}=\frac{e_{14}\Phi_0}{c_{44}}\frac{f^*(\alpha,\beta)}{\beta}\frac{1}{\Delta(\alpha,\beta)}\begin{Bmatrix}A_1^*(\alpha,\beta)\\ A_2^*(\alpha,\beta)\\ C_1^*(\alpha,\beta)\\ C_2^*(\alpha,\beta)\end{Bmatrix}, \quad (18)
$$

where

$$
\left.\begin{aligned}
\Delta(\alpha,\beta) &\equiv 2\left[2\mu_3-(\upsilon_1+\upsilon_2)\right](k_1-k_2)\alpha^2\gamma_{\mu1}\gamma_{\mu2}(\gamma_{\upsilon1}-\gamma_{\upsilon2})\\
&\quad+(\upsilon_1-\upsilon_2)\begin{Bmatrix}2(k_1-k_2)\alpha^2\left[\gamma_{\mu1}\gamma_{\mu2}(\gamma_{\upsilon1}+\gamma_{\upsilon2})-\mu_3\beta^2(\gamma_{\mu1}+\gamma_{\mu2})\right]\\ -\mu_3(1+k_1)(1+k_2)\left(2\alpha^2+\mu_3\beta^2\right)\beta^2(\gamma_{\mu1}-\gamma_{\mu2})\\ -4\alpha^4(k_1\gamma_{\mu1}-k_2\gamma_{\mu2})\end{Bmatrix},\\
A_1^*(\alpha,\beta) &\equiv -\alpha\begin{Bmatrix}2(\upsilon_1-\upsilon_2)\alpha^2\left[2\alpha^2+(1+k_2)\mu_3\beta^2\right]\\ +2\left[2\alpha^2-\mu_3(1+k_2)\beta^2\right]\gamma_{\mu2}(\upsilon_2\gamma_{\upsilon1}-\upsilon_1\gamma_{\upsilon2})\\ -4\mu_3(1+k_2)\alpha^2\gamma_{\mu2}(\gamma_{\upsilon1}-\gamma_{\upsilon2})\end{Bmatrix},\\
C_1^*(\alpha,\beta) &\equiv -\begin{Bmatrix}4(k_1-k_2)\mu_3\alpha^2\gamma_{\mu1}\gamma_{\mu2}\gamma_{\upsilon2}\\ +\left(2\alpha^2+\mu_3\beta^2\right)(\gamma_{\mu1}-\gamma_{\mu2})\left[2\alpha^2(\upsilon_2-\mu_3)-\mu_3^2(1+k_1)(1+k_2)\beta^2\right]\\ -2\mu_3(k_1-k_2)\upsilon_2\alpha^2\beta^2(\gamma_{\mu1}+\gamma_{\mu2})-4\alpha^4\mu_3(k_1\gamma_{\mu1}-k_2\gamma_{\mu2})\end{Bmatrix}
\end{aligned}\right\}, \quad (19)
$$

and $A_2^*(\alpha,\beta)$ and $C_2^*(\alpha,\beta)$ are obtained by interchanging subscripts "1" and "2" in $A_1^*(\alpha,\beta)$ and $C_1^*(\alpha,\beta)$, respectively.

3 NUMERICAL CALCULATION

To illustrate the numerical results, the following nondimensional quantities are introduced:

$$(\hat{x},\hat{y},\hat{z}) \equiv \frac{(x,y,z)}{\delta},\ \ (\hat{E}_x,\hat{E}_y) \equiv \frac{(E_x,E_y)}{\left(\frac{\Phi_0}{\delta}\right)},\ \ (\hat{\sigma}_{yz},\hat{\sigma}_{zx}) \equiv \frac{(\sigma_{yz},\sigma_{zx})}{\left(e_{14}\frac{\Phi_0}{\delta}\right)}. \tag{20}$$

Numerical parameters are chosen as

$$\frac{(c_{11},\, c_{12},\, c_{13},\, c_{33})}{c_{44}} = (1.2,\, 0.4,\, 0.6,\, 15),\ \ \eta = 1.5,\ \ k_{\text{couple}} = 0.1\,. \tag{21}$$

Figures 2－4 show the distributions of the electric fields and the resulting shear stresses in $\hat{x}$, $\hat{y}$, and $\hat{z}$ directions, respectively. Figure 2 shows that the electric field $\hat{E}_x$ decreases monotonically toward zero with $\hat{x}$ and that the resulting shear stress $\hat{\sigma}_{yz}$ exhibits similar behavior. Figure 3 shows that the electric field $\hat{E}_x$ and shear stress $\hat{\sigma}_{yz}$ are maximum on the $\hat{x}$-axis and, roughly speaking, decrease toward zero with $\hat{y}$ and that, on the other hand, the electric field $\hat{E}_y$ and the resulting shear stress $\hat{\sigma}_{zx}$ are zero on the $\hat{x}$-axis, reach their maxima around the periphery of the surface electric potential, and decrease toward zero with $\hat{y}$. From Fig. 4 it is found that the distributions of the electric field $\hat{E}_x$ and shear stress $\hat{\sigma}_{yz}$ in $\hat{z}$ direction exhibit similar behavior to those in $\hat{y}$ direction.

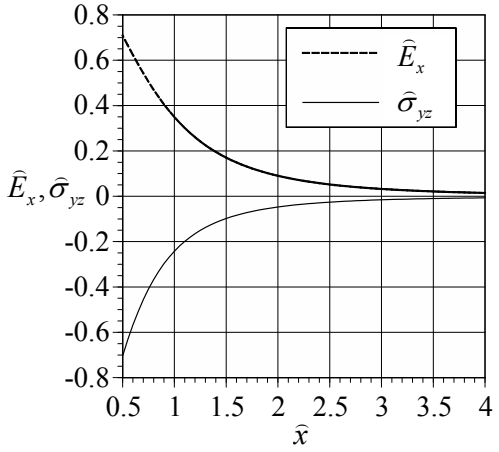


Figure 2: Distrubution of electric field and stress in $\hat{x}$ direction $(\hat{y}=0,\ \hat{z}=0)$

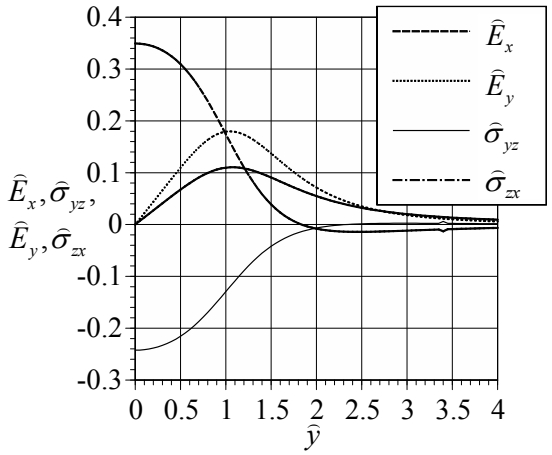


Figure 3: Distrubution of electric field and stress in $\hat{y}$ direction $(\hat{x}=1, \hat{z}=0)$

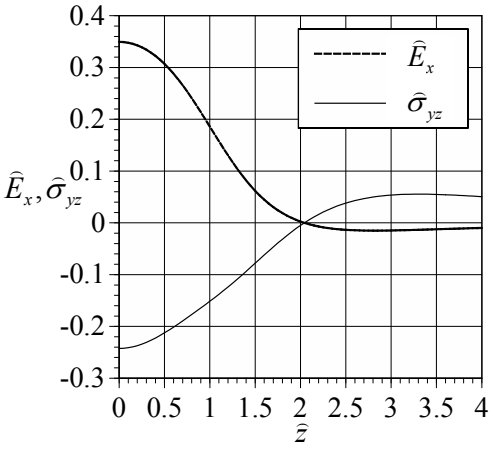


Figure 4: Distrubution of electric field and stress in $\hat{z}$ direction $(\hat{x}=1, \hat{y}=0)$

4 CONCLUSIONS

- The analytical solution of the electroelastic field in an semi-infinite body with D_∞ symmetry subjected locally uniform electric potential is formulated.
- For the analytical model above-mentioned, the electroelastic field inside the body, which is of great significance for sound operation of nondestructive evaluation techniques, is elucidated.

REFERENCES

[1] Fukada, E. Piezoelectricity of wood. *J. Phys Soc. Jpn* (1955) **10**:149-154.
[2] Smetana, J.A. and Kelso, P.W. Piezoelectric charge density measurements on the surface of Douglas-fir. *Wood Sci.* (1971) **3**:161-171.
[3] Knuffel, W and Pizzi, A. The piezoelectric effect in structural timber. *Holzforsch.* (1986)

40:157-162.
[4] Knuffel, W.E. The piezoelectric effect in structural timber - part II. The influence of natural defects. *Holzforsch.* (1988) **42**:247-252.
[5] Kim. S.K. *Group theoretical methods and applications to molecules and crystals*. Cambridge University Press, (1999).
[6] Elliott, H.A. Axial symmetric stress distributions in aeolotropic hexagonal crystals. *Math. Proc. Camb. Phil. Soc.* (1949) **45**:621-630.
[7] Chen, W.T. On some problems in transversely isotropic elastic materials. *J. Appl. Mech.* (1966) **33**:347-355.
[8] Atsumi, A. and Itou, S. Stresses in a transversely isotropic slab having a spherical cavity. *J. Appl. Mech.* (1973) **40**:752-758.
[9] Nakai, T., Igushi, N., and Ando, K. Piezoelectric behavior of wood under combined compression and vibration stresses I: Relation between piezoelectric voltage and microscopic deformation of a Sitka spruce (Picea sitchensis Carr.). *J. Wood Sci.* (1998) **44**:28-34.
[10] Nakai, T. and Ando, K. Piezoelectric behavior of wood under combined compression and vibration stresses II: Effect of the deformation of cross-sectional wall of tracheids on changes in piezoelectric voltage in linear-elastic region. *J. Wood Sci.* (1998) **44**:255-259.
[11] Nakai, T., Hamatake, M., and Nakao, T. Relationship between piezoelectric behavior and the stress – strain curve of wood under combined compression and vibration stresses. *J. Wood Sci.* (2004) **50**:97-99.
[12] Nakai, T., Yamamoto, H., Hamatake, M., and Nakao, T. Initial shapes of stress-strain curve of wood specimen subjected to repeated combined compression and vibration stresses and the piezoelectric behavior. *J. Wood Sci.* (2006) **52**:539-543.
[13] Sneddon, I.N. *The use of integral transforms.* McGraw-Hill, (1972).

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis(Eds)

ANANLYSIS OF ACTUATOR MADE OF FGM USING NEW ELECTRO-THERMO-MECHANICAL FINITE ELEMENTS

JURAJ PAULECH*, JURAJ HRABOVSKÝ*, JUSTÍN MURÍN* AND VLADIMÍR KUTIŠ*

*Department Applied Mechanics and Mechatronics
Institute of Automotive Mechatronics
Faculty of Electrical Engineering and Information Technology
Slovak University of Technology in Bratislava
812 19 Bratislava, Slovak Republic
e-mail: juraj.paulech@stuba.sk, web page: http://www.stuba.sk

Key words: Actuator, coupled electro-thermo-mechanical analysis, Functionally Graded Materials, new FEM equations

Abstract. Actuator is a mechatronic system that transforms one type of energy (e.g. electric energy) into the mechanical displacement and mechanical force (mechanical energy). Nowadays, these actuators can be made of Functionally Graded Materials (FGM) to ensure simple shape of the actuator and to improve its effectiveness, particularly for micro systems. FGM is built as a mixture of two or more constituents which have almost the same geometry and dimensions. The variation of macroscopic material properties can be induced by variation of both the volume fractions and material properties (e.g. by a non-homogeneous temperature field) of the FGM constituents.
The paper deals with a new approach in analysing of the systems made of FGM using our new beam finite elements. Multiphysical analysis (weak coupled electro-thermo-mechanical analysis) and spatial continuous variation of material properties are supported. The analysis of the micro actuator with constant cross section made of FGM is presented in the paper. This simple-shaped actuator is supplied by electric current and the efficiency of the actuator is optimised. The solution results will be compared with those obtained by using solid elements of a FEM commercial program.

1 INTRODUCTION

Nowadays, the scientific and technological progress are already at such level that for the development of new systems in classical way (such as mechanical and heating systems, systems in construction industry, etc.) it is not enough to propose new shapes of components and their optimization, but it requires use of new materials with the desired properties that lie outside the parameters of materials commonly used for that purpose.

New materials like Functionally Graded Material (FGM) are necessary for sophisticated structures like Micro-Electro-Mechanical Systems (MEMS), advanced electronic devices, etc. In all these applications, using new materials like FGM can greatly improve efficiency of a system e.g. classic shape of actuator (Fig. 1a) can be replaced by new type – simply-shaped actuator (Fig. 1b) where functionality is caused by varying material properties.

Figure 1: a) Classic shape of MEMS actuator, b) New shape of FGM actuator

FGM is built as a mixture of two or more constituents which have almost the same geometry and dimensions. From macroscopic point of view, FGM is isotropic in each material point but the material properties can vary continuously or discontinuously in one, two or three directions. The variation of macroscopic material properties can be caused by varying the volume fraction of the constituents or with varying of the constituents' material properties (e.g. by non-homogeneous temperature field). The methods based on the homogenization theory have been designed and successfully applied to determine the effective material properties of heterogeneous materials from the corresponding material behavior of the constituents (and of the interfaces between them) and from the geometrical arrangement of the phases.
Coupled electro-thermo-mechanical analysis of actuator made of FGM using our new beam finite elements will be presented.

2 FEM EQUATIONS FOR COUPLED ELECTRO - THERMO - MECHANICAL ANALYSIS

Derivation process of the new FEM equations for coupled electro-thermo-mechanical element is based on differential equations for electric thermal and structural fields for 1D type of analysis, respectively. All quantities in following equations are the polynomial functions of x. Homogenization process of the varying material properties and the calculation of other effective finite element parameters have been done by extended mixture rule [1] and multilayer method is fully described in [2],[3].

2.1 Differential equations

Homogeneous 1D static differential equation for FGM (with non-constant coefficients on on the left-hand side) for electric field with boundary conditions has a form:

$$\begin{gathered} -\sigma(x)\frac{d^2\varphi(x)}{dx^2} - \frac{d\sigma(x)}{dx}\frac{d\varphi(x)}{dx} = 0 \\ \varphi(0) = \varphi_0 \qquad J(L) = J_L \end{gathered} \tag{1}$$

where $\varphi(x)$ [V] is the electric potential, $\sigma(x)$ [S/m] is the specific electric conductivity and $J(x)$ [Am^{-2}] is the current density.

Static differential equation for heat transfer with non-constant auxiliary thermal source $Q(x)$ [Wm^{-3}] in the volume, with non-constant convective heat transfer coefficient $\alpha_t(x)$ [K^{-1}] and with coupled to the electric field has a form (2). One-way coupling between the electric and thermal field is provided by Joule heat $P_{J1}(x)$ [W/m^3], that can be calculated as one of the outputs from electric analysis and it enters the thermal analysis as volume heat (beside or instead of $Q(x)$).

$$-\lambda(x)\frac{d^2T(x)}{dx^2} - \frac{d\lambda(x)}{dx}\frac{dT(x)}{dx} + \alpha_t(x)T(x)\frac{o}{A} = P_{J1}(x) + Q(x) + \alpha_t(x)T_{amb}\frac{o}{A} \tag{2}$$

with boundary conditions, e.g.:

$$T(0) = 0 \qquad q(L) = q_L \tag{3}$$

where $\lambda(x)$ [Wm^{-1}K^{-1}] is the thermal conductivity, $T(x)$ [K] is the temperature, o [m] is the perimeter, A [m^2] is the cross section area, T_{amb} [K] is the ambient temperature and $q(x)$ [Wm^{-2}] is the heat flux.

Homogeneous differential equation for structural analysis with effect of thermal expansion (coupling with the electro-thermal analysis) for pure tensile and compressive stress has a form:

$$\begin{aligned} &E_L^{NH}(x)\frac{d^2u(x)}{dx^2} + \frac{dE_L^{NH}(x)}{dx}\frac{du(x)}{dx} = \\ &= \frac{n(x)}{A} + \alpha_t(x)\Delta t(x)\frac{dE_L^{NH}(x)}{dx} + \Delta T(x)E_L^{NH}(x)\frac{d\alpha_t(x)}{dx} + E_L^{NH}(x)\alpha_t(x)\frac{d\Delta t(x)}{dx} \end{aligned} \tag{4}$$

with boundary conditions, e.g.:

$$u(0) = u_0 \qquad \sigma_N^s(L) = \sigma_{N,L}^s \tag{5}$$

where $E_L^{NH}(x)$ [Pa] is the Young modulus for tension/compression, $u(x)$ [m] is the displacement, $n(x)$ [Nm^{-1}] are the distributed axial forces, $N(x)$ [N] is the normal force and α_t [K^{-1}] is the coefficient of thermal expansion.

Homogeneous differential equation for structural analysis for bending has a form:

$$\frac{d^2w(x)}{dx^2} = \frac{M(x)}{E_L^{MH}}(x)I_y \tag{6}$$

with boundary conditions, e.g.:

$$w(0) = w_0 \qquad \sigma_y(L) = \sigma_{y,L} \tag{7}$$

where $w(x)$ [m] is the transversal displacement, $M(x)$ is the bending moment, $E_L^{MH}(x)$ is the Young modulus for bending, $\varphi_y(x)$ [rad] is the angle of the cross section rotation (around y axis), I_y [m^4] is the quadratic moment of the cross section.

The solution of these differential equations is based on numerical method for solving 1D differential equation with non-constant coefficients and with right-hand side described in [4] in detail.

2.2 New beam/link FGM finite element equations

The finite element equations for electric analysis in FGM link have a form:

$$\begin{bmatrix} c_0(L) & -1 \\ -\left(c_0(L) - \dfrac{c_1(L)c_0'(L)}{c_1'(L)}\right) & 1 \end{bmatrix} \cdot \begin{bmatrix} \varphi_0 \\ \varphi_L \end{bmatrix} = \begin{bmatrix} \dfrac{c_1(L)}{\sigma_0} J_0 \\ \dfrac{c_1(L)}{c_1'(L)\sigma_L} J_L \end{bmatrix} \tag{8}$$

FEM equations for thermal analysis considering the convective effect, generated heat and Joule heat (coupling between the electric and thermal field) have a form:

$$\begin{bmatrix} c_0(L) & -1 \\ -\left(c_0(L) - \dfrac{c_1(L)c_0'(L)}{c_1'(L)}\right) & 1 \end{bmatrix} \cdot \begin{bmatrix} T_0 \\ T_L \end{bmatrix} =$$

$$= \begin{bmatrix} \dfrac{c_1(L)}{\lambda_0} q_0 - \displaystyle\sum_{j=0}^{g} \varepsilon_j b_{j+2}(L) \\ \dfrac{c_1(L)}{c_1'(L)\lambda_L} q_L - \dfrac{c_1(L)}{c_1'(L)} \displaystyle\sum_{j=0}^{g} \varepsilon_j b_{j+2}'(L) + \displaystyle\sum_{j=0}^{g} \varepsilon_j b_{j+2}(x) \end{bmatrix} \tag{9}$$

Derived FEM equations for the structural analysis for pure tensile and compressive stress with coupling to the electro-thermal analysis (thermal expansion coefficient) have a form:

$$\begin{bmatrix} c_0(L) & -1 \\ -\left(c_0(L) - \dfrac{c_1(L)c_0'(L)}{c_1'(L)}\right) & 1 \end{bmatrix} \cdot \begin{bmatrix} u_0 \\ u_L \end{bmatrix} =$$

$$= \begin{bmatrix} \dfrac{c_1(L)}{E_L^{NH} A} N_0 - c_1(L)\alpha_{t0}\Delta T_0 - \displaystyle\sum_{j=0}^{g} \varepsilon_j b_{j+2}(L) \\ \dfrac{c_1(L)}{c_1'(L)} \left(\dfrac{N_L}{E_L^{NH} A} + \alpha_{tL}\Delta T_L - \displaystyle\sum_{j=0}^{g} \varepsilon_j b_{j+2}'(L) \right) + \displaystyle\sum_{j=0}^{g} \varepsilon_j b_{j+2}(L) \end{bmatrix} \tag{10}$$

and FEM equations for bending of the beam have general form:

$$\begin{bmatrix} K_{11} & K_{12} & K_{13} & K_{14} \\ K_{21} & K_{22} & K_{23} & K_{24} \\ K_{31} & K_{32} & K_{33} & K_{34} \\ K_{41} & K_{42} & K_{43} & K_{44} \end{bmatrix} \cdot \begin{bmatrix} w_0 \\ \varphi_{y,0} \\ w_L \\ \varphi_{y,L} \end{bmatrix} = \begin{bmatrix} T_{z,0} \\ M_0 \\ T_{z,L} \\ M_L \end{bmatrix} \tag{11}$$

where $T_z(x)$ [N] is the transversal force.
The terms $c_i(x)$, c_i', $b_i(x)$, $b_i'(x)$, $i \in \langle 0, 1\rangle$ are the transfer functions (for particular solution and for uniform solution) of the differential equations (1) - (6) which can be calculated by simple numerical algorithm [4].

3 NUMERICAL EXPERIMENT

Let us consider actuator with constant cross section made of FGM according to Fig. 2. It consists of 3 parts (beams) that lengths are: $L_1 = 5$ mm, $L_2 = 0.25$ mm and $L_3 = 5$ mm. Their constant rectangular cross-section is $b = 0.2$ mm and $h = 0.1$ mm.

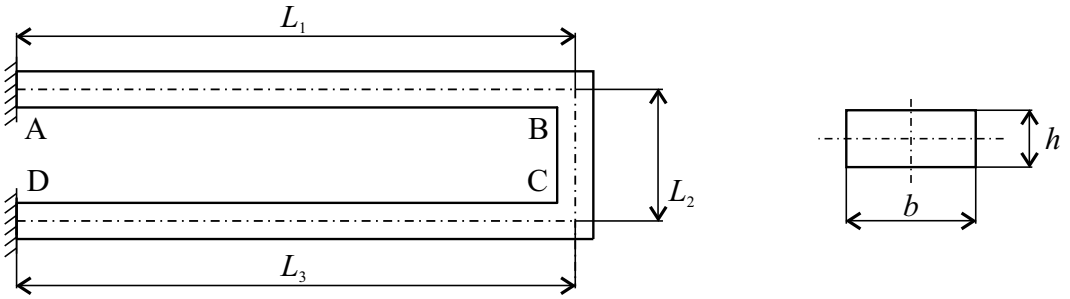


Figure 2: FGM microactuator

Actuator is made of FGM that consist of two components: NiFe - named as matrix and denoted with index m and Tungsten - named as fibre and denoted with index f. Material properties of the components are constant (not temperature dependent): Nife (matrix): Young modulus $E_m = 255$ GPa, thermal conductivity $\lambda_m = 100$ W/mK, electric conductivity $\gamma_m = 1.31 \times 10^7$ S/m, thermal expansion coefficient $\alpha_m = 1.5 \times 10^{-5}$ K^{-1}; Thungsten (fibre): Young modulus $E_f = 400$ GPa, thermal conductivity $\lambda_f = 160$ W/mK, electric conductivity $\gamma_f = 2.84 \times 10^7$ S/m, thermal expansion coefficient $\alpha_m = 5.3 \times 10^{-5}$ K^{-1}. The variation of material properties is caused by varying volume fraction. Variation of the fibre's volume fraction has been chosen as the polynomial function of x, y.

- Part 1

$$\begin{aligned} v_f(x,y) =& 1/10 + 220 \times 10^3 x^2 - 232 \times 10^6 x^3 + 73600 \times 10^6 x^4 - 6912 \times 10^9 x^5 \\ & - 112 \times 10^{12} x^2 y^2 + 108800 \times 10^{12} x^3 y^2 - 33280 \times 10^{15} x^4 y^2 \\ & + 3072 \times 10^{18} x^5 y^2 - 112 \times 10^{15} x^2 y^3 + 108800 \times 10^{15} x^3 y^3 \\ & - 33280 \times 10^{18} x^4 y^3 + 3072 \times 10^{21} x^5 y^3 + 22400 \times 10^{18} x^2 y^4 \\ & - 21760 \times 10^{21} x^3 y^4 + 6656000 \times 10^{21} x^4 y^4 - 614400 \times 10^{24} x^5 y^4 \\ & + 44800 \times 10^{21} x^2 y^5 - 43520 \times 10^{24} x^3 y^5 + 13312 \times 10^{27} x^4 y^5 \\ & - 1228800 \times 10^{27} x^5 y^5 \end{aligned}$$

- Part 2

$$v_f(x,y) = 0.766683$$

- Part 3

$$\begin{aligned} v_f(x,y) =& 0.766683 + 4.26633 \times 10^6 x^2 - 1.84858 \times 10^{10} x^3 + 3.89101 \times 10^{13} x^4 \\ & - 4.99269 \times 10^{16} x^5 + 4.25073 \times 10^{19} x^6 - 2.49749 \times 10^{22} x^7 \\ & + 1.0285 \times 10^{25} x^8 - 2.96182 \times 10^{27} x^9 + 5.83847 \times 10^{29} x^{10} \\ & - 7.49901 \times 10^{31} x^{11} + 5.64709 \times 10^{33} x^{12} - 1.88879 \times 10^{35} x^{13} \end{aligned}$$

Variation of the fibres volume fraction $v_f(x,y)$ for part 1 and 3 are shown in Figure 3. The same variation of the fibres volume fraction at the points B and C has been assumed.

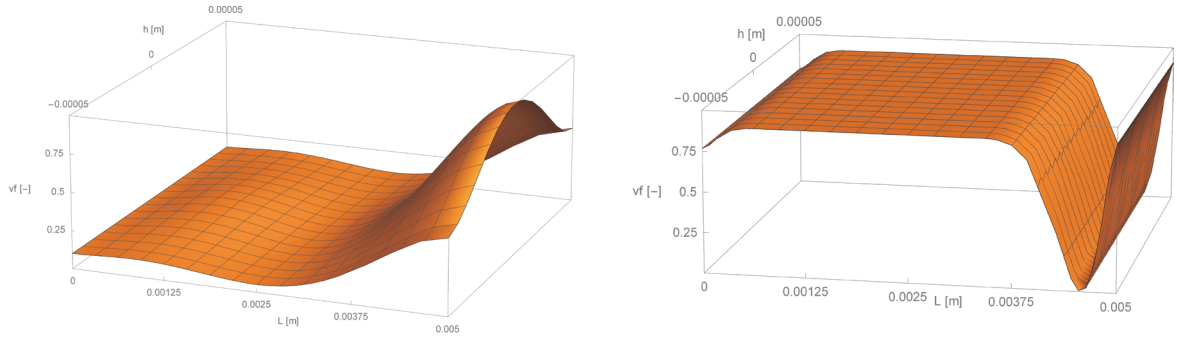


Figure 3: Variation of the fibres volume fraction, left - part 1, right - part 3

The effective material properties (Young modulus for tension/compression $E_L^{NH}(x)$ [Pa], Young modulus for bending $E_L^{MH}(x)$ [Pa], thermal conductivity $\lambda_L^H(x)$ [$\mathrm{Wm^{-1}K^{-1}}$], electric conductivity $\rho_L^H(x)$ [S/m], thermal expansion coefficient α_{TL}^H [$\mathrm{K^{-1}}$]) of the homogenized beam have been calculated by multilayered method [2], [3] and the results are:

- Beam 1 (the local axis x begins at node A and ends at point B - see Fig. 2):

$$\begin{aligned} E_L^{NH} =& 2.695 \times 10^{11} + 2.24267 \times 10^{16} x^2 - 2.44373 \times 10^{19} x^3 \\ & + 7.85707 \times 10^{21} x^4 - 7.424 \times 10^{23} x^5 \end{aligned}$$

$$\begin{aligned} E_L^{MH} =& 2.695 \times 10^{11} + 1.70473 \times 10^{8} x + 1.62396 \times 10^{16} x^2 \\ & - 1.84272 \times 10^{19} x^3 + 6.01868 \times 10^{21} x^4 - 5.72704 \times 10^{23} x^5 \end{aligned}$$

$$\begin{aligned} \lambda_L^H =& 106 + \frac{135868725000}{14641} x^2 - \frac{148050030000000}{14641} x^3 \\ & + \frac{47600892000000000}{14641} x^4 + \frac{4497721920000000000}{14641} x^5 \end{aligned}$$

$$\begin{aligned} \rho_L^H =& 1.463 \times 10^7 + 2.3664 \times 10^{12} x^2 - 2.57856 \times 10^{15} x^3 \\ & + 8.29057 \times 10^{17} x^4 - 7.83361 \times 10^{19} x^5 \end{aligned}$$

$$\begin{aligned}\alpha_{TL}^{H} =& 0.00001356 - 0.000286032x - 1.52203x^2 + 1923.69x^3 \\ & - 649717x^4x^5 - 6.34908 \times 10^7 x^5\end{aligned}$$

- Beam 2 (the local axis x begins at node B and ends at point C - see Fig. 2):

$$E_L^{NH} = 3.66169 \times 10^7$$

$$E_L^{MH} = 3.66169 \times 10^7$$

$$\lambda_L^{H} = 146.001$$

$$\rho_L^{H} = 2.48302 \times 10^7$$

$$\alpha_{TL}^{H} = 0.0000118$$

- Beam 3 (the local axis x begins at node C and ends at point D - see Fig. 2):

$$\begin{aligned}E_L^{NH} =& 3.66169 \times 10^{11} + 6.18618 \times 10^{17}x^2 - 2.68044 \times 10^{21}x^3 \\ & + 5.64197 \times 10^{24}x^4 - 7.2394 \times 10^{27}x^5 + 6.16356 \times 10^{30}x^6 \\ & - 3.62135 \times 10^{33}x^7 + 1.49132 \times 10^{36}x^8 - 4.29464 \times 10^{38}x^9 \\ & + 8.46578 \times 10^{40}x^{10} - 1.08736 \times 10^{43}x^{11} + 8.18829 \times 10^{44}x^{12} \\ & - 2.73875 \times 10^{46}x^{13}\end{aligned}$$

$$\begin{aligned}E_L^{MH} =& 3.66169 \times 10^{11} + 6.18618 \times 10^{17}x^2 - 2.68044 \times 10^{21}x^3 \\ & + 5.64197 \times 10^{24}x^4 - 7.2394 \times 10^{27}x^5 + 6.16356 \times 10^{30}x^6 \\ & - 3.62135 \times 10^{33}x^7 + 1.49132 \times 10^{36}x^8 - 4.29464 \times 10^{38}x^9 \\ & + 8.46578 \times 10^{40}x^{10} - 1.08736 \times 10^{43}x^{11} + 8.18829 \times 10^{44}x^{12} \\ & - 2.73875 \times 10^{46}x^{13}\end{aligned}$$

$$\begin{aligned}\lambda_L^{H} =& 146.001 + 2.5598 \times 10^{8}x^2 - 1.10915 \times 10^{12}x^3 \\ & + 2.33461 \times 10^{15}x^4 - 2.99561 \times 10^{18}x^5 + 2.55044 \times 10^{21}x^6 \\ & - 1.49849 \times 10^{24}x^7 + 6.17099 \times 10^{26}x^8 - 1.77709 \times 10^{29}x^9 \\ & + 3.50308 \times 10^{31}x^{10} - 4.49941 \times 10^{33}x^{11} + 3.38826 \times 10^{35}x^{12} \\ & - 1.13328 \times 10^{37}x^{13}\end{aligned}$$

$$\begin{aligned}\rho_L^{H} =& 2.48302 \times 10^{7} + 6.52749 \times 10^{13}x^2 - 2.82832 \times 10^{17}x^3 \\ & + 5.95325 \times 10^{20}x^4 - 7.63882 \times 10^{23}x^5 + 6.50362 \times 10^{26}x^6 \\ & - 3.82115 \times 10^{29}x^7 + 1.5736 \times 10^{32}x^8 - 4.53158 \times 10^{34}x^9 \\ & + 8.93286 \times 10^{36}x^{10} - 1.14735 \times 10^{39}x^{11} + 8.64006 \times 10^{40}x^{12} \\ & - 2.88985 \times 10^{42}x^{13}\end{aligned}$$

$$\alpha_{TL}^{H} = 6.87608 \times 10^{-6} - 0.000186889x$$

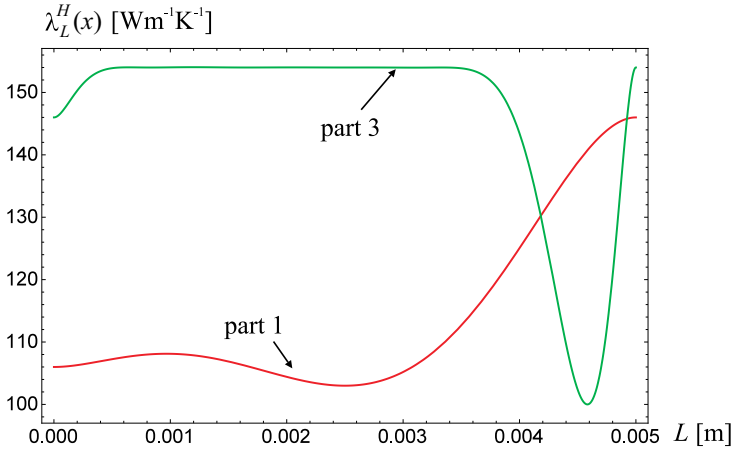


Figure 4: Homogenized thermal conductivity

The homogenized thermal conductivity $\lambda_L^H(x)$ at part 1 and 3 is shown in Fig. 4.

The applied constrains and loads:

- electric potential and current: $V_A = 0$ V, $I_D = 5$ A;
- temperatures: $T_A = 25$ °C, TD = 25 °C;
- fixed support: $u_A = 0$ m, $u_D = 0$ m,
  $w_A = 0$ m, $w_D = 0$ m,
  $\varphi_A = 0$ rad, $\varphi_D = 0$ rad;

The coupled electro-thermo-mechanical analysis of FGM actuator has been done using our new FGM beam/link finite elements. The calculation has been done using software MATHEMATICA [5]. Only three our new finite elements have been used (one for each part). The same problem has been solved using a fine mesh - 12998 of PLANE223 elements of the FEM program ANSYS [6]. The average relative difference Δ [%] between quantities calculated by our method and the ANSYS solution has been evaluated.
Electric analysis was performed as the first solution and the nodal electric variables φ have been obtained (see Table 1).

Table 1: The results of electric analysis

electric potential [V]	new element	ANSYS	Δ [%]
φ_B	0.077782	0.0775569	0.28
φ_C	0.080298	0.0796372	0.82
φ_D	0.131995	0.131098	0.67

Thermal analysis was performed as the second one. Distributed thermal load - Joule heat caused by electric current, was included into the analysis. The results of thermal alalysis are presented in Table 2.

Table 2: The results of thermal analysis

temperature [°C]	new element	ANSYS	Δ [%]
T_B	346.33	340.42	1.71
T_C	342.08	336.77	1.55

Structural analysis is performed as the last analysis, where thermal forces caused by thermal expansion were included into the model. The solution of structural analysis are the displacements u for longitudinal direction and w for transversal direction (see Table 3).

Table 3: The results of structural analysis

displacement [m]	new element	ANSYS	Δ [%]
u_B	0.13713×10^{-4}	0.13588×10^{-4}	0.91
w_B	-0.70975×10^{-4}	-0.72665×10^{-4}	2.38
u_C	0.67744×10^{-5}	0.65903×10^{-5}	2.71
w_C	-0.71917×10^{-4}	-0.73204×10^{-4}	1.78

As it can be seen in Tables 1 - 3, a very good agreement of both solution results has been obtained. The comparison of total deformation of the FGM actuator calculated by our new approach and comercial FEM program ANSYS is shown in Fig. 5.

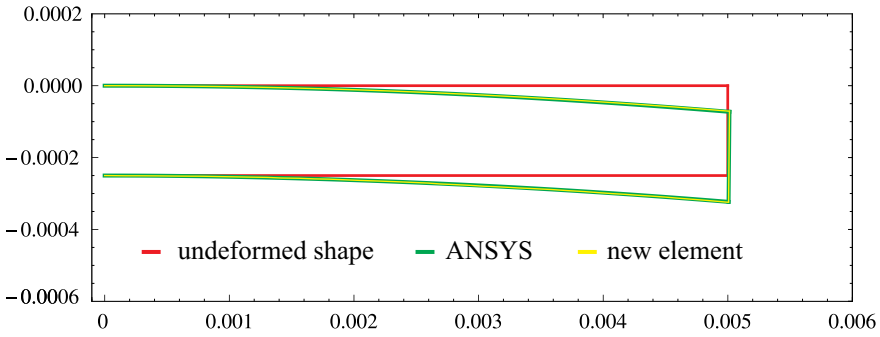


Figure 5: Total deformation of the FGM actuator

4 CONCLUSIONS

New FEM equations for weak coupled static electro-thermo-mechanical analysis of the FGM beam structures have been presented in this contribution. The numerical experiment – multiphysical analysis of micro actuator made of FGM has been done using our new approach and obtained results have been compared with ones obtained by solution with software ANSYS. The effectiveness and accuracy of the new finite elements have been shown.

Acknowledgement: This paper has been supported by Grant Agency APVV-0246-12, Grant Agency VEGA No. 1/0228/14, Grant Agency VEGA No. 1/0453/15 and Agency KEGA No. 007STU-4/2015.

REFERENCES

[1] Murín, J., Kutiš, V. *Improved mixture rules for composite (FGMs) sandwich beam finite element.* In Computational Plasticity IX. Fundamentals and Applications. Barcelona, Spain, (2007): 647-650.

[2] Kutiš, V., J. Murín, J., Belák, R. and Paulech, J. *Beam element with spatial variation of material properties for multiphysics analysis of functionally graded materials.* Computers and Structures, 89:11921205, 2011.

[3] Murín, J., Kutiš, V., Paulech, J. and Hrabovský, J. *Electric-Thermal Link Finite Element Made of a FGM with Spatial Variation of Material Properties.* Composites Part B: Engineering, 42:19661979, 2011.

[4] Rubin, H. *Solution of differential equations of arbitrary order with polynomial coefficients and application to a statics problem.* ZAMM 76 (1996), 105-117.

[5] S. Wolfram MATHEMATICA 5, Wolfram research, Inc., 2003.

[6] ANSYS Swanson Analysis System, Inc., 201 Johnson Road, Houston, PA 15342/1300, USA.

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis(Eds)

COMPARING KINETIC AND HYDRODYNAMICAL MODELS FOR ELECTRON TRANSPORT IN MONOLAYER GRAPHENE

MARCO COCO*, ARMANDO MAJORANA†, GIOVANNI MASCALI‡, VITTORIO ROMANO†

* Department of Mathematics and Computer Science
University of Catania
Viale A. Doria 6, 95125 Catania, Italy
e-mail: mcoco@dmi.unict.it

† Department of Industrial Engineering
University of Catania
Viale A. Doria 6, 95125 Catania, Italy
e-mail: majorana@dii.unict.it, romano@dii.unict.it

‡ Department of Mathematics and Computer Science
University of Calabria
and INFN-Gruppo c. 87036 Cosenza, Rende, Italy
e-mail: giovanni.mascali@unical.it

Key words: Graphene, Boltzmann Transport Equation, Discontinuous Galerkin Method, Hydrodynamical Models for Semiconductors.

Abstract. The aim of this work is to compare, in monolayer graphene, solutions of the electron Boltzmann equation, obtained with a discontinuous Galerkin method, with those of a hydrodynamical model based on the Maximum Entropy Principle.

1 INTRODUCTION

Graphene is a gapless semiconductor made of a single layer of carbon atoms arranged into a honeycomb hexagonal lattice [1]. In view of applications in graphene-based electron devices, it is crucial to understand the basic transport properties of this material.

A physically accurate model is given by a semiclassical transport equation whose scattering terms have been deeply analyzed recently [2, 3, 4]. Due to the computational difficulties, the most part of the available solutions have been obtained by direct Monte Carlo simulations. A different approach has been employed in [5].

For computer aided design (CAD) purposes, it could be useful to have macroscopic

models like drift-diffusion, energy-transport and hydrodynamical ones. Macroscopic models have been proposed, for example, in [6, 7, 8].

The aim of this work is to assess the validity of the hydrodynamical model based on the Maximum Entropy Principle (MEP) [7], by comparing the solutions of this model with those of the transport equation for electrons in suspended monolayer graphene. A numerical scheme based on the discontinuous Galerkin method is used for finding the solutions of the electron Boltzmann equation because the same method has been already successfully applied to a more conventional semiconductor material like silicon [9, 10].

Comparison of the physically average quantities, electron energy and velocity, shows that the MEP model is reasonable even if the introduction of some improvements regarding additional moments or nonlinear effects is needed.

2 THE TRANSPORT EQUATIONS FOR ELECTRONS IN GRAPHENE

The electron energy in graphene depends on a two dimensional wave vector $\mathbf{k}$ belonging to a bi-dimensional Brillouin zone which has an hexagonal shape.

The most part of electrons are in the valleys around the vertices of the Brillouin zone, called Dirac points or K and K' points. Usually the K and K' valleys are treated as a single equivalent one.

In a semiclassical kinetic setting, the charge transport in graphene is described by four Boltzmann equations, one for electrons in the valence band (π) and one for electrons in the conduction band (π^*), that in turn can belong to the K or K' valley,

$$\frac{\partial f_{\ell,s}(\mathrm{t},\mathbf{x},\mathbf{k})}{\partial t} + \mathbf{v}_{\ell,s}\cdot\nabla_{\mathbf{x}} f_{\ell,s}(\mathrm{t},\mathbf{x},\mathbf{k}) - \frac{e}{\hbar}\mathbf{E}\cdot\nabla_{\mathbf{k}} f_{\ell,s}(\mathrm{t},\mathbf{x},\mathbf{k}) = \left.\frac{df_{\ell,s}}{dt}(\mathrm{t},\mathbf{x},\mathbf{k})\right|_{e-ph}, \qquad (1)$$

where $f_{\ell,s}(\mathrm{t},\mathbf{x},\mathbf{k})$ represents the distribution function of charge carriers, band π or π^* ($s=-1$ or $s=1$), in the valley ℓ (K or K'), at position $\mathbf{x}$, time t and wave-vector $\mathbf{k}$. We denote by $\nabla_{\mathbf{x}}$ and $\nabla_{\mathbf{k}}$ the gradients with respect to the position and the wave vector, respectively. The microscopic velocity $\mathbf{v}_{\ell,s}$ is related to the band energy $\varepsilon_{\ell,s}$ by

$$\mathbf{v}_{\ell,s} = \frac{1}{\hbar}\nabla_{\mathbf{k}}\,\varepsilon_{\ell,s}\,.$$

With a very good approximation [1] a linear dispersion relation holds for the band energies $\varepsilon_{\ell,s}$ around the equivalent Dirac points; so that $\varepsilon_{\ell,s} = s\,\hbar\,v_F\,|\mathbf{k}-\mathbf{k}_\ell|$, where v_F is the (constant) Fermi velocity, $\hbar$ the Planck constant divided by $2\,\pi$, and $\mathbf{k}_\ell$ is the position of the Dirac point ℓ. The elementary (positive) charge is denoted by e, and $\mathbf{E}$ is the electric field obtained by the Poisson equation, which must be coupled with Eq. (1). The right hand side of Eq. (1) is the collision term representing the interaction of electrons with acoustic, optical and K phonons. Acoustic phonon scattering is intra-valley and intra-band. Optical phonon scattering is intra-valley and can be longitudinal optical (LO) and transversal optical (TO); it can be intra-band, leaving the electron in the same band, or

inter-band, pushing the electron from an initial band to another one. Scattering with optical phonon of K type pushes electrons from a valley to a nearby one (inter-valley scattering). We assume that phonons are a bath at thermal equilibrium. Hence, the general form of the collision term can be written as

$$\left.\frac{df_{\ell,s}}{dt}(\mathrm{t},\mathbf{x},\mathbf{k})\right|_{e-ph} = \sum_{\ell',s'}\left[\int S_{\ell',s',\ell,s}(\mathbf{k}',\mathbf{k})\, f_{\ell',s'}(\mathrm{t},\mathbf{x},\mathbf{k}')\,(1-f_{\ell,s}(\mathrm{t},\mathbf{x},\mathbf{k}))\,d\mathbf{k}' - \int S_{\ell,s,\ell',s'}(\mathbf{k},\mathbf{k}')\, f_{\ell,s}(\mathrm{t},\mathbf{x},\mathbf{k})\,(1-f_{\ell',s'}(\mathrm{t},\mathbf{x},\mathbf{k}'))\,d\mathbf{k}'\right],$$

where the total collision term is given by the sum of the contributions of the several types of scatterings described above

$$S_{\ell',s',\ell,s}(\mathbf{k}',\mathbf{k}) = \sum_{\nu}\left|G^{(\nu)}_{\ell',s',\ell,s}(\mathbf{k}',\mathbf{k})\right|^2\left[\left(n^{(\nu)}_{\mathbf{q}}+1\right)\delta\left(\varepsilon_{\ell,s}(\mathbf{k})-\varepsilon_{\ell',s'}(\mathbf{k}')-\hbar\,\omega^{(\nu)}_{\mathbf{q}}\right) + n^{(\nu)}_{\mathbf{q}}\,\delta\left(\varepsilon_{\ell,s}(\mathbf{k})-\varepsilon_{\ell',s'}(\mathbf{k}')+\hbar\,\omega^{(\nu)}_{\mathbf{q}}\right)\right]. \tag{2}$$

The index ν labels the ν-th phonon mode, $G^{(\nu)}_{\ell',s',\ell,s}(\mathbf{k}',\mathbf{k})$ is the scattering rate, which describes the scattering mechanism, due to the ν-th phonons, between electrons belonging to valley ℓ' and band s', and electrons belonging to valley ℓ and band s. The symbol δ denotes the Dirac distribution function, $\omega^{(\nu)}_{\mathbf{q}}$ is the ν-th phonon frequency, $n^{(\nu)}_{\mathbf{q}}$ is the Bose-Einstein distribution for the ν-type phonons

$$n^{(\nu)}_{\mathbf{q}} = \frac{1}{e^{\hbar\,\omega^{(\nu)}_{\mathbf{q}}/k_BT}-1},$$

where k_B is the Boltzmann constant and T the constant graphene lattice temperature. If, for a phonon ν_*-type, $\hbar\,\omega^{(\nu_*)}_{\mathbf{q}} \ll k_BT$, then the corresponding scattering can be assumed elastic. In this case, we eliminate in Eq. (2) the term $\hbar\,\omega^{(\nu_*)}_{\mathbf{q}}$ inside the delta distribution and we use the Laurent approximation $n^{(\nu_*)}_{\mathbf{q}} \approx k_BT/\hbar\,\omega^{(\nu_*)}_{\mathbf{q}}$.

Electrons which contribute to the charge transport in graphene are those in the conduction and valence band, and it is preferable to treat the latter as holes for insuring the integrability of the corresponding distribution function. Electrons and holes mostly populate the states near to the K and K' valleys. In this paper we consider the case of a high value of the Fermi energy which is equivalent for conventional semiconductors to a n-type doping. Under such a condition electrons belonging to the conduction band do not move to the valence band and vice versa. Therefore the hole dynamics is neglected.

A reference frame centered in the K-point will be used and in order to simplify the notation the indices s and ℓ will be omitted.

Under the above hypotheses the scattering rates read as follows.

For acoustic phonons, we consider the elastic approximation

$$2\,n^{(ac)}_{\mathbf{q}}\left|G^{(ac)}(\mathbf{k}',\mathbf{k})\right|^2 = \frac{1}{(2\,\pi)^2}\,\frac{\pi\,D^2_{ac}\,k_B\,T}{\hbar\,\sigma_m\,v_p^2}\,(1+\cos\vartheta_{\mathbf{k},\mathbf{k}'}), \tag{3}$$

where D_{ac} is the acoustic phonon coupling constant, v_p is the sound speed in graphene, σ_m is the graphene areal density, and $\vartheta_{\mathbf{k},\mathbf{k}'}$ is the convex angle between $\mathbf{k}$ and $\mathbf{k}'$.
There are three relevant optical phonon scatterings due to the longitudinal optical (LO), the transversal optical (TO) and the K (K) phonons. The related scattering rates are

$$\left|G^{(LO)}(\mathbf{k}',\mathbf{k})\right|^2 = \frac{1}{(2\,\pi)^2}\,\frac{\pi\,D_O^2}{2\,\sigma_m\,\omega_O}\,(1-\cos(\vartheta_{\mathbf{k},\mathbf{k}'-\mathbf{k}}+\vartheta_{\mathbf{k}',\mathbf{k}'-\mathbf{k}})) \tag{4}$$

$$\left|G^{(TO)}(\mathbf{k}',\mathbf{k})\right|^2 = \frac{1}{(2\,\pi)^2}\,\frac{\pi\,D_O^2}{2\,\sigma_m\,\omega_O}\,(1+\cos(\vartheta_{\mathbf{k},\mathbf{k}'-\mathbf{k}}+\vartheta_{\mathbf{k}',\mathbf{k}'-\mathbf{k}})) \tag{5}$$

$$\left|G^{(K)}(\mathbf{k}',\mathbf{k})\right|^2 = \frac{1}{(2\,\pi)^2}\,\frac{\pi\,D_K^2}{\sigma_m\,\omega_K}\,(1-\cos\vartheta_{\mathbf{k},\mathbf{k}'})\,, \tag{6}$$

where D_O is the optical phonon coupling constant, ω_O the optical phonon frequency, D_K is the K-phonon coupling constant and ω_K the K-phonon frequency. The angles $\vartheta_{\mathbf{k},\mathbf{k}'-\mathbf{k}}$ and $\vartheta_{\mathbf{k}',\mathbf{k}'-\mathbf{k}}$ denote the convex angles between $\mathbf{k}$ and $\mathbf{k}'-\mathbf{k}$ and between $\mathbf{k}'$ and $\mathbf{k}'-\mathbf{k}$, respectively.

3 THE NUMERICAL METHOD FOR THE TRANSPORT EQUATION

A numerical approach based on the discontinuous Galerkin method for solving the kinetic model described in Sec. 2 is used.

Since we are interested to the transport properties in a homogeneous suspended monolayer, we look for spatially homogeneous solutions to Eq. (1) with a constant electric field. The Boltzmann equation in the K valley reduces to

$$\frac{\partial f(\mathrm{t},\mathbf{k})}{\partial t}-\frac{e}{\hbar}\,\mathbf{E}\cdot\nabla_{\mathbf{k}} f(\mathrm{t},\mathbf{k}) = \int S(\mathbf{k}',\mathbf{k})\,f(\mathrm{t},\mathbf{k}')\,(1-f(\mathrm{t},\mathbf{k}))\,d\mathbf{k}' - \int S(\mathbf{k},\mathbf{k}')\,f(\mathrm{t},\mathbf{k})\,(1-f(\mathrm{t},\mathbf{k}'))\,d\mathbf{k}'\,. \tag{7}$$

A similar equation holds for the K' valley. As initial condition we take the Fermi-Dirac distribution

$$f(0,\mathbf{k}) = \frac{1}{1+\exp\left(\dfrac{\varepsilon(\mathbf{k})-\varepsilon_F}{k_B\,T}\right)},$$

where $T = 300\,K$, and ε_F is the Fermi energy, which is related to the initial charge density by

$$\rho(0) = \frac{2}{(2\,\pi)^2}\int f(0,\mathbf{k})\,d\mathbf{k}\,. \tag{8}$$

Eq. (7) is discretized by adopting a discontinuous Galerkin scheme. We choose a bounded domain $\Omega\subset\mathbb{R}^2$ such that $f(\mathrm{t},\mathbf{k})\approx 0$ for every $\mathbf{k}\notin\Omega$ and $\mathrm{t}>0$, and introduce a finite

decomposition $\{C_\alpha\}_{\alpha=1}^N$ of Ω, with C_α appropriate open sets, such that

$$C_\alpha \cap C_\beta = \emptyset \quad \text{if } \alpha \neq \beta, \text{ and } \quad \bigcup_{\alpha=1}^{N} \overline{C_\alpha} = \Omega\,.$$

The distribution function is assumed to be constant in each cell C_α. If we denote by $\chi_\alpha(\mathbf{k})$ the characteristic function over the cell C_α, then the approximation of the distribution function f is given by

$$f(\mathrm{t},\mathbf{k}) \approx f^\alpha(\mathrm{t}) \quad \forall\, \mathbf{k} \in C_\alpha \iff f(\mathrm{t},\mathbf{k}) \approx \sum_{\alpha=1}^{N} f^\alpha(\mathrm{t})\,\chi_\alpha(\mathbf{k}) \quad \forall\, \mathbf{k} \in \bigcup_{\alpha=1}^{N} C_\alpha\,.$$

This assumption replaces the unknown f, which depends on the two variables t and $\mathbf{k}$, with a set of N unknowns f^α, which depend only on time t. In order to obtain a set of N equations for the new unknowns f^α, we integrate Eq. (7) with respect to $\mathbf{k}$ over every cell C_α and replace f with its approximation. The derivative of f with respect to the time is treated easily. We have

$$\int_{C_\alpha} \frac{\partial f(\mathrm{t},\mathbf{k})}{\partial t}\, d\mathbf{k} \approx M_\alpha\, \frac{d\, f^\alpha(t)}{dt}$$

where M_α is the measure of the cell C_α. It is clear that the numerical method yields a system of ordinary differential equations by discretizing the collision operator and the drift term as discussed below.

Discretization of the collision operator.

Since for each $\mathbf{k} \in C_\alpha$

$$\begin{aligned}
&\int S(\mathbf{k}',\mathbf{k})\, f(\mathrm{t},\mathbf{k}')\,(1-f(\mathrm{t},\mathbf{k}))\, d\mathbf{k}' - \int S(\mathbf{k},\mathbf{k}')\, f(\mathrm{t},\mathbf{k})\,(1-f(\mathrm{t},\mathbf{k}'))\, d\mathbf{k}' \\
&\approx \sum_{\beta=1}^{N} \left[\int_{C_\beta} S(\mathbf{k}',\mathbf{k})\, f^\beta(\mathrm{t})\,(1-f^\alpha(\mathrm{t}))\, d\mathbf{k}' - \int_{C_\beta} S(\mathbf{k},\mathbf{k}')\, f^\alpha(\mathrm{t})\,\big(1-f^\beta(\mathrm{t})\big)\, d\mathbf{k}' \right] \\
&= \sum_{\beta=1}^{N} \left[f^\beta(\mathrm{t})\,(1-f^\alpha(\mathrm{t})) \int_{C_\beta} S(\mathbf{k}',\mathbf{k})\, d\mathbf{k}' - f^\alpha(\mathrm{t})\,\big(1-f^\beta(\mathrm{t})\big) \int_{C_\beta} S(\mathbf{k},\mathbf{k}')\, d\mathbf{k}' \right],
\end{aligned}$$

if we define

$$A^{\alpha,\beta} = \int_{C_\alpha} \left[\int_{C_\beta} S(\mathbf{k},\mathbf{k}')\, d\mathbf{k}' \right] d\mathbf{k}\,, \tag{9}$$

then we obtain

$$\begin{aligned}
&\int_{C_\alpha} \left[\int S(\mathbf{k}',\mathbf{k})\, f(\mathrm{t},\mathbf{k}')\,(1-f(\mathrm{t},\mathbf{k}))\, d\mathbf{k}' - \int S(\mathbf{k},\mathbf{k}')\, f(\mathrm{t},\mathbf{k})\,(1-f(\mathrm{t},\mathbf{k}'))\, d\mathbf{k}' \right] d\mathbf{k} \\
&\approx \sum_{\beta=1}^{N} \left[A^{\beta,\alpha}\,(1-f^\alpha(\mathrm{t}))\, f^\beta(\mathrm{t}) - A^{\alpha,\beta}\, f^\alpha(\mathrm{t})\,\big(1-f^\beta(\mathrm{t})\big) \right].
\end{aligned}$$

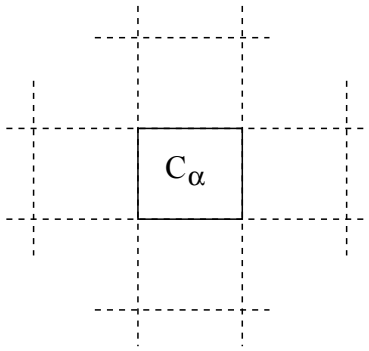


Figure 1: Cells employed for the numerical flux in the case of a simple rectangular grid.

So doing, the integral collision operator is replaced by quadratic polynomials. We note that the numerical coefficients $A^{\alpha,\beta}$ depend only on the scattering terms and the domain decomposition.

Discretizaton of the force term.

We must approximate the term

$$-\frac{e}{\hbar}\,\mathbf{E}\cdot\int_{C_\alpha}\nabla_{\mathbf{k}}f(\mathrm{t},\mathbf{k})\,d\mathbf{k}=\;-\frac{e}{\hbar}\,\mathbf{E}\cdot\int_{\partial C_\alpha}f(t,\mathbf{k})\,\mathbf{n}\,d\sigma$$

where $\mathbf{n}$ is the external unit normal to the boundary ∂C_α of the cell C_α. Since, due to the Galerkin method, the approximation of f is not defined on the boundary of the cells, we must introduce a *numerical flux*, that furnishes reasonable values of f on every ∂C_α, depending on the values of the approximation of f in the nearest neighborhood of the cell C_α and on the sign of $\mathbf{E}\cdot\mathbf{n}$. In Fig. 1 we show a simple picture of the cells that can be involved to find the numerical flux. The simplest numerical flux is given by the *upwind rule*, that uses only the four nearest adjacent cells.

4 CARRIER MOMENT EQUATIONS AND CLOSURE RELATIONS

Macroscopic quantities can be defined as moments of the distribution functions with respect to some suitable weight functions $\psi(\mathbf{k})$, assuming a sufficient regularity for the existence of the involved integrals. In particular for electrons and holes, we propose a set of moment equations consisting of the balance equations of the following quantities

$$\text{average density} \quad \rho_i = \frac{4}{(2\,\pi)^2}\int_{\mathbb{R}^2} f_i(\mathrm{t},\mathbf{x},\mathbf{k})\,d\mathbf{k},$$
$$\text{average velocity} \quad \rho_i\mathbf{V}_i = \frac{4}{(2\,\pi)^2}\int_{\mathbb{R}^2} f_i(\mathrm{t},\mathbf{x},\mathbf{k})\,\mathbf{v}\,d\mathbf{k},$$
$$\text{average energy} \quad \rho_i W_i = \frac{4}{(2\,\pi)^2}\int_{\mathbb{R}^2} f_i(\mathrm{t},\mathbf{x},\mathbf{k})\,\varepsilon\,d\mathbf{k},$$
$$\text{average energy-flux} \ \rho_i\mathbf{S}_i = \frac{4}{(2\,\pi)^2}\int_{\mathbb{R}^2} f_i(\mathrm{t},\mathbf{x},\mathbf{k})\,\varepsilon\mathbf{v}\,d\mathbf{k},$$

(i = electron, hole), where the factor 4 arises from taking into account both the spin states and the two equivalent valleys.

By integrating the Boltzmann equations with respect to $\mathbf{k}$, one has the following balance equations for the above-defined macroscopic quantities

$$\frac{\partial}{\partial t}\rho_i + \nabla_{\mathbf{r}}\cdot(\rho_i\,\mathbf{V}_i) = \rho_i\,C_i, \tag{10}$$
$$\frac{\partial}{\partial t}(\rho_i\,\mathbf{V}_i) + \nabla_{\mathbf{r}}\cdot\left(\rho_i\,\mathbf{F}_i^{(0)}\right) + \mathsf{e}_i\,\rho_i\mathbf{G}_i^{(0)}\mathbf{E} = \rho_i C_{\mathbf{V}_i}, \tag{11}$$
$$\frac{\partial}{\partial t}(\rho_i W_i) + \nabla_{\mathbf{r}}\cdot(\rho_i\,\mathbf{S}_i) + \mathsf{e}_i\rho_i\mathbf{E}\cdot\mathbf{V}_i = \rho_i C_{W_i}, \tag{12}$$
$$\frac{\partial}{\partial t}(\rho_i\,\mathbf{S}_i) + \nabla_{\mathbf{r}}\cdot\left(\rho_i\mathbf{F}_i^{(1)}\right) + \mathsf{e}_i\rho_i\mathbf{G}_i^{(1)}\mathbf{E} = \rho_i C_{\mathbf{S}_i}, \tag{13}$$

where the G's and F's are extra-fluxes and the terms at the right hand sides are productions (the reader is referred to [7] for details) and e_i is equal to e for electrons and $-e$ for holes.

The extra fluxes and the production terms are additional unknown quantities. For them constitutive relations in terms of the fundamental variables are needed in order to get a closed system of balance equations. A well theoretically founded way to get the desired closure relations is to resort to the Maximum Entropy Principle (MEP) [11], according to which the electron and hole distribution functions can be estimated by the distributions $f_{e,MEP}$ and $f_{h,MEP}$ solving the following problem

$$(f_{e,MEP}, f_{h,MEP}) = \max_{f_e(\mathrm{t},\mathbf{x},),f_h(\mathrm{t},\mathbf{x},)\in\mathcal{F}(\mathbb{R}^2)} S[f_e, f_h],$$

under the constraints

$$\begin{pmatrix}\rho_i\\ \rho_i W_i\end{pmatrix} = \frac{4}{(2\,\pi)^2}\int_{\mathbb{R}^2}\begin{pmatrix}1\\ \varepsilon\end{pmatrix} f_i(\mathrm{t},\mathbf{x},\mathbf{k})\,d\mathbf{k},$$
$$\begin{pmatrix}\rho_i\mathbf{V}_i\\ \rho_i\mathbf{S}_i\end{pmatrix} = \frac{4}{(2\,\pi)^2}\int_{\mathbb{R}^2} f_i(\mathrm{t},\mathbf{x},\mathbf{k})\begin{pmatrix}\mathbf{v}\\ \varepsilon\mathbf{v}\end{pmatrix} d\mathbf{k},$$

where $S[f_e, f_h]$ is the total entropy of the system (remind that the phonons are assumed to represent a thermal bath kept at constant temperature and therefore they add a constant contribution to the entropy) given by

$$-k_B \left\{ \frac{4}{(2\pi)^2} \int_{\mathbb{R}^2} \left[f^e \ln f^e + (1-f^e) \ln (1-f^e) \right] d\,\mathbf{k} + \frac{4}{(2\pi)^2} \int_{\mathbb{R}^2} \left[f^h \ln f^h + \right. \right.$$
$$\left. \left. \left(1-f^h\right) \ln \left(1-f^h\right) \right] d\,\mathbf{k} \right\},$$

$\mathcal{F}(\mathbb{R}^2)$ being the space of the distribution functions that admit the moments required as constraints.

By solving the above maximization problem we get

$$f_i = \frac{1}{1 + \exp\left(\lambda_i + \lambda_{W_i}\varepsilon + \mathbf{v} \cdot \left(\lambda_{\mathbf{V}_i} + \varepsilon \lambda_{\mathbf{S}_i}\right)\right)}.$$

As in [12] we linearize the distributions around their isotropic part, obtaining

$$f_i \;\approx\; \frac{1}{e^{\lambda_i + \lambda_{W_i}\varepsilon} + 1} \left[1 - \frac{e^{\lambda_i + \lambda_{W_i}\varepsilon}}{e^{\lambda_i + \lambda_{W_i}\varepsilon} - 1}\, \mathbf{v} \cdot \left(\lambda_{\mathbf{V}_i} + \varepsilon \lambda_{\mathbf{S}_i}\right) \right],$$

where the λ's are Lagrange multipliers which have to be expressed as functions of the state variables by taking into account the constraints.

After that, these distributions are inserted into the kinetic definitions of the additional variables, so closing the system of the balance equations (see [7] for the details).

5 NUMERICAL SOLUTIONS

We want to assess the validity of the MEP hydrodynamical model by a comparison with the solutions furnished by the direct integration of the transport equation. For solving the latter let us consider a circle as domain Ω. A TVD third order Runge-Kutta scheme is used to solve the resulting ODE system similarly to Ref. [13]. We remark that the numerical scheme guarantees the mass conservation.

The Fermi level is set equal to 0.4 eV, a value high enough for neglecting the inter-band interactions, and the lattice constant is kept equal to 300 K. In the literature there are several values for the coupling constants entering into the collision terms. For example for the acoustic deformation potential one can find values ranging from 2.6 eV to 29 eV. Similar degree of uncertainty is found for the optical and K phonon coupling constants as well.

We have performed numerical simulations of a suspended monolayer graphene by considering the parameters used in [14]. The solutions do not depend on $\mathbf{x}$ and therefore we neglect the terms in divergence form in the balance equations (10)-(13), that become a

system of ODEs. Moreover only the component of the velocity and the energy-flux along the direction of the electric field, which we assume to be the x axis, is changing with time if we set the initial velocity equal to zero. Regarding the initial conditions of the other macroscopic variables, consistently with an initial Fermi-Dirac distribution for Eq. (7), we assume zero energy-flux while the initial density and the average energy density are calculated from the initial Fermi-Dirac distribution.

Eq. (7) and the system (10)-(13) have been solved for different values of the applied electric field and the results for the average velocity and the energy are shown in Fig.s 2, 3.

Regarding the average energy the results of the MEP model are quite satisfactory. In the steady regime the maximum relative error is about 5 % and is reached when $E = 10$ kV/cm. In the other case we have a relative error of 1.4 % for $E = 2$ kV/cm, 2.4 % for $E = 4$ kV/cm, 3 % for $E = 20$ kV/cm. Note that the error is not monotone with respect to the electric field.

The behavior of the relative error for the velocity is different. This latter, instead, is decreasing by increasing the applied field. The discrepancy has a maximum for $E = 2$ kV/cm of about 32 %. For $E = 4$ kV/cm the relative error is 28 %, for $E = 10$ kV/cm 19 %, for $E = 20$ kV/cm 4 %.

In order to understand if the Fermi energy influences the accuracy of the MEP model, we have performed the same simulations with $\varepsilon_F = 0.6$ eV. The qualitative behaviour is similar to the case $\varepsilon_F = 0.4$ eV. Again one finds a relative error for the energy not greater than 5 %. Instead the relative error of the velocity is higher: 37 % if $E = 2$ kV/cm, 32 % if $E = 4$ kV/cm, 30 % $E = 10$ kV/cm, 24 % $E = 20$ kV/cm.

Although the overall discrepancy is reasonable for the applications, it is likely that one needs to include some nonlinear terms in the velocity and the energy-flux or additional moments. Apparently the MEP model under consideration is not able to give a correct equilibrium limit. This is an open problem we are working on.

Acknowledgements

The authors acknowledge the financial support by the project FIR 2014, University of Catania.

REFERENCES

[1] Castro Neto, A.H., Guinea, F., Peres, N.M.R., Novoselov, K.S., Geim, A.K. The electronic properties of graphene. *Rev. of Mod. Phys.* (2009) **81**:109-162.

[2] Shishir, R. S. and Ferry, D.K. Velocity saturation in intrinsic graphene. *J. Phys. Condens. Matter* (2009) **21**:344201 .

[3] Fang, T., Konar, A., Xing, H. and Jena, D. High-field transport in two-dimensional graphene. *Phys. Rev. B* (2011) **84**:125450.

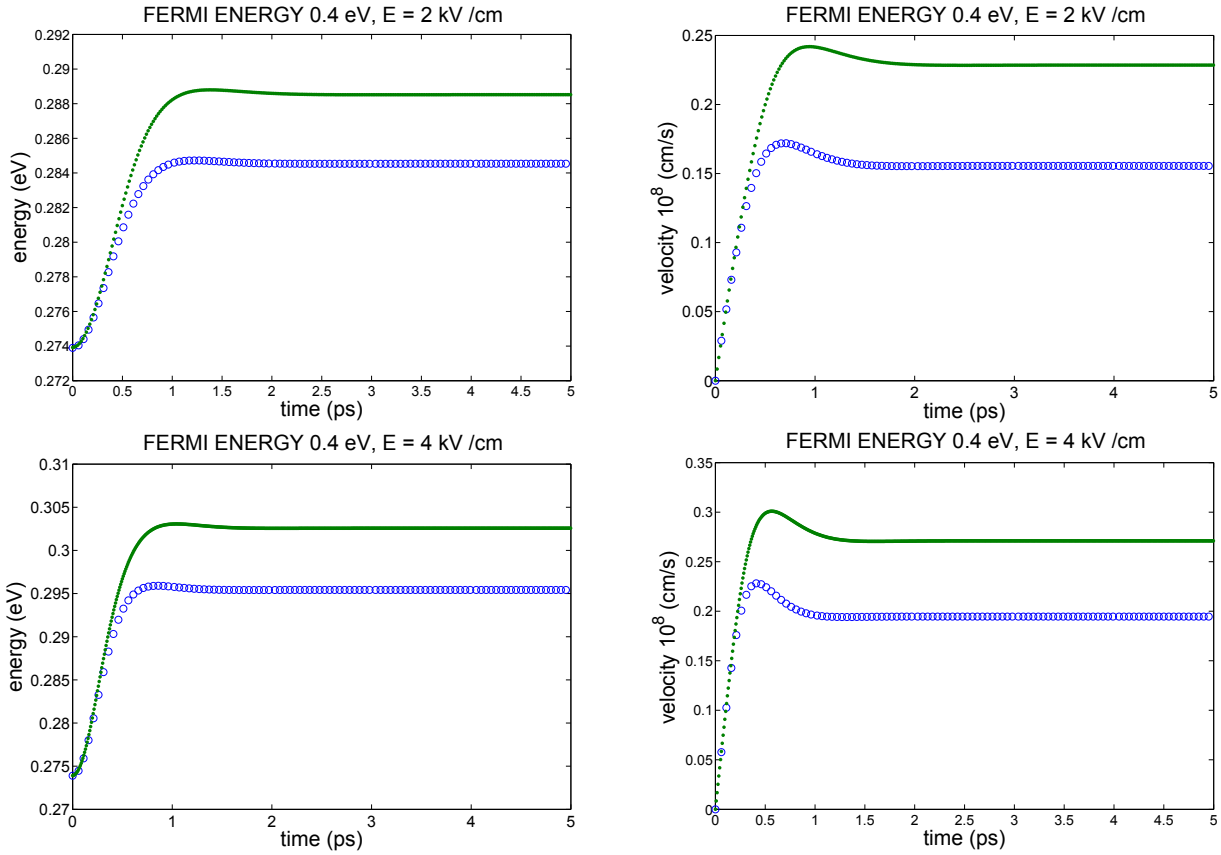


Figure 2: Comparison of the energy and the velocity obtained with a direct solution of the Boltzmann equation (points) and the MEP hydrodynamical model (circles) for the electric fields $E = 2$ kV/cm, $E = 4$ kV/cm by using the same values of the scattering parameters as in [14], by considering a constant lattice temperature of 300 K and a Fermi energy equal to 0.4 eV.

[4] Tomadin, A., Brida, D., Cerullo, G., Ferrari, A.C. and Polini, M. Nonequilibrium dynamics of photoexcited electrons in graphene: Collinear scattering, Auger processes, and the impact of screening. *Phys. Rev. B* (2013) **88**:035430.

[5] Lichtenberger, P., Morandi, O. and Schürrer, F. High-field transport and optical phonon scattering in graphene. *Physical Review B* (2011) **84**:045406.

[6] Zamponi, N. and Barletti, L. Quantum electronic transport in graphene: a kinetic and fluid-dynamical approach. *Math. Methods Appl. Sci.* (2011) **34**:807-818 .

[7] Camiola, V.D. and Romano, V. Hydrodynamical model for charge transport in graphene. *Journal of Statistical Physics* (2014) **157**:11141137.

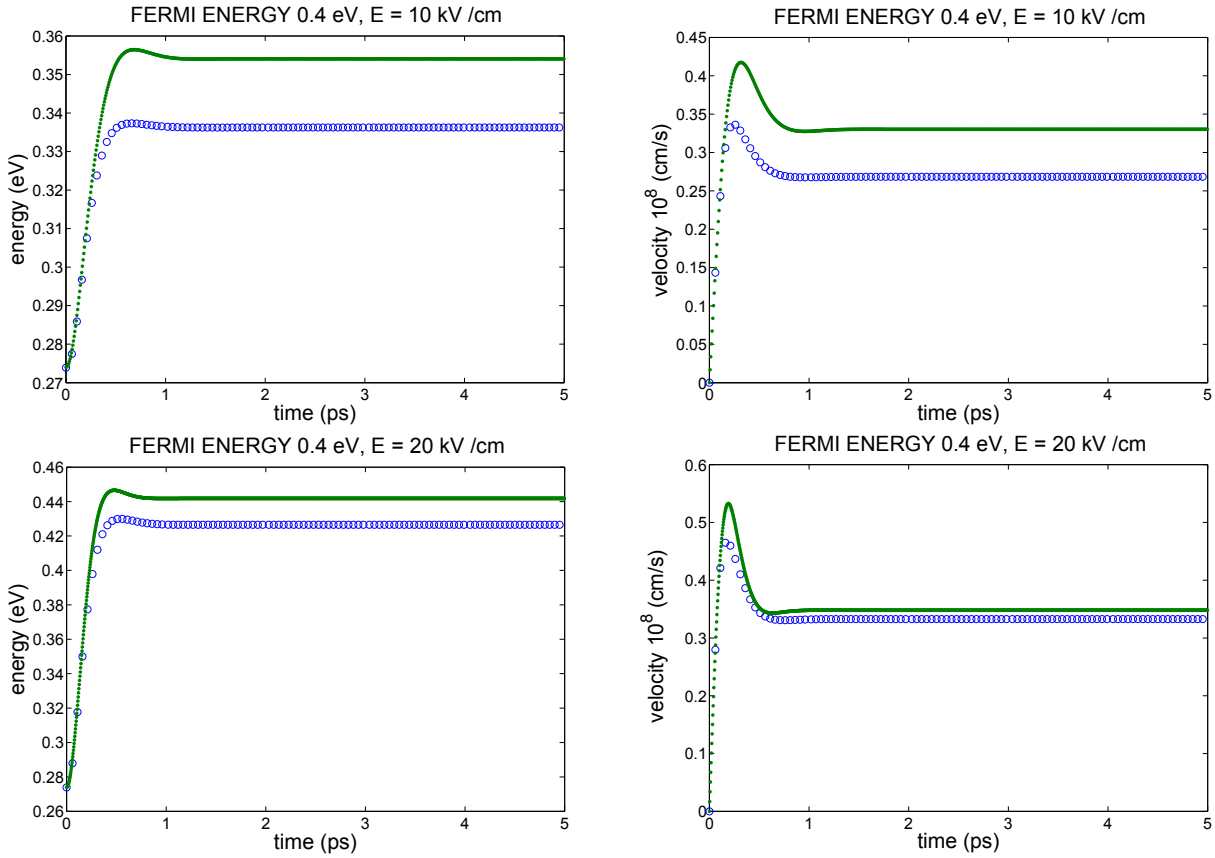


Figure 3: Comparison of the energy and the velocity obtained with a direct solution of the Boltzmann equation (points) and the MEP hydrodynamical model (circles) for the electric fields $E = 10$ kV/cm, $E = 20$ kV/cm by using the same values of the scattering parameters as in [14], by considering a constant lattice temperature of 300 K and a Fermi energy equal to 0.4 eV.

[8] Mascali, G. and Romano, V. A comprehensive hydrodynamical model for charge transport in graphene. 978-1-4799-5433-9/14/$31.00 © 2014 IEEE, IWCE-2014 Paris (2014).

[9] Cheng, Y., Gamba, I.M., Majorana, A., Shu, C.W. A discontinuous Galerkin solver for Boltzmann-Poisson systems in nano devices. *Computer Methods in Applied Mechanics and Engineering* (2009) **198 (37-40)**:3130–3150.

[10] Cheng, Y., Gamba, I.M., Majorana, A., Shu, C.W. A brief survey of the discontinuous Galerkin method for the Boltzmann-Poisson equations. *Boletin de la Sociedad Espanola de Matematica Aplicada* (2011) **54**:47-64 .

[11] Jaynes, E.T. Information Theory and Statistical Mechanics. *Phys. Rev. B* (1957) **106**:620-630.

[12] Camiola, V.D., Mascali, G., Romano, V. Simulation of a Double-Gate MOSFET by a Nonparabolic Hydrodynamical Subband Model for Semiconductors Based on the Maximum Entropy Principle. *Mathematical and Computer Modelling* (2013) **58**:321-343.

[13] Galler, M. and Majorana, A. Deterministic and Stochastic Simulations of Electron Transport in Semiconductors. *Bulletin of the Institute of Mathematics - Academia Sinica* - New Series (2007) **2**:349-365.

[14] Borysenko, K.M., Mullen, J.T., Barry, E.A., Paul, S., Semenov, Y.G., Zavada, J.M., Buongiorno Nardelli, M. and Kim, K.W. First-principles analysis of electron-phonon interactions in graphene. *Phys. Rev. B* (2010) **11**:121412(R).

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis(Eds)

ELECTRO-THERMO-MECHANICAL ANALYSIS OF SMA ACTUATOR

VLADIMÍR KUTIŠ*, PETER DRAHOŠ*, JUSTÍN MURÍN*, JURAJ PAULECH* AND JURAJ HRABOVSKÝ*

*Institute of Automotive Mechatronics
Faculty of Electrical Engineering and Information Technology
Slovak University of Technology in Bratislava
81219 Bratislava, Slovak Republic
e-mail: vladimir.kutis@stuba.sk, web page: http://www.fei.stuba.sk/

Key words: Shape Memory Alloy, Electro-Thermal Analysis, Structural Analysis, Phase Change, Finite Element Analysis

Abstract. The paper deals with electro-thermal and structural analysis, modeling and simulation of the Shape Memory Alloy (SMA) actuator made of Nickel-Titanium alloy. Primary goal is focused on electro-thermal analysis and investigation of SMA actuator temperature distribution near mechanical connection. The secondary goal is focused on actuation deformation and the influence of nonhomogenous temperature distribution on the deformation of SMA actuator. All analyzes are performed by finite element code ANSYS.

1 INTRODUCTION

Shape Memory Alloy (SMA) is a smart material converting thermal energy to mechanical work, which can be used in many different mechatronic systems like actuators as well as sensors [1, 2]. Transformation behaviour exploited in SMA material is thermally induced shape change, often labeled the shape-memory effect. A material component may be deformed or strained at low temperatures and when heated, it reverses this strain and remembers its prestrained shape. Deformable martensite phase transforms to a more stable austenite phase at higher temperatures. Thermally activated SMA applications require temperature control to optimize the effect of shape memory. Typical application areas of SMA are aerospace industry – composite structures that have SMA wires embedded can be used to change the shape of an airplane wing; vibration damper systems – vibrational dampers, comprised of composite materials using prestrained, embedded SMA wire; medical applications – medical applications use the superelastic property of SMA, and many of them are in the expanding field of minimally invasive surgery.

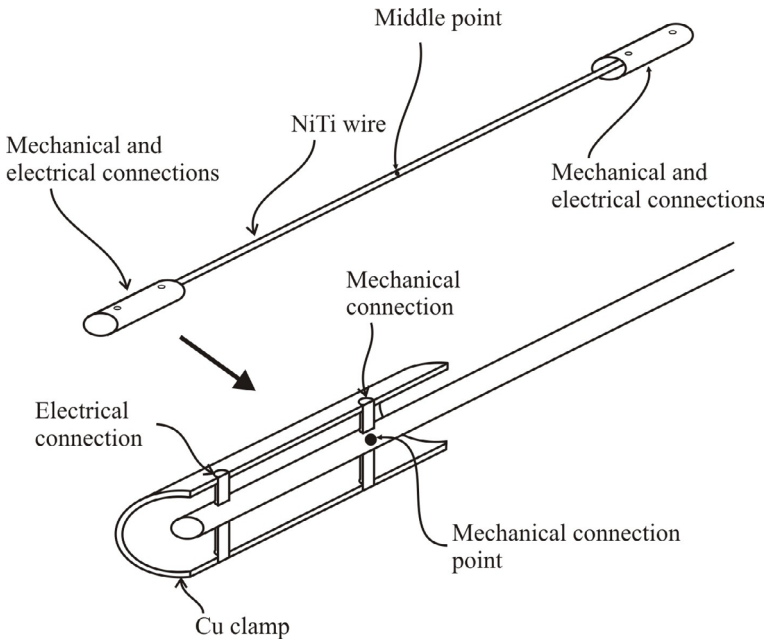


Figure 1: SMA actuator with clamp

The paper deals with coupled electro-thermo-structural analysis of thermal SMA wire actuator made of Nickel-Titanium (NiTi) alloy in FEM code ANSYS [3] – see Fig 1. The main idea of thermal SMA actuators is to cross transformation temperature from martensite phase to austenite phase and than required actuation deformation of SMA actuator is induced by this phase change. Both phases from electro-thermal viewpoint have different electrical and thermal material properties and there is also latent heat during phase transformation. Phase change is not performed at one temperature but there is temperature range where the one phase is transformed to other. Start and finish temperature of phase change depends on mechanical stress of NiTi wire as well as on increasing or decreasing of temperature, i.e. heating or cooling of NiTi wire.

Investigated SMA NiTi wire actuator is loaded by external force which causes axial stress in NiTi wire with value 200 MPa. The transformation of martensite phase to austenite phase (i.e heating process) in investigated SMA wire under tensile stress 200 MPa starts at temperature 119 °C and finishes at temperature 129 °C. In the reverse cooling process, where austenite phase transforms to martensite phase, the transformation under the same stress conditions starts at temperature 81 °C and finishes at temperature 69 °C. In martansite phase under transformation temperature and under prescribed mechanical loading, the strain of SMA wire has value 4.8% and in austenite phase (i.e over the transformation temperature) the strain has value 0.4%. These values in real system and in modeled approximation in material model are shown in Fig. 2.

NiTi actuator is loaded through mechanical connection and actuation motion is performed by heating up of actuator. Active length of NiTi wire is length between two

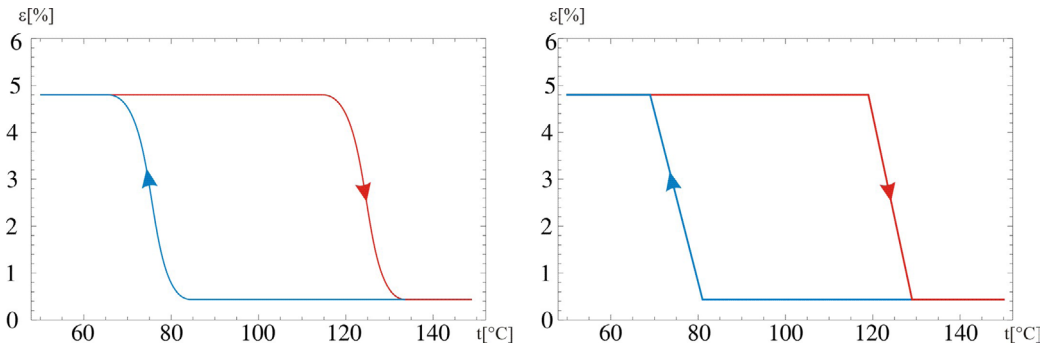


Figure 2: Transformation temperatures during the phase change, Left – real behaviour, Right – modeled approxiamtion in material model

mechanical connections – see Fig. 1. SMA wire is heated by Joule loss heat caused by electric current that is passing through the NiTi wire through clamp and electrical connection. System is naturally cooled by surrounding air.

Critical part of SMA actuator from thermal and also mechanical point of view is mechanical connection of NiTi wire to clamp. This part of NiTi wire is cooled more than the rest of the wire due to connections to clamp. It is necessary to investigate the influence of actuator clamps and specially mechanical connection to temperature distribution along NiTi wire and its effect on actuation deformation of SMA wire, because unsuitably designed connection can cause loss of power or smaller actuation deformation and other characteristics in this part of SMA actuator.

2 ELECTRO-THERMAL ANALYSIS

The goal of steady-state electro-thermal analysis is to calculate the spatial temperature distribution on SMA actuator and special attention is paid to temperature distribution in NiTi wire near mechanical connection, where the influence of different thermal conductivities of mechanical connection is analysed – see Fig. 1. Also transient electro-thermal analysis is performed in order to investigate thermal dynamics of SMA actuator.

2.1 Geometry and material parameters

Geometry of NiTi wire with clamp is shown in Fig. 1. In this figure, there is also shown mechanical and electrical connections.
NiTi wire has diameter 0.3 mm and the active length of actuator is 100 mm – this is the length between two mechanical connections. The length of clamp is 11 mm and the inner and the outer diameter of clamp is 1.5 mm and 2 mm, respectively. Screw diameter of mechanical and electrical connection is 0.4 mm.
Considered electrical and thermal material parameters of NiTi wire, which are function of temperature are defined in Tab. 1. Because not only steady-state but also transient

Table 1: Electrical and thermal material parametrs of NiTi wire as function of temperature

Martensite to Austenite [°C] (Increasing temperature)	0	119	129	180
Austenite to Martensite [°C] (Decreasing temperature)	0	69	81	180
Thermal conductivity [W/mK]	8	8	18	18
Electrical resistivity$\times 10^8$ [Ωm]	76	76	82	82

analysis was performed, thermal capacity and density of NiTi wire has to be included into material model. Both material properties were considered as constant in all investigated thermal conditions. Thermal capacity of NiTi wire has value 460 J/kgK and density of NiTi wire is 6540 kg/m^3.
All other components – clamp, screws for mechanical and electrical connections were considered with constant material properties – see Tab. 2. Clamp and screws for electrical connection is made of copper. In order to investigate the influence of mechanical connection, thermal conductivity of screws for mechanical connection is considered in some range – from 0.5 W/mK to 2 W/mK. Other thermal and electrical material parameters of screws for mechanical connection were compatible with material properties of teflon.

Table 2: Electrical and thermal material parametrs of clamp and connections

Component	Material	Thermal conductivity [W/mK]	Electrical resitivity [Ωm]	Thermal capacity [J/kgK]	Density [kg/m^3]
Clamp	copper	385	1.67×10^{-8}	396	8900
Mechanical connection		0.5/1/2	1×10^{12}	1000	2100
Electrical connection	copper	385	1.67×10^{-8}	396	8900

In the inner space of clamp, there was considered air with constant thermal properties – thermal conductivity 0.02 W/mK, thermal capacity 700 J/kgK and density 1.2 kg/m^3.

2.2 FEM simulation

Electro-thermal steady-state and transient analysis of SMA actuator was performed by code ANSYS Multiphysics. Due to symmetry of actuator and connections only 1/8 of geometry model was considered - see Fig. 3 Top. There were used 2 element types, 3D solid element for coupled analysis SOLID 226, where elctro-thermal capabilities were set up and 2D surface effect element for convection and radiation heat transfer SURF 152.
Boundary conditions were prescribed as follows:

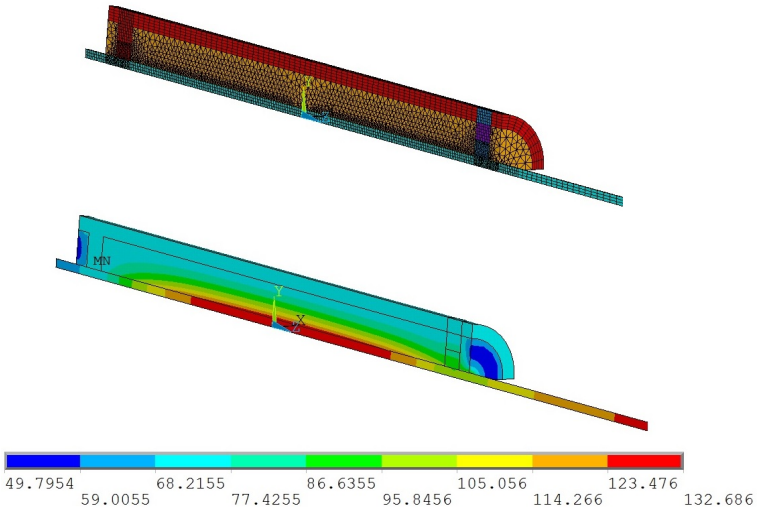


Figure 3: Top – mesh of SMA actuator, Bottom – temperature distribution near mechanical connection

- electrical boundary conditions – electric current, that flows from clamp through electrical connection to NiTi wire had value 0.8 A and in transient analysis starts at time 0.01 s
- thermal boundary conditions – convection was prescribed in all outer surfaces of NiTi wire and clamp. The coefficient of heat transfer by convection was calculated analytically separately for outer cylindrical surface of NiTi wire and separately for clamp. All analytical calculations were performed for horizontal configuration of SMA actuator using dimensionless parameters. Calculated coefficients of heat transfer were included in the model as a function of surface temperature. Ambient temperature was 27 °C.

In steady-state thermal analysis, three analyses with different thermal conductivities of mechanical connection were performed. Longitudinal distributions of temperature in active part of NiTi wire for all three thermal conductivities of mechanical connection are shown in Fig. 4 Left. As we can see from this figure, if material of mechanical connection has thermal conductivity 2 W/mK or 1 W/mK, not all points of NiTi wire reach starting transformation temperatures. If we consider that thermal conductivity of mechanical connection is 2 W/mK, than almost 16% of NiTi wire length does not cross finish phase change temperature. The spatial temperature distribution in NiTi wire and also in clamp and in connections for thermal conductivity of mechanical connection with value 2 W/mK is shown in Fig. 3 Bottom. As we can see from the Fig. 3 Bottom and from Fig. 4 Left, mechanical connection can strongly affect of NiTi wire temperature near mechanical connection.

Fig. 4 Right and Fig. 5 Left and Right shows thermal dynamics of NiTi actuator. As we can see from all three graphs, after 30 seconds the middle point of NiTi wire (see Fig. 1)

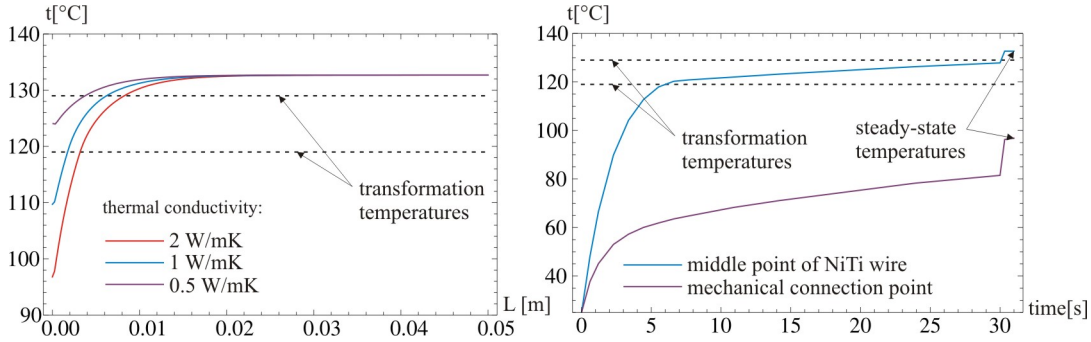


Figure 4: Left – Longitudinal distributions of temperature in active part of NiTi wire for all three thermal conductivities of mechanical connection, Right – thermal dynamics of NiTi wire with thermal conductivity of mechanical connection 2 W/mK

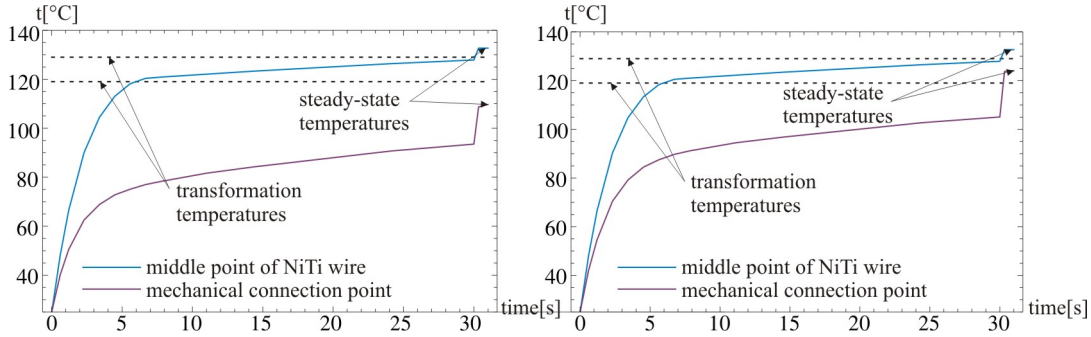


Figure 5: Left – thermal dynamics of NiTi wire with thermal conductivity of mechanical connection 1 W/mK, Right – thermal dynamics of NiTi wire with thermal conductivity of mechanical connection 0.5 W/mK

almost reaches the finish phase change temperature for all three investigated thermal conductivities of mechanical connection, but NiTi wire near mechanical connection does not cross even start phase change temperature after 30 second.

3 STRUCTURAL ANALYSIS

The goal of static structural analysis is to calculate actuation deformation of SMA thermal actuator, that is caused by phase change induced by increasing temperature. As was mentioned in chapter 2, only material, whose temperature cross transformation temperature transforms its phase and only this part of actuator works effectively. All three longitudinal temperature distributions in NiTi wire, which are shown in Fig. 4 Left, are considered.

3.1 Material parameters

In FEM code ANSYS, Souza-Auricchio material model of shape memory alloy with shape memory effect is implemented [4]. The model contains six parameters, which has

to be determined by experiment.
Our SMA wire can be described by following material parameters: Young modulus of austenite phase 45×10^9 Pa, Poisson ratio 0.3, hardening parameter 1000×10^6 Pa, temperature below which no twinned martensite is observed 343 K, elastic limit 110×10^6 Pa, material parameter β 4.49×10^6 Pa/K and maximal transformation strain 0.06.

3.2 FEM simulation

Static structural analysis of SMA actuator was performed by FEM code ANSYS Multiphysics. Because only NiTi wire is analyzed, 1D beam element BEAM 188, that supports SMA model with shape memory effect, was used. Only half of NiTi wire geometry was modeled (NiTi wire between middle and mechanical connection point). Left end of beam model, which represents mechanical connection of NiTi wire, was supported. At the right end of beam model there was prescribed force 14.1372 N, which causes tension stress in NiTi wire with value 200 MPa.

In order to simulate actuation deformation of SMA actuator, each simulation has to contain two loading states. In the first state, NiTi wire is loaded by structural force. The temperature considered in the first loading state is equal surrounding temperature. In the second state, temperature distribution, that was calculated in 3D electro-thermal analysis, is included into the model as thermal loading. This temperature distribution causes phase change in the NiTi wire locations, where temperature cross transformation temperature. The result of the phase change is the actuation deformation of SMA wire.

The longitudinal displacement of SMA actuator middle point caused by external force after the first loading state has value 2.4 mm. The middle point longitudinal actuation displacements of SMA wire after the second loading state for different temperature distributions are shown in Tab. 3.

Table 3: Dependence of SMA middle point actuation displacement on temperature distribution

Temperature distribution for mechanical connection with:	Actuation displacement of SMA middle point [mm]
thermal conductivity 0.5 W/mK	2.1
thermal conductivity 1 W/mK	1.9
thermal conductivity 2 W/mK	1.8

Fig. 6 Left and Right shows stress-strain relations for middle point and mechanical connection point with thermal conductivity of mechanical connection 0.5 W/mK, respectively. As we can see from this figure, actuation deformation of mechanical connection point is smaller than middle point deformation. This smaller deformation is caused by the fact, that the temperature of mechanical connection point does not cross finish phase change transformation temperature – this point has temperature only 124 °C and only

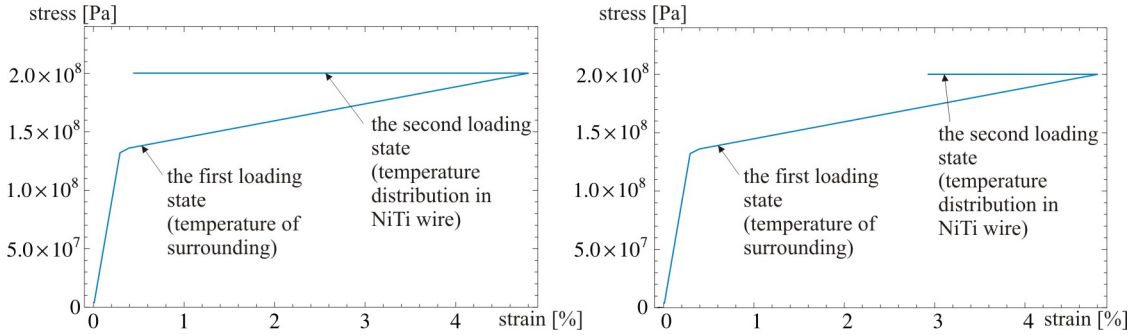


Figure 6: Stress-strain relations for following NiTi wire points: Left – middle point, Right – mechanical connection point, thermal conductivity of mechanical connection is 0.5 W/mK

adequate part of material is transformed from martensite phase to austenite phase.

4 CONCLUSIONS

Coupled electro-thermo-structural analysis of thermal SMA wire actuator made of Nickel-Titanium alloy was presented in the paper. The primary goal of the paper was focused on analysis of temperature distribution near mechanical connection of NiTi wire. In the paper steady-state and transient electro-thermal analysis of NiTi wire was presented. The secondary goal of the paper was focused on structural analysis of NiTi wire and on the investigation of influence of temperature distribution on actuation deformation of SMA actuator. Our next research will be focused on shape optimization of mechanical connection of NiTi wire.

ACKNOWLEDGEMENTS

This work was supported by the Slovak Research and Development Agency under the contract No. APVV-0246-12, by Grant Agency VEGA, grant No. 1/0228/14 and 1/0453/15 and by Grant Agency KEGA, grant No. 007STU-4/2015.

REFERENCES

[1] Pons, J. L. Emerging Actuator Technologies – A Micromechatronic Approach, John Wiley&Sons, (2005).

[2] Schwartz, M. Encyclopedia of Smart Materials, John Wiley&Sons, (2002).

[3] ANSYS Theory manual, (2014).

[4] Auricchio, F., Scalet, G. and Urbano, M. A Numerical/Experimental Study of Nitinol Actuator Springs. *Journal of Materials Engineering and Performance* (2014) **23**:2420–2428.

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis (Eds)

FINITE ELEMENT ANALYSIS USING A HIERARCHAL DECOMPOSITION FOR THE INTERACTION OF STRUCTURE, FLUID AND ELECTROSTATIC FIELD IN MEMS

D. ISHIHARA[*], T. HORIE[*], T. NIHO[*] AND A. BABA[†]

[*] Department of Mechanical Information Science and Technology
Kyushu Institute of Technology
680-4 Kawazu, Iizuka, Fukuoka 820-8502, Japan
e-mail: ishihara@mse.kyutech.ac.jp

[†]Center for Microelectronic System (CMS)
Kyushu Institute of Technology
680-4 Kawazu, Iizuka, Fukuoka 820-8502, Japan

Key words: Structure-Fluid-Electric Interaction, Decomposition, Partitioning, Splitting, Finite Element Analysis, Micro-Electro-Mechanical System (MEMS).

Abstract. In this study, a hierarchal decomposition for the interaction of the structure, fluid and electrostatic field or the structure-fluid-electrostatic interaction, which is one of typical phenomena in micro-electro-mechanical system (MEMS), is proposed in order to solve it efficiently. The proposed decomposition partitions the structure-fluid-electrostatic interaction into the fluid-structure interaction (FSI) and the electrostatic field, and, moreover, splits the FSI into the velocity and fluid pressure fields. In this way, the whole interaction system is decomposed into the three fields in a hierarchal way. The proposed decomposition is implemented using a finite element method and is applied to a micro cantilever beam actuated by the electrostatic force in air. It follows from the comparison among the results for the structure-fluid-electrostatic interaction, the FSI and the experiment that the proposed method taking into account the full-interaction can predict the vibration characteristic of the MEMS accurately.

1 INTRODUCTION

MEMS are typically smaller than 1 millimeter and larger than 1 micrometer in size. At these size scales, the surface forces are superior to the body forces due to the scale effect. Therefore, MEMS are typically driven by the electrostatic force, and their vibrations are damped significantly due to the fluid viscous force from air. In addition, both the electrostatic and fluid viscous forces are sensitive to the structural behavior. Therefore, the structure-fluid-electrostatic interaction is one of typical phenomena in MEMS. Vibration characteristics such as the resonance frequency and the damping ratio are the key design parameters, and the interaction analysis is required in order to predict them accurately [1, 2].

In this study, a finite element analysis for the structure-fluid-electrostatic interaction is proposed. The proposed analysis consists of the partitioned analysis, where the structure-fluid-electrostatic interaction is partitioned into the FSI and the electrostatic field, and the

monolithic analysis for the FSI. Moreover, the monolithic FSI system is solved using a projection method [5, 6, 7], where the FSI is split into the velocity and fluid pressure fields. Its application can be seen in [8, 9]. The decomposition used in the proposed analysis consists of the partitioning and the splitting in a hierarchal way. Therefore, it is called as the hierarchal decomposition.

The proposed analysis is applied to the micro cantilever beam driven by the electrostatic force in air. It is demonstrated from the comparison among the results for the structure-fluid-electrostatic interaction, the FSI, and the experiment that the proposed analysis taking into account the full interaction can predict the dynamic behavior of MEMS accurately.

2 FINITE ELEMENT ANALYSIS USING HIERARCHAL DECOMPOSITION

2.1 Finite element discretized equations

The finite element discretized equilibrium equation system for the structure-fluid-electrostatic interaction is schematically written as

$$\mathbf{Q}^{\mathrm{s}} + \mathbf{Q}^{\mathrm{f}} = \mathbf{g}^{\mathrm{e}}, \tag{1}$$

where $\mathbf{Q}$ is the equivalent internal force vector including all effects such as the structural and fluid inertia, the structural elasticity, the fluid diffusion, and the fluid pressure gradient, $\mathbf{g}$ is the external force vector, and the superscripts s, f and e denote the quantities corresponding to the structure, the fluid, and the electrostatic field, respectively. Since MEMS can undergo large deformation, the finite deformation is taken into account using the total Lagrangian formulation, where the Hooke's law is used for the relation between the second Piola-Kirchhoff stress and the Green-Lagrange strain under the assumption of small strain. The fluid is assumed to be Newtonian.

The electrostatic force $\mathbf{g}^{\mathrm{e}}$ is the nonlinear function with respect to the structural displacement vector $\mathbf{u}^{\mathrm{s}}$. In the present study, $\mathbf{g}^{\mathrm{e}}$ is obtained as follows: The electrostatic potential $\boldsymbol{\varphi}$ is obtained from the Laplace equation with the boundary including the structural surface. Next, the electrostatic field $\mathbf{E}$ is given by the gradient of $\boldsymbol{\varphi}$. Finally, $\mathbf{g}^{\mathrm{e}}$ is obtained as

$$\mathbf{g}^{\mathrm{e}} = -(\varepsilon/2)(\mathbf{E} \cdot \mathbf{n})^2\mathbf{n}, \tag{2}$$

where ε is the dielectric constant, $\mathbf{n}$ is the unit vector outward normal to the structural surface.

2.2 Coupled analysis using hierarchal decomposition

In order to partition the structure-fluid-electrostatic interaction into the FSI and the electrostatic field, Eq. (1) is evaluates as

$${}^{t+\Delta t}\mathbf{Q}^{\mathrm{s}\,(k+1)} + {}^{t+\Delta t}\mathbf{Q}^{\mathrm{f}\,(k+1)} = \mathbf{g}^{\mathrm{e}}\,({}^{t+\Delta t}\mathbf{u}^{\mathrm{s}\,(k)}), \tag{3}$$

where the superscript $t+\Delta t$ denotes the current time step, and the superscript k with the round brackets denotes the iteration at the current time step. After deriving the electrostatic force $\mathbf{g}^{\mathrm{e}}$ using the previous structural displacement ${}^{t+\Delta t}\mathbf{u}^{\mathrm{s}\,(k)}$, Eq. (3) can be solved monolithically.

Moreover, the FSI is split into the velocity and pressure fileds using a projection method [5, 6, 7]. In this way, the structure-fluid-electrostatic interaction is decomposed into the three fields in a hierarchal way. Therefore, the proposed decomposition is called as the hierarchal decomposition.

The predictor multi-corrector algorithm based on the Newmark's method is used for the time integration. Fig. 1 shows the analysis flow of the proposed analysis.

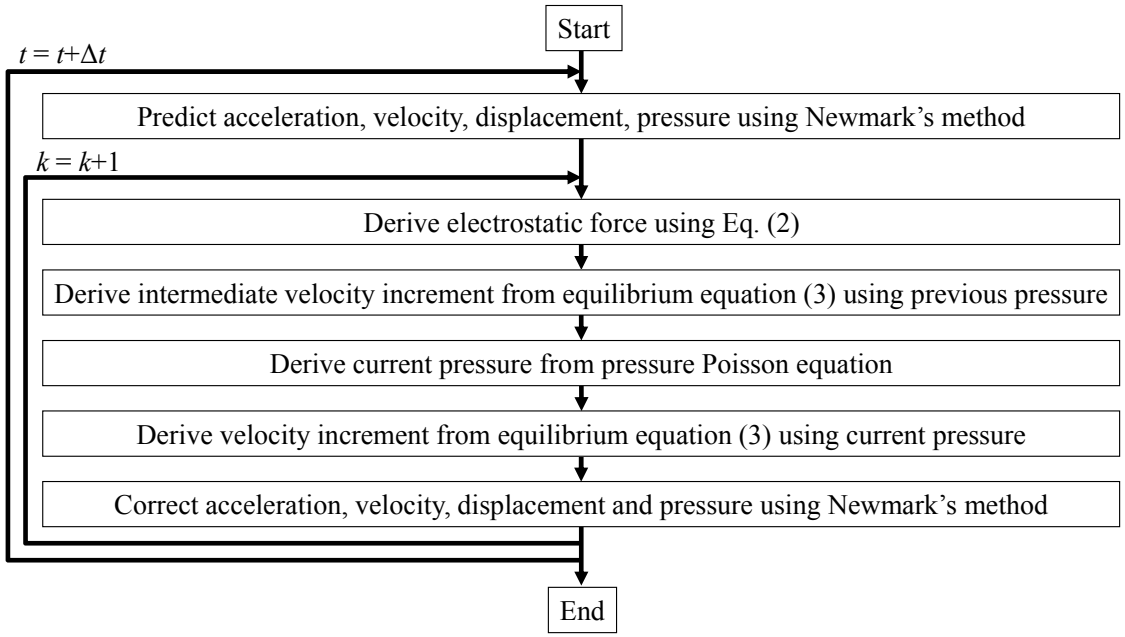


Figure 1: Analysis flow of the prpposed method.

3 NUMERICAL EXAMPLE

3.1 Problem setup

The present problem is schematically shown in Fig. 2. A micro cantilever beam is made using a chemical etching for the SOI wafer (the upper and bottom layers are made of arsenic-doped Si and conductive, while the middle layer is made of SiO_2 and isolated). The step voltage is applied between the beam (the upper layer) and the base (the bottom layer). The beam has the dimensions of the length $L = 1.00\times10^3$ μm, the width $W = 36$ μm, the thickness $H = 3.0$ μm and the gap $G = 5.0$ μm, which are identified using the SEM images.

The structural mass density ρ^s is assumed to be equal to that of the bulk material of Si 2328 kg/m^3. The Poisson's ratio ν is assumed to be 0 [3]. Since the Young's modulus E is usually different from that of the bulk material of Si due to the scale effect, it is evaluated as follows: The natural frequency of the beam is evaluated from the experimental result of its free vibration in vacuum (under 2 Pa). Then, E is evaluated using the natural frequency from the experiment and its theoretical solution under the assumption of the Euler-Bernoulli beam. E is evaluated as 184.2 GPa, which is consistent with that in the previous study [4]. The material properties of air (26 degrees C) are the mass density $\rho^s = 1.18\times10^{-3}$ g/cm^3 and the viscosity $\mu = 1.82\times10^{-4}$ g/(cm sec). The dielectric constant of air is 8.859×10^{-12} F/m.

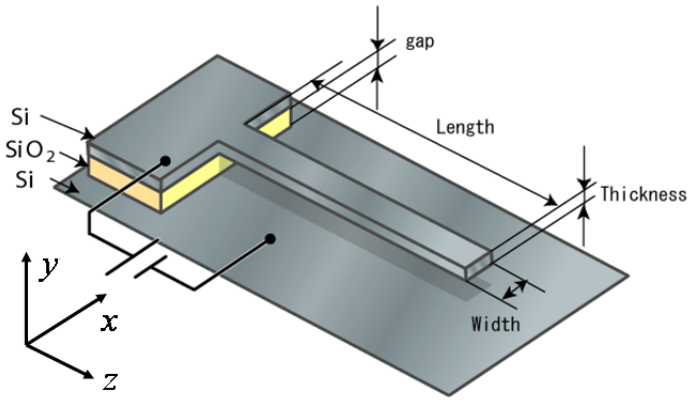


Figure 2: Schematic view of the problem setup. The micro cantilever beam is driven and vibrated by the electrostatic force generated by the step voltage between the beam and the base.

Each voltage to the chip is supplied by the high speed DC power supply (the risetime is approximately 50μsec). The velocity in the *y*-direction of the free end of the beam is measured using the laser Doppler vibrometer (Ono Sokki Co., Ltd., Japan). The data of the supplied voltage and the velocity of the beam are simultaneously collected using a data acquisition system with a sampling speed of 500,000Hz. The former is used to obtain the applied voltage between the beam and the base, and the latter is used for the validation of the present numerical results. The time histories of the maximum applied voltages are shown in Fig. 3.

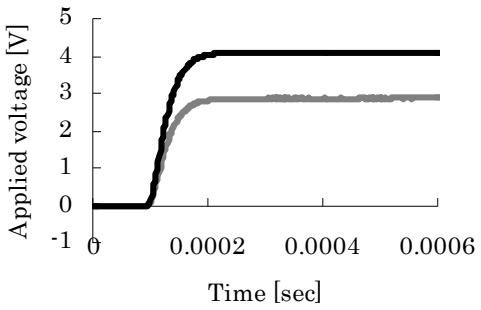


Figure 3: Time histories of the maximum applied voltages. The black line indicates the applied voltage in air, while the gray line indicates that in vacuum.

The symmetry with respect to the *y*-*z* plane is taken into account for the analysis domain. The mesh for the structural domain consists of 661 nodes and 240 elements, where the quadratic hexahedral elements (20 nodes) are used, the mesh for the fluid domain consists of 5,916 nodes and 26,698 elements, where the linear tetrahedral elements (4 nodes) are used, and the mesh for the electrostatic field domain consists of 195 nodes and 480 elements, where the linear tetrahedral elements (4 nodes) are used. The time increment used in the numerical analyses is 1μsec.

3.2 Results and discussion

Fig. 4 shows the time histories of the velocity in the y-direction of the free end of the beam for the applied voltage in Fig. 3. As shown in this figure, the air damping was significant. Therefore, the FSI must be taken into account in the present problem. Fig. 4 also shows the necessity of the full interaction analysis. As shown in this figure, the numerical result for the structure-fluid-electrostatic interaction was consistent with the experimental result, while the numerical result for the FSI was inconsistent with the other results. The vibration from the full interaction analysis was over-damping, while that from the FSI analysis was under-damping. This qualitative difference was caused by the high sensitivity of damping to the gap. At equilibrium, the displacement from the full interaction analysis was -1.36μm, while that from the FSI analysis was -0.899μm. Therefore, the damping in the full interaction analysis was far larger than that in the FSI analysis.

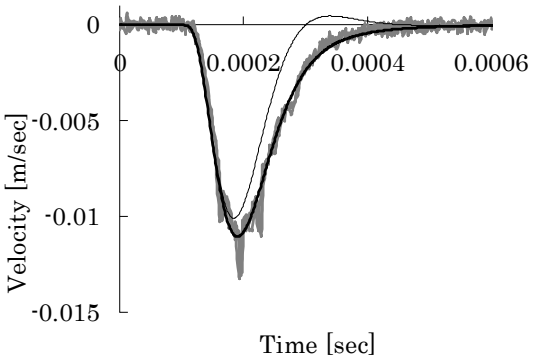


Figure 4: Time histories of the y-velocity of the free end of the micro cantilever beam for the applied voltage in Fig.3. The black line indicate the numerical result for the structure-fluid-electrostatic interaction, and the narrow black line indicates the numerical result for the FSI, while the gray line indicates the experimental result in air.

4 CONCLUSIONS

In the present study, the finite element analysis using the hierarchal decomposition was proposed in order to solve the structure-fluid-electric interaction. The proposed analysis was applied to the micro cantilever beam driven by the electrostatic force in air. It follows from the comparison among the results for the structure-fluid-electrostatic interaction, the FSI, and the experiment that the proposed analysis taking into account the full-interaction can predict the MEMS dynamic behavior accurately.

ACKNOWLEDGEMENT

This work was supported by JSPS KAKENHI 26390133.

REFERENCES

[1] De SK, Aluru NR. Coupling of hierarchical fluid models with electrostatic and mechanical models for the dynamic analysis of MEMS. *Journal of Micromechanics and Microengineering* 2006;16:1705-1719.

[2] Ghosh R, Mukherjee S. Fully Lagrangian modeling of dynamics of MEMS with thin beams-part 2: undamped vibration. *Journal of Applied Mechanics* 2009;76:051008.
[3] Namazu T, Tanaka M, Inoue S. A simple determination method of in-plane Poisson's ratio for MEMS materials by means of on-chip pure bending test. *Proceedings of 20th IEEE International Conference on Micro Electro Mechanical Systems* 2007; 235-238.
[4] Sato K, Yoshioka T, Ando T, Shikida M, Kawabata T. Tensile testing of silicon film having different crystallographic orientations carried out on a silicon chip. *Sensors and Actuators A* 1998;70:148-152.
[5] Ishihara D, Horie T. A projection method for the interaction of an incompressible fluid and a structure using new algebraic splitting. To appear in *Computer Modeling in Engineering and Sciences*.
[6] Ishihara D, Yoshimura S. A monolithic approach for interaction of incompressible viscous fluid and an elastic body based on fluid pressure Poisson equation. *International Journal for Numerical Methods in Engineering* 2005;64:167-203.
[7] Ishihara D, Kanei S, Yoshimura S, Horie T. Efficient parallel analysis of shell-fluid interaction problem by using monolithic method based on consistent pressure Poisson equation. *Journal of Computational Science and Technology* 2008;2(1):185-196.
[8] Ishihara D, Horie T, Denda M. A two dimensional computational study on fluid-structure interaction cause of wing pitch changes in dipteran flapping flight. *The Journal of Experimental Biology* 2009;212:1-10.
[9] Ishihara D, Horie T, Niho T. An experimental and three-dimensional computational study on the aerodynamic contribution to the passive pitching motion of flapping wings in hovering flies. *Bioinspiration and Biomimetics* 2014;9(4):046009(23pp).

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis (Eds)

THE NANOCOPS PROJECT ON ALGORITHMS FOR NANOELECTRONIC COUPLED PROBLEMS SOLUTIONS

RICK JANSSEN [1], JAN TER MATEN [2], CAREN TISCHENDORF [3], HANS-GEORG BRACHTENDORF [4], KAI BITTNER [4], WIM SCHOENMAKER [5], PETER BENNER [6], LIHONG FENG [6], ROLAND PULCH [7], FREDERIK DELEU [8], AARNOUT WIEERS [8]

[1] Corresponding author. NXP Semiconductors
High Tech Campus 46, 5656 AE Eindhoven, the Netherlands.
E-mail: rick.janssen@nxp.com

[2] Bergische Univ. Wuppertal, Germany; [3]Humboldt Univ. zu Berlin; [4] FH Oberösterreich, Hagenberg, Austria; [5] Magwel NV, Leuven, Belgium; [6] Max Planck Gesellschaft, Magdeburg, Germany; [7] Ernst-Moritz-Arndt Univ. Greifswald, Germany, [8] ON Semiconductor, Oudenaarde, Belgium

Key words: nanoelectronics, coupled problems, power devices, RF circuits

Abstract. The nanoCOPS project [1, 2] is a collaborative research project within the FP7-ICT research program funded by the European Union. The consortium comprises experts in mathematics, physics and electrical engineering from seven universities (BU Wuppertal, HU Berlin, Brno UT, TU Darmstadt, FH OÖ Hagenberg, U Greifswald, KU Leuven), one research institute (MPG Magdeburg), two industrial partners (NXP Semiconductors Netherland and ON Semiconductor Belgium) and two SMEs (MAGWEL and ACCO Semiconductor). We present an overview of the project subjects addressing the "bottlenecks" in the currently-available infrastructure for nanoelectronic design and simulation. In particular, we discuss the issues of an electro-thermal-stress coupled simulation for Power-MOS device design and of simulation approaches for transceiver designs at high carrier frequencies and baseband waveforms such as OFDM (Orthogonal Frequency Division Multiplex).

1 INTRODUCTION

Designs in nanoelectronics often lead to large-size simulation problems and include strong feedback couplings. Industry demands the provisions of variability to guarantee quality and yield. It also requires the incorporation of higher abstraction levels to allow for system simulation in order to shorten the design cycles, while at the same time preserving accuracy. The nanoCOPS project addresses the simulation of two technically and commercially important problem classes identified by the industrial partners:

- Power-MOS devices, with applications in energy harvesting, that involve couplings between electromagnetics (EM), heat, and stress, and
- RF-circuitry in wireless communication, which involves EM-circuit-heat coupling and

multirate behaviour, together with analogue-digital signals.

To meet market demands, the scientific challenges are to:

- create efficient and robust simulation techniques for strongly coupled systems, that exploit the different dynamics of sub-systems and that allow designers to predict reliability and ageing;
- include a variability capability such that robust design and optimization, worst case analysis, and yield estimation with tiny failures are possible (including large deviations like 6-sigma);
- reduce the complexity of the sub-systems while ensuring that the parameters can still be varied and that the reduced models offer higher abstraction models that are efficient to simulate.

Our solutions are:

- to develop advanced co-simulation/multirate/monolithic techniques, combined with envelope/wavelet approaches;
- to produce new generalized techniques from Uncertainty Quantification (UQ) for coupled problems, tuned to the statistical demands from manufacturability;
- to develop enhanced, parameterized Model Order Reduction techniques for coupled problems and for UQ.

The best (efficient, robust) algorithms produced are currently being implemented and transferred to SME partner MAGWEL. Validation is conducted on industrial designs provided by the industrial partners. A thorough comparison to measurements on real devices will be made. A public online progress report can be found in [3].

2 SIMULATION OF COUPLED PROBLEMS: CO-SIMULATION, MULTIRATE, AND MONOLITHIC

The coupling of various physical effects in nanoelectronics plays an important role in the operational reliability, at both circuits and systems level. This is the case for high-performance applications (CPUs, RF-circuits) as well as applications in hostile environments (e.g., such as high voltages and/or high currents in automotive applications, RF Power and Base Stations applications). Various types of coupled phenomena exist. For example, electro-thermal coupling is a key concern during operational cycles in industry where a substantial amount of heat is generated that (1) will affect the voltage and current distributions and (2) will indirectly impact the sources of the heat itself. The extent and impacts of electro-thermal-stress coupling is studied in the modelling of power-MOS devices in DC and in the transient regime (time domain), taking environmental aspects like metal stack and package into account. The determination of both reliability and ageing needs to be more effectively addressed by the combined simulation of these coupled effects. Another challenging coupling mechanism concerns Radio Frequency (RF) designs that have to involve with circuit-EM-heat couplings,

where parasitic long-range electromagnetic (EM) effects induce substantial distortion at the circuit level, which can lead to the sudden malfunction of the circuit. In order to address both these types of problems, companies need to have a capability for the simulation of multi-physics with dynamics involving different time scales.

Co-simulation techniques are natural approaches in efficiently solving coupled problems. Field-circuit couplings have been considered in [4, 5]. Dynamic iteration can be performed within each time window [6]. In [7], for the field-thermal coupling this is combined with a time-averaging for the heat source, thus exploiting multirate difference in the dynamics between the field and the heat quantities. When strong couplings arise one will reside to efficient monolithic algorithms to solve the coupled problem (see also Fig. 1).

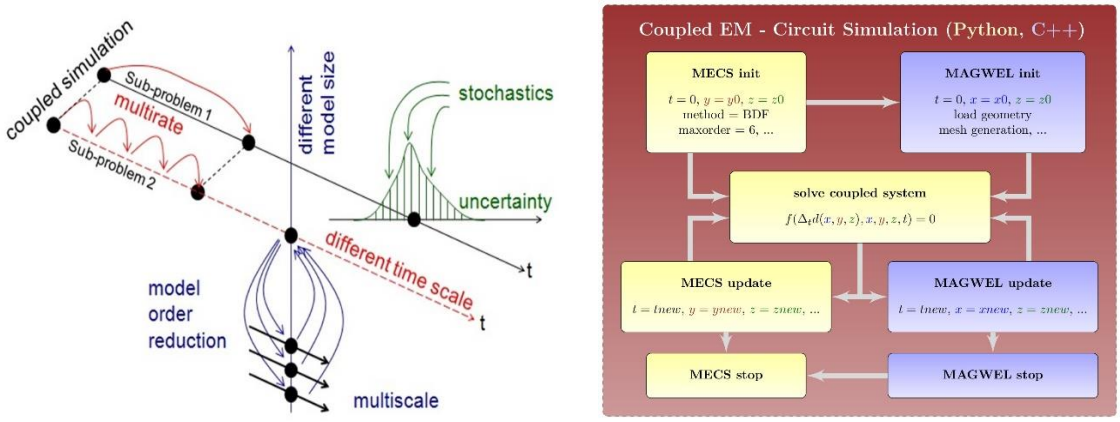


Fig. 1. **Left:** Schematic of a coupled problem (consisting of two sub-problems), including uncertainties. In nanoCOPS, those problems are efficiently solved in time domain and probability space with exploitation of their multirate (different time steps) and multiscale behaviour (different discretizations). The discretized models will lead to different reduced models by techniques from MOR. Parameterized MOR will guarantee the ability to properly deal with the uncertainties in parameters, geometries, and coupling quantities. **Right:** Overview flow for strongly coupled problems, yielding a monolithic coupled simulation.

Multirate time integration for circuit simulation has been studied for circuit decomposition as well as for signals with a broad difference in the frequency domain. When different signal shapes are present in the circuit, these may be approximated more efficiently if individual grids are used for each of the signals. The problem is cast in a hyperbolic time-domain formulation involving two time directions, one mimicking a low-frequency behavior, the other one a high-frequency one with periodic boundary conditions: an MPDAE system (Multirate Partial Differential-Algebraic Equations). This splitting can be adaptively optimized during the process. The grid points along the fast-varying direction may vary when progressing in the low-frequency time direction (see Fig. 2 (left)). Spline/wavelet methods are exploited for reasons of compactness of the support of the basis functions.

As an example we consider a chain of 5 frequency dividers (as part of a PLL). In each step the frequency is reduced by a factor 2 as one can see in Fig. 2 (right). From the solution in (τ_1, τ_2)-time-domain space, a 1-dimensional solution depending on $(t, \phi(t))$ (for a suitable phase-function ϕ) can be constructed, which provides an envelope solution. Recently, the

method has been extended to deal with circuit partitions as well [8, 9]. Currently, one considers coupling with heat as well.

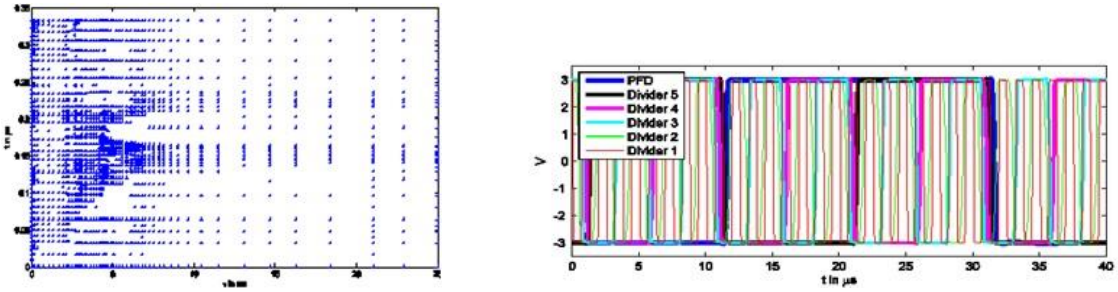


Fig. 2. **Left:** Adaptive grid for a Colpitss oscillator. **Right:** Several signals in a frequency divider chain as part of a PLL.

A highlighting monolithic simulation for a coupled electromagnetic-heat problem is shown in Fig. 3. It couples large-scale (millimeter) structures to small-scale (sub-micron) finger details of a power MOS.

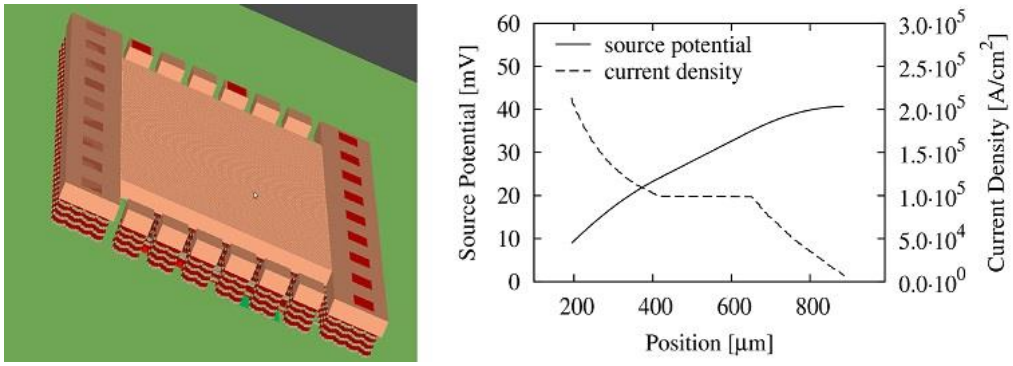


Fig. 3: Analysis of a power MOS (left) resulting into an asymmetric current density due to thermally induced conductance variations in the metallic interconnect (right).

3 MODEL ORDER REDUCTION, UNCERTAINTY QUANTIFICATION

The involvement of varying parameters, needed in the design process, affects the performance of time-domain simulation techniques: the dynamics can change (resulting in different time steps in a monolithic simulation, or in more iterations for Dynamic Iteration in co-simulation, or in different multirate behavior of MPDAEs. Also projection matrices in Model Order Reduction techniques depend on chosen parameters. In [10] a robust algorithm for parametrized Model Order Reduction (pMOR), based on implicit moment matching involving Krylov-space techniques, has been derived, for linear systems based on state-space formulations, which directly applies to circuit equations. In [11] the method has been extended to second order systems coming from electromagnetic field discretizations. In a monolithic coupling the electromagnetic vector potential generates a heat coupling to heat by a power term

giving a quadratic term. For such couplings special MOR methods have been developed recently [12]: bilinear and quadratic-linear models were considered that allow treatments in the frequency domain (involving multiple frequencies). Additionally in [11] an a posteriori output error bound for reduced order models of micro- and nano-electrical(-mechanical) systems is derived. The error bound is independent of the discretization method (finite difference, finite element, finite volume) applied to the original PDEs. Secondly, the error bound can be directly used in the discretized vector space, without going back to the PDEs, and especially to the bilinear form (weak formulation) associated with the finite element discretization, which must be known a priori for deriving/using the error bound for the reduced basis method. This can be combined with adaptive selection of expansion points. These techniques enable automatic generation of the reduced models computed by parametric model reduction methods based on approximation (interpolation) of the transfer function, e.g., Krylov subspace based methods.

Fig. 4: Relative error of the expectation of the electrical field in a coplanar waveguide after using a pMOR Krylov-space technique within Uncertainty Quantification based on generalized Polynomial Chaos expansions.

In [13, 14] methods for Uncertainty Quantification (UQ) via generalized Polynomial Chaos (gPC) expansions have been proposed. These methods can greatly benefit when being combined with methods for pMOR [15] (see also Fig. 4). Assuming that the discretization of the underlying structure of the electromagnetic problem is fixed, in [16] UQ-results are obtained involving parameterized MOR. In [17] the sensitivity of the variance with respect to parameters is considered. This gives an indication of dominant parameters, see also [15]. With Stroud-quadrature [18] one can deal with a number of parameters that are of interest for industrial purposes.

In [19] stochastically varying domains are considered, leading to topology optimization for a permanent magnet (PM) synchronous machine with material uncertainties. These techniques are now applied to problems in nanoscale. In [20] for variation of the solution due to varying the thickness of a layer in a power MOS transistor model, or due to varying the conductivity in the layer, gPC was used to efficiently estimate mean and standard variation.

In [21] the effect of the number of parameter variations on the impact of noise from digital parts on the isolation sensitive RF domains was investigated, i.e., the number of downbonds, the number of ground pins, the domain spacing and shape, the application of deep-Nwell and

exposed diepad, and the number of exposed diepad vias.

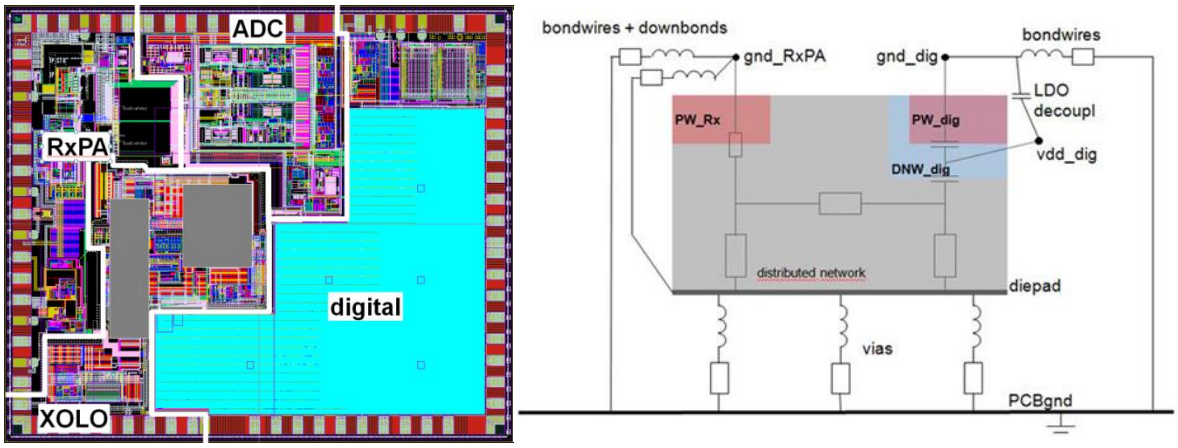


Fig. 5: **Left:** Integrated RF-CMOS automotive transceiver design. **Right:** Floorplan model for isolation and grounding strategies.

In order to minimize interference issues and coupling effects in RF products, it is essential to apply proper floorplanning and grounding strategies. The interaction of the IC with its physical environment needs to be accounted for, so as to certify that the final packaged and mounted product meets the specifications.
The first focus is on the key requirements to address physical design issues in the early design phases of complex RF designs. Typical physical design issues encountered, such as on-chip coupling effects, chip-package interaction, substrate coupling and co-habitation, have been investigated.
The main challenges are the first order prediction of cross-domain coupling. Therefore we apply a floorplan methodology to quantify the impact of floorplanning choices and isolation grounding strategies. This methodology is based on a very high level floorplan EM/circuit simulation model, including the most important interference contributors and including on-chip, package and PCB elements, to be applied in the very early design phases (initial floorplanning).
The overall model of a complete RF product contains the following parts (see Fig. 5):

- On-chip: domain-regions, padring, sealring, splittercells, substrate effects.
- Package: ground and power pins, bondwires/downbonds, exposed diepad.
- PCB: ground plane and exposed diepad connections.

ACKNOWLEDGEMENTS

We acknowledge the Support from the Project nanoCOPS, nanoelectronic COpled Problems Solutions (FP7-ICT-2013-11/619166), http://www.fp7-nanoCOPS.eu/.

REFERENCES

[1] http://www.fp7-nanocops.eu/.
[2] Janssen, H.H.J.M., Benner, P., Bittner, K., Brachtendorf, H.-G., Feng, L., ter Maten, E.J.W.,

Pulch, R., Schoenmaker, W., Schöps, S., Tischendorf, C.: The European Project nanoCOPS for Nanoelectronic Coupled Problems Solutions, Paper accepted for proceedings ECMI-2014, 18th European Conference on Mathematics for Industry, Taormina, Sicily, June 9-13, 2014. An extended version will be submitted to a special issue of the Journal of Mathematics in Industry.

[3] ter Maten, E.J.W., Günther, M., Putek, P., Benner, P., Feng, L., Schneider, J., Brachtendorf, H.-G., Bittner, K., Deleu, F., Wieers, A., Janssen, R., Kratochvil, T., Gotthans, T., Pulch, R., Liu, Q., Reynier, P., Schoenmaker, W., Meuris, P., Schöps, S., De Gersem, H., Tischendorf, C., Strohm, C., nanoCOPS: Nanoelectronic COupled Problem Solutions. ECMI Newsletter 56, pp. 62-67, 2014.
Online: http://www.mafy.lut.fi/EcmiNL/issues.php?action=viewart&ID=351

[4] Bartel, A., Brunk, M., Günther, M., Schöps, S.: Dynamic iteration for coupled problems of electric circuits and distributed devices. SIAM J. Sci. Comput., 35(2) pp. B315–B335, 2013.

[5] Tischendorf, C., Schoenmaker, W., De Smedt, B., Meuris, P., Baumanns, S., Matthes, M., Jansen, L., Strohm, C.: Dynamic Coupled Electromagnetic Field Circuit Simulation. Presented in Minisymposium on "Simulation Issues for Nanoelectronic Coupled Problems" at ECMI-2014, 18th European Conference on Mathematics for Industry, Taormina, Sicily, June 11, 2014.

[6] Schöps, S.: Iterative Schemes for Coupled Multiphysical Problems in Electrical Engineering. Invited talk SCEE-2014 (Scientific Computing in Electrical Engineering), Wuppertal, Germany, 2014. IMACM Report 2014-28, pp. 11–12, Bergische Universität Wuppertal, 2014,
http://www.imacm.uni-wuppertal.de/fileadmin/imacm/preprints/2014/imacm_14_28.pdf.

[7] Kaufmann, C., Günther, M., Klagges, D., Knorrenschild, M., Richwin, M., Schöps, S., ter Maten, E.J.W.: Efficient frequency-transient co-simulation of coupled heat-electromagnetic problems. Journal of Mathematics in Industry, 4:1, 2014.
http://www.mathematicsinindustry.com/content/4/1/1.

[8] Bittner, K., Brachtendorf, H.-G.: Adaptive multi-rate wavelet method for circuit simulation. Radioengineering, 23-1, pp. 300–307, 2014,
http://www.radioeng.cz/fulltexts/2014/14_01_0300_0307.pdf.

[9] Bittner, K., Brachtendorf, H.-G.: Fast algorithms for grid adaptation using non-uniform biorthogonal spline wavelets. SIAM J. Scient. Computing, submitted, 2014.

[10] Benner, P., Feng, L.: A robust algorithm for parametric model order reduction based on implicit moment matching. In: A. Quarteroni, G. Rozza (Eds): Reduced Order Methods for modeling and computational reduction, MS &A series, 9, Springer, pp. 159–186, 2014.

[11] Feng, L., Benner, P., Antoulas, A.C.: An a posteriori error bound for reduced order modeling of micro- and nano-electrical(-mechanical) systems. Presented at SCEE-2014 (Scientific Computing in Electrical Engineering), Wuppertal, Germany, 2014. IMACM Report 2014-28, pp. 99–100, Bergische Universität Wuppertal, 2014,
http://www.imacm.uni-wuppertal.de/fileadmin/imacm/preprints/2014/imacm_14_28.pdf.

[12] Benner P., Breiten T., Krylov-subspace based model reduction of nonlinear circuit models using bilinear and quadratic-linear approximations, in: M. Günther, A. Bartel, M. Brunk, S. Schöps, M. Striebel (Eds.), Progress in Industrial Mathematics at ECMI 2010, Series Mathematics in Industry Vol. 17, Springer, pp. 153–160, 2012.

[13] Le Maître, O.P., Knio, O.M.: Spectral methods for uncertainty quantification, with

applications to computational fluid dynamics. Springer, Science+Business Media B.V., Dordrecht, 2010.
[14] Xiu, D.: Numerical methods for stochastic computations - A spectral method approach. Princeton Univ. Press, Princeton, NJ, USA, 2010.
[15] ter Maten, E.J.W., Pulch, R., Schilders, W.H.A., Janssen, H.H.J.M.: Efficient calculation of Uncertainty Quantification. In: M. Fontes, M. Günther, N. Marheineke (Eds): Progress in Industrial Mathematics at ECMI 2012, Series Mathematics in Industry Vol. 19, Springer, pp. 361–370, 2014.
[16] Benner, P., Schneider, J.: Uncertainty quantification for Maxwell's equations using stochastic collocation and model order reduction. Report MPIMD/13-19, Max Planck Institute Magdeburg, 2013. http://www.mpi-magdeburg.mpg.de/preprints.
[17] Pulch, R., ter Maten, E.J.W., Augustin, F.: Sensitivity analysis and model order reduction for random linear dynamical systems. Accepted for Mathematics and Computers in Simulation, 2015. See also: CASA-Report 2013-15, TU Eindhoven, 2013, http://www.win.tue.nl/analysis/reports/rana13-15.pdf.
[18] Stroud, A.H.: Remarks on the disposition of points in numerical integration formulas. Mathematical Tables and Other Aids to Computation, 11(60), pp. 257–261, 1957.
[19] Putek, P., Gausling, K., Bartel, A., Gawrylczyk, K.M., ter Maten, J., Pulch, R., Günther, M.: Robust topology optimization of a Permanent Magnet synchronous machine using level set and stochastic collocation methods. Submitted to proceedings SCEE-2014 (Scientific Computing in Electrical Engineering), Wuppertal, Germany, 2014. See also: IMACM Report 2014-28, pp. 83–84, Bergische Universität Wuppertal, 2014, http://www.imacm.uni-wuppertal.de/fileadmin/imacm/preprints/2014/imacm_14_28.pdf.
[20] Putek, P., Meuris, P., Günther, M., ter Maten, J., Pulch, R., Wieers, A., Schoenmaker, W., Uncertainty quantification in electro-thermal coupled problems based on a power transistor device. Accepted for presentation at MATHMOD 2015, Vienna, Feb. 18-20, 2015.
[21] Di Buccianico, A., ter Maten, J., Pulch, R., Janssen, R., Niehof, J., Hanssen, M., Kapora, S.: Robust and efficient uncertainty quantification and validation of RFIC isolation. Radioengineering, Volume 23, Issue 1, pp. 308-318, 2014.
Online: http://www.radioeng.cz/fulltexts/2014/14_01_0308_0318.pdf

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis (Eds)

ACOMPARISION OF CONFORMING AND NON-CONFORMING MESH METHODS FOR FLOW AROUND A CIRCULAR CYLINDER IN NON-INERTIAL FRAME OF REFERENCES

S.H. MADANI, J. WISSINK, H. BAHAI

College of Engineering, Design and physical study
Brunel University London, UK
Hossein.Madani@Brunel.ac.uk

Key words: FSI, Vortex shedding, Immersed boundary, interpolation/reconstruction method, cylindrical coordinates, circular cylinder

Abstract. For applications with large physical domains and large body displacements it is of vital importance to use an accurate and computationally affordable numerical method. The objective of the present study is to compare the accuracy and computational expenses of an immersed boundary method (using IB interpolation) with those of a boundary-conforming numerical method. For the latter, the Navier-Stokes equations were solved using cylindrical coordinates [1]. The same boundary conditions for inlet and outlet were applied in both simulations. In both cases a non-inertial frame of reference was applied to be able to model moving boundaries [2]. The vortical structures that appear behind the cylinder, as well as the drag and lift coefficients and the Strouhal number for forced and vortex-Induced-Vibrations (VIV) are compared under various conditions. Although in cylindrical coordinates the definition of the boundary condition at the cylindrical wall is more accurate, the definition of the outflow condition was found to be problematic due to the usage of the moving frame of reference. The simulation results show that both approaches produce acceptable results. When using a similar number of mesh points, the simulation using cylindrical coordinates is less expensive, though the need for a larger computational domain when using cylindrical coordinates (to overcome difficulties in the definition of the outflow boundary conditions) again increases the computational costs of this method.

INTRODUCTION

The flow over moving bluff bodies received significant attention in the recent decade and obtaining accurate solutions for Fluid-Structure Interaction (FSI) problems is of interest in many engineering and scientific applications. A broad classification of FSI methods can be based on the type of mesh employed in the discretization, where we can differentiate between boundary-conforming and non-boundary-conforming mesh methods [3]. A well-known conforming mesh method is the curvilinear coordinate method which is combined with the Arbitrarily Lagrangian-Eulerian method (ALE) to model the FSI problem. For non-conforming mesh methods usually an immersed boundary method is employed and the most recent developments in FSI methods are based on this approach. On the other hand, in some

applications, when there is only one moving structure, it is possible to adjust coordinates on the moving body and solve the flow governing equation in the relative frame of reference. In this case, there is no need to deform the mesh (curvilinear coordinate approach) or update the interpolation formulas (immersed boundary) while simulating flow structure interaction. In this study the objective is to compare the results using an immersed boundary (non-conforming approach) with those using a conforming approach (curvilinear coordinates). For the latter, to minimize the discretization overhead cylindrical coordinates are used. To be able to compare our results with other experimental and numerical results, a low-Reynolds number flow around a cylinder is used as the test case.

The immersed-boundary (IB) method is a technique for solving flow problems in regions with irregular boundaries using a simple structured grid solver. The term "immersed boundary method" was initially introduced by Peskin [4] while presenting simulation of blood flow in a cardiovascular system. It was specifically designed to handle deforming (elastic) boundaries interacting with low Reynolds number flow. The simulation was carried out on a Cartesian grid and at those locations where the boundary did not align with a mesh line the solution algorithm was locally modified to enforce the desired boundary conditions on the flow. More recently, numerous modifications and refinements have been proposed to enhance the accuracy, stability, and application range of the IB method [5].

Depending on the way that the boundary conditions are imposed on the immersed boundary, the IB methods can be generally categorized into continuous and discrete forcing approaches. In the continuous forcing method [3], a forcing function is applied to the Navier-Stokes equation in order to maintain the boundary condition on the structure (e.g. enforcing a no-slip boundary condition on a stationary body). The most important issue in this method is the definition of the continuous forcing function needed to enforce the correct boundary conditions [6].

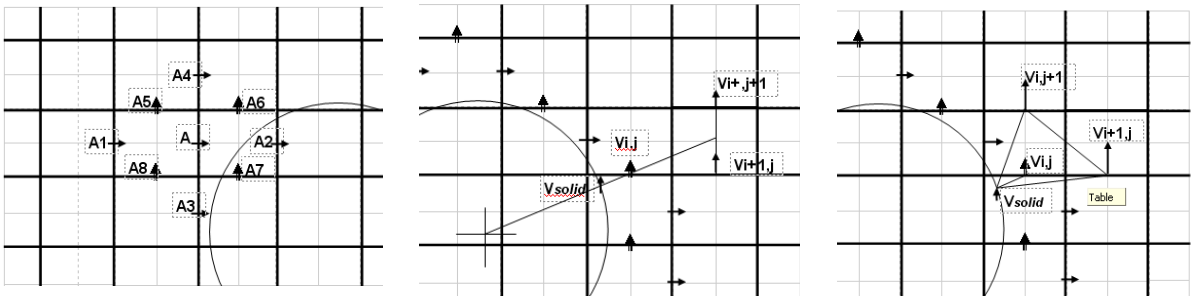


Figure 1: A 2D Cartesian mesh with a solid boundary (circle) left). Interface point A requires interpolation, points A1 to A8 are all neighbouring points of A of which A2 and A7 are inside the solid domain. Right) Standard Reconstruction Method (SRM)[7]. Middle) Bilinear interpolation perpendicular to the solid surface [3].

In this paper, the indirect discrete forcing approach is used for the IB method. In this case, the flow velocity is interpolated at the interface cells and the forcing term is imposed indirectly to the discrete equations. The interface points are defined as the points in the fluid domain near the solid boundary for which one of the neighbouring points in the discretized equations is inside the solid domain [3]. Therefore, the parameters related to these points cannot be updated through solving the governing equation. Any cells that contain one or more interface points are called interface cells (Figure 1). In the IB interpolations approach, each

time step the flow parameters in the interface cells are updated by direct interpolation and used as boundary condition for the flow solver.

In the next section, the governing equations of the fluid flow and the structural problem, which has been used in the simulation, are introduced.

FORMULATION AND NUMERICAL METHODS

Flow governing equations

The Navier–Stokes equation for an unsteady, incompressible fluid flow in vector form and relative frame of references is given by equations (1) and (2).

$$\rho\left(\frac{\partial \mathbf{V}}{\partial t} + \mathbf{V}.\nabla\mathbf{V}\right) = -\nabla p + \mu\nabla^2\mathbf{V} + \rho G(v,t) \quad \textbf{(1)}$$

$$\nabla.\mathbf{V} = \mathbf{0} \quad \textbf{(2)}$$

Where G(v,t) in equation (1) is the additional acceleration term due to deriving the Navier-Stokes equation in a moving frame of reference which is defined as equation (3). For more information see [2].

$$G(v,t) = 2\dot{\theta}I_0 V + \left(\dot{\theta}\right)^2 X + \theta I_0 X - A^T\ddot{d} \quad \textbf{(3)}$$

In equation (3), the terms including $\dot{\theta}$ and $\ddot{\theta}$ are omitted so that the rotation matrix A^T becomes a unit vector, hence only the transverse degree of freedom of the cylinder is considered here. As a result, equation (3) reduces to describing the acceleration of a moving solid body in the transverse direction,- a_{ysolid}.

In addition, the following terms are used to non-dimensionalize the flow governing equations.

$$t^* = \frac{t\,V}{D} \quad x^* = \frac{x}{D} \quad y^* = \frac{y}{D} \quad p^* = \frac{P}{\rho V^2} \quad u^* = \frac{u}{V} \quad v^* = \frac{v}{V} \quad Re = \frac{\rho V D}{\mu} \quad \textbf{(4)}$$

After applying the non-dimensionalisation and simplifications given above, the incompressible 2D Navier-stokes equations in a Cartesian domain and moving frame of reference become:

$$\frac{\partial u}{\partial t} + \frac{\partial u^2}{\partial x} + \frac{\partial uv}{\partial y} = -\frac{\partial p}{\partial x} + \frac{1}{Re}\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}\right) \quad \textbf{(5)}$$

$$\frac{\partial v}{\partial t} + \frac{\partial vu}{\partial x} + \frac{\partial v^2}{\partial y} = -\frac{\partial p}{\partial y} + \frac{1}{Re}\left(\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2}\right) - a_{ysolid} \quad \textbf{(6)}$$

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} = 0 \quad \textbf{(7)}$$

For simplification "*" is dropped for the non-dimensionalized terms.

The non-dimensional, incompressible Navier-stokes equations in 2D cylindrical coordinates and moving frame of reference are given in eqns. (8) –(10).

$$\frac{\partial u_r}{\partial t} + u_r\frac{\partial u_r}{\partial r} + \frac{u_\varphi}{r}\frac{\partial u_r}{\partial \varphi} - \frac{{u_\varphi}^2}{r} = -\frac{\partial p}{\partial r} + \frac{1}{Re}\left[\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial u_r}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2 u_r}{\partial \varphi^2} - \frac{u_r}{r^2} - \frac{2}{r^2}\frac{\partial u_\varphi}{\partial \varphi}\right] - a_{ysolid}\sin(\varphi) \quad \textbf{(8)}$$

$$\frac{\partial u_\varphi}{\partial t} + u_r \frac{\partial u_\varphi}{\partial r} + \frac{u_\varphi}{r}\frac{\partial u_\varphi}{\partial \varphi} - \frac{u_r u_\varphi}{r} = -\frac{1}{r}\frac{\partial p}{\partial \varphi} + \frac{1}{Re}\left[\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial u_\varphi}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2 u_\varphi}{\partial \varphi^2} - \frac{u_\varphi}{r^2} + \frac{2}{r^2}\frac{\partial u_r}{\partial \varphi}\right] - a_{ysolid}\sin(\varphi) \quad \textbf{(9)}$$

$$\frac{\partial}{\partial r}(r u_r) + \frac{\partial u_\varphi}{\partial \varphi} = 0, \quad \textbf{(10)}$$

where 'p' is the generalised pressure which is defined by the static pressure divided by the density. Hence, to obtain the correct the static pressure we need to multiply 'p' by the density.

A staggered variable arrangement, as introduced by Harlow and Welsh [8], is used to discretize the governing equations on a Cartesian mesh and in the similar way the equations in cylindrical coordinates are discretised. The continuity equation is enforced by taking the divergence of the momentum equation and using the continuity equation to obtain a Poisson equation for the pressure field. This equation is solved by using the strongly implicit procedure (SIP), Stone's method, at every time step [9]. To maintain a consistent implementation, the pressure equation is discretized in a similar way as the momentum equations.

The extent of the computational domain was selected to be relatively large to ensure that the location of the boundaries does not affect the simulation. The size of domain in the y direction is important to be able to minimize the blocking effect and the size of the domain in the x direction should be large enough to be able to minimize the effect of outlet boundary on the vortex shedding behind the cylinder. The size of mesh near the solid boundary was selected in a way to maintain the second order accuracy of the methods while minimising the computational cost which was checked in a mesh refinement study.

A polar coordinate system was used around a circular cylinder in the conforming approach in which the grid is uniform in the circumferential direction and stretched in radial direction.

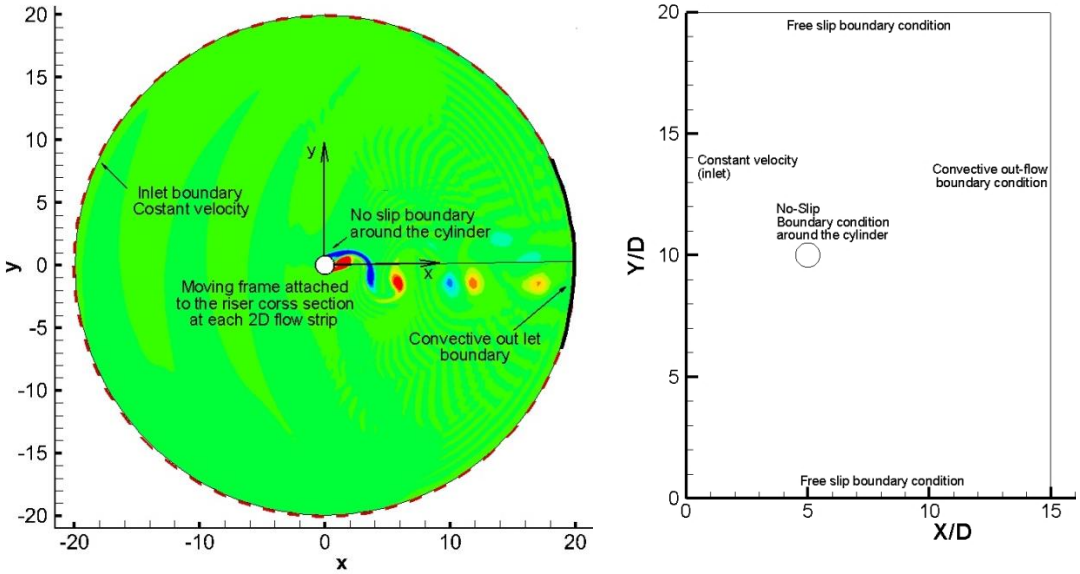


Figure 2: fluid domain size and boundary conditions, left) Cylindrical coordinates Right) Cartesian Coordinates.

Structural equations

The governing equation for a system with one degree of freedom (Figure 3) in the transverse direction is given by equation (11) which directly follows from Newton's second law applied to a moving body. In this equation 'm' is the mass of the cylinder per unit length, 'C' is the damping coefficient and 'k' is the stiffness of the structure. F_L represents the hydrodynamic force in transverse direction from the fluid flow (lift force).

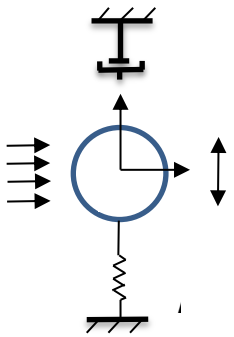


Figure 3: Schematic of the vibrating system

$$m\frac{d^2y}{dt^2} + C\frac{dy}{dt} + ky = F_L(t) \quad \textbf{(11)}$$

Non-dimensionalized scales similar to the Navier-Stokes equations are applied to the structural equation, as shown in equations (12) to (15), where V_r is the reduced velocity and f_N is the natural frequency of the structure in vacuum.

$$V_r = \frac{U_\infty}{f_N D} \quad and \quad V_r = \left(\frac{1}{2\pi}\right)\sqrt{\frac{K}{m}} \quad (12)$$

$$C_L = \frac{F_L}{\frac{1}{2}\rho U_\infty^2 D L_c} \quad (13)$$

$$m^* = \frac{m}{m_f} = \frac{m}{\rho_f \pi\left(\frac{D^2}{4}\right)L_c} \quad (14)$$

$$\xi = \frac{c}{c_c} \quad and \quad c_c = 2\sqrt{Km} \quad (15)$$

The non-dimensional form of the structural system becomes as equation (16).

$$\frac{d^2y^*}{dt^{*2}} + 2 \times \xi \times \left(\frac{2\pi}{V_r}\right)\frac{dy^*}{dt^*} + \left(\frac{2\pi}{V_r}\right)^2 y^* = \frac{2 \times C_L(t)}{\pi m^*} \quad (16)$$

In the next section a comparison is presented of simulation results obtained by using the two approaches described above.

SIMULATION RESULTS

The results are presented in three main categories. Flow over stationary cylinder, cylinder moving with a prescribed motion and Vortex-Induced –Vibration of the cylinder. In all of the cases the Reynolds number is below 200 so that the flow remains two dimensional. For the flow around the stationary cylinder the results of a mesh refinement study are presented. In the case of forced vibration, the results of a parametric study that shows the effect of the size

of the convective outlet boundary for the cylindrical coordinates are presented. In addition the results of the cylindrical coordinates are compared with the IB boundary approach as well. In the final parts the results of vortex-induced vibration for low mass ratio cases are compared with the literature.

Stationary cylinder results

Before comparing the two methods a series of parametric studies are performed to show that the results for each case are not dependent on the sizes of grid and domain. The size of the mesh near the solid boundary plays a significant role both in the accuracy of the results and in the computational expenses. In non-conforming boundary approaches, when the grid gets finer near the immersed boundary, the shape of the cylinder (IB) is approximated more accurately. To find the proper mesh size for the numerical simulation while maintaining the second-order accuracy of the model 6 grid sizes ranging from 0.2D to 0.00625D are used. The simulation results for the flow over a stationary cylinder show that the Strouhal number, the drag and the lift coefficients for sufficiently fine grids (approximately finer than 0.025) are less dependent on the grid size. Also, using the conforming boundary method, the results show that the sizes of 0.005D in radial direction and 0.015D in the circumferential direction are optimum values.

In addition the size of domain plays a significant role on the lift, drag and Strouhal number while flow is passing a circular cylinder. Here, a series of parametric studies is conducted for both the IB method and the cylindrical coordinates method.

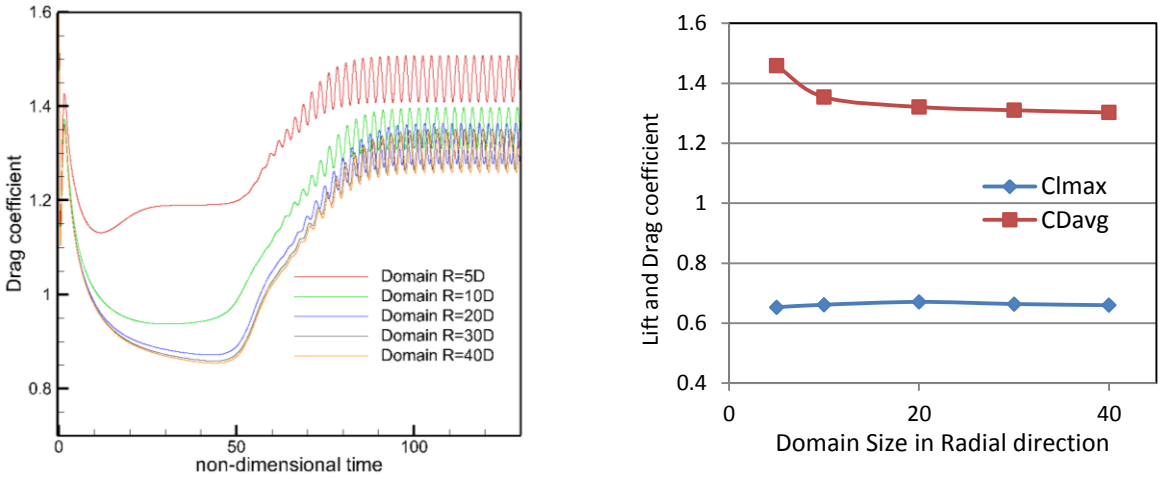


Figure 4: cylindrical coordinate (conforming boundary method), left) drag coefficient verse time of iteration, Right) the lift and drag coefficient verse size of domain in radial direction

The simulation results show that using cylinder coordinates the optimum domain size in radial direction is about 20D (figure 4). In the case of the IB method, the blockage effect becomes negligible when using a domain size of 40D in the cross flow direction (figure 5). A comparison between figures 4 and 5 indicates that the lift and drag has a similar trend in both approaches (conforming and non-conforming boundary) when using a larger domain.

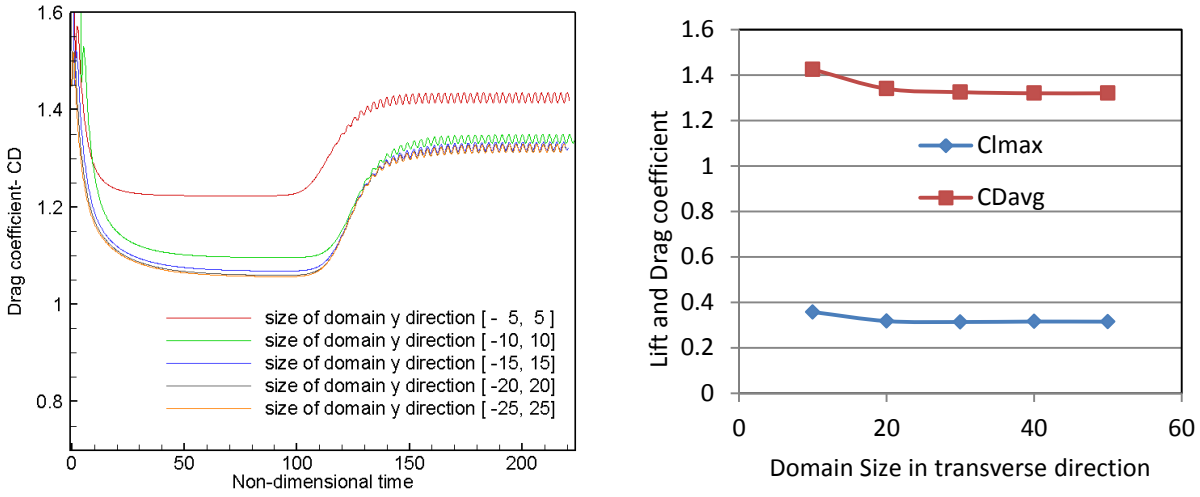


Figure 5: IB (non-conforming boundary) method, left) drag coefficient verse time of iteration, Right) the lift and drag coefficient verse size of domain in radial direction

Table 1 shows the drag and lift coefficients for the two methods at various Reynolds number using the optimum grid and domain sizes for both cases. The results show that for Reynolds numbers exceeding 150 the IB boundary approach predict higher lift and drag coefficients. However for lower Reynolds numbers, using cylindrical coordinates results in larger lift and drag coefficients. In addition, both approaches show acceptable results when compared to literature. Some of the results presented in the literature are provided for comparison in the table 1.

Table 1: Lift and Drag coefficient for the flow over a stationary cylinder

Methods	**Coefficient**	**Re=80**	**Re=100**	**Re=120**	**Re=150**	**Re=180**	**Re=200**
Cylindrical coordinate	**Lift**	0.25	0.337	0.41	0.52	0.62	0.668
	Drag	1.38	1.36	1.345	1.34	1.35	1.335
IB, Cartesian coordinate	**Lift**	0.21	0.315	0.42	0.535	0.642	0.701
	Drag	1.325	1.32	1.325	1.34	1.365	1.383
Literature	**Lift**	0.245[12]	0.329[13]	0.41[12]	----------	----------	0.8[10]
	Drag	1.35[12]	1.33[14]	1.32[12]	----------	----------	1.23[11]

Forced vibration results

One of the shortcomings of using cylindrical meshes in combination with the moving frame of reference for studying fluid structure interaction is the definition of the convective

outlet boundary conditions. While for the IB method the velocity of the structure has to be given at the top, bottom and inlet boundaries, for the boundary conforming method using cylindrical coordinates, it is not straight forward to define the optimum location of the boundary and outlet conditions. To show the effect of the location of the convective outflow boundary for the conforming method, 6 simulations using various ranges were performed and lift and drag coefficients were calculated. The size of the convective boundary is given in terms of the angle that the convective boundary arc subtends at the centre of the cylinder. The cylinder was moving with the prescribed motion $y = 0.6\sin(2\pi Ft)$. In which F= 0.75, F=f/fs, fs =0.198 is the vortex Shedding frequency at Re=200.

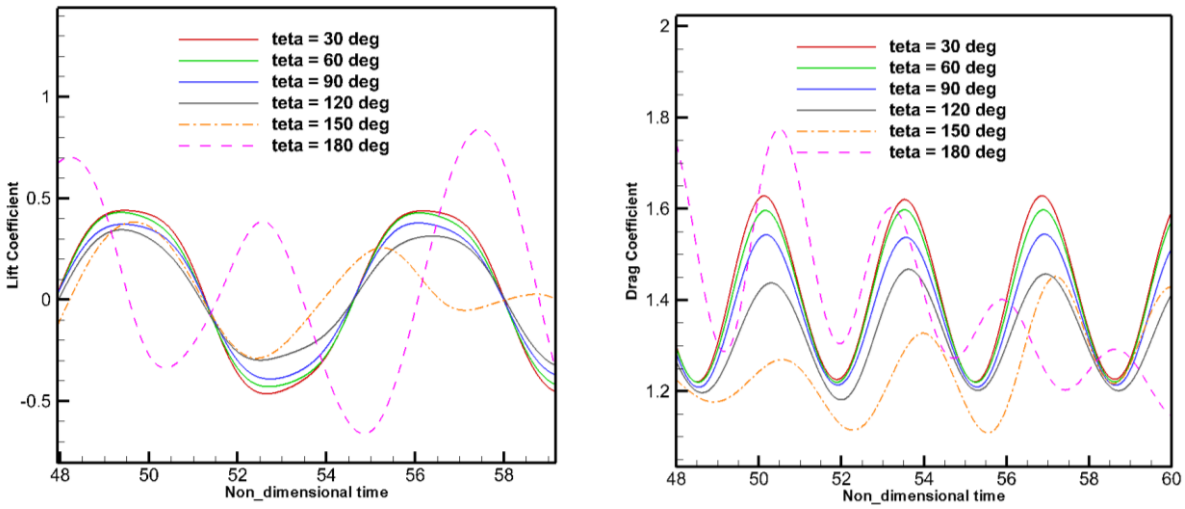


Figure 6: lift and drag coefficient. The effect of size of the convective boundary cylindrical coordinates.

Figure 6 shows that the lift and drag coefficient is converged for the cases that the convective boundary angle is less than 120 degrees, and the results are matching well with literature [11] and with the result based on immersed boundary (Figure 7).

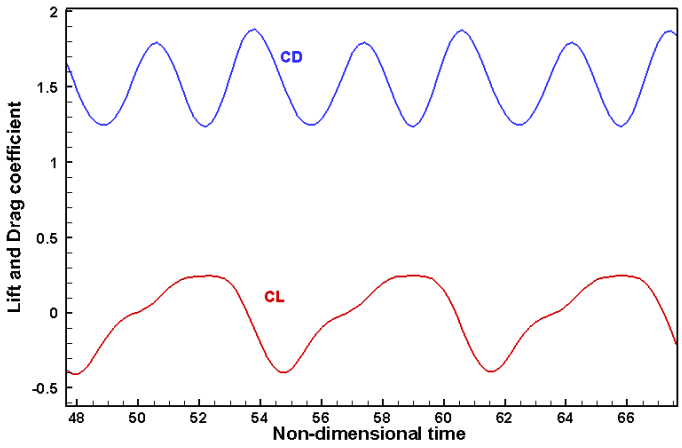


Figure 7: lift and drag coefficient, prescribed motion, Immersed boundary and moving frame of reference

Free vibration

In this stage, the simulation results for the low mass ratio are compared to the results provided by Borazjani [15] and Ahn and kallinderis [16] who employed the IB method and the ALE approach, respectively. In this case the Reynolds number, mass ratio and damping ratio are fixed at 150, 2, 0, respectively and the stiffness of the structural system is changed by varying the reduced velocity from 3 to 8.

Simulation of the flow around a stationary cylinder shows that the non-dimensional vortex shedding frequency or Strouhal number is St=0.2 at Re=150, therefore the lock-in phenomenon should occur around this frequency and hence at a reduced frequency of Vr=5.

The simulation results show that the applied IB reconstruction method accurately predicted the lock-in range, however, the maximum amplitude is predicted to be lower than the one obtained by Borazjani and Ahn & Kallinderis. This might be due to the definition of the pressure boundary for the pressure Poisson equation.

Table 2: Amplitude of oscillation (ymax/D) at various reduced velocity at constant Reynolds number, Re=150, and low mass ratio, m^*=2.

Reduced Velocity	Vr=3	Vr=4	Vr=5	Vr=6	Vr=7	Vr=8	Vr=25
Recent Research	0.04	0.42	0.38	0.30	0.20	0.06	0.03
Borazjani et al. 2008	0.06	0.52	0.48	0.43	0.38	0.08	-----
Ahn & Kallinderis 2006	0.06	0.56	0.52	0.42	0.37	0.08	-----

CONCLUSION

Two methods, one using conforming (cylindrical coordinates) and the other nonconforming (Immersed boundary) coordinate systems were employed to perform three test cases 1) flow around a stationary cylinder and 2) flow around a cylinder undergoing a forced vibration and 3) flow around a cylinder undergoing vortex induced vibrations.

For the stationary cylinder, a mesh refinement study for the IB method shows that the lift and drag coefficients become virtually grid independent as soon as the mesh size is less than 0.025D. In addition, the lift and drag coefficients obtained with the two methods at various Reynolds number ranging from 80 to 200 are found to be in good agreement with one another as well as with results found in the literature. For the forced vibration of a circular cylinder at Re=200, both approaches were found to produce acceptable results. However, when using a similar number of mesh points, the simulation the boundary conforming method turns out to be less expensive. Finally, the simulation results for the VIV case suggest that the IB method (interpolation method) properly predicted the range of lock-in. However, this method also predicted a reduced maximum vibration amplitude compared to the experimental results.

REFERENCES

[1] Verzicco, R., & Orlandi, P. A finite-difference scheme for three-dimensional incompressible flows in cylindrical coordinates. *Journal of Computational Physics*, *123*(2) (1996), 402-414.

[2] Li, L., Sherwin, S. J., & Bearman, P. W. (2002). A moving frame of reference algorithm for fluid/structure interaction of rotating and translating bodies. *International Journal for Numerical Methods in Fluids*, *38*(2), 187-206.
[3] Madani, S. H., Wissink, J., & Bahai, H. (2013). A comparative study of immersed-boundary interpolation methods for a flow around a stationary cylinder at low Reynolds number, Coupled Problems 2013, 2013, pp. 585 - 596
[4] Peskin, C. S., (1982), "The Fluid Dynamics of Heart Valves: Experimental, Theoretical, and Computational Methods," Annual Review of Fluid Mechanics, Vol. 14, pp. 235–259.
[5] Mittal, R., and Iaccarino, G. , (2005), "Immersed Boundary Methods," Annual Review of Fluid Mechanics, Vol. 37, pp. 239–261. doi:10.1146/annurev.fluid.37.061903.175743
[6] Lee, I., & Choi, H. (2015). A discrete-forcing immersed boundary method for the fluid–structure interaction of an elastic slender body. *Journal of Computational Physics*, *280*, 529-546.
[7] Kang S., Iaccarino G., Moin P., (2009), Accurate immersed-boundary reconstructions for viscous flow simulations, AIAA J. 47 (7) 1750–1760.
[8] Harlow F.H. and welsh J.E., Numerical Calculation of time-dependent viscous incomperissible Flow of Fluid with Free surface, The Physics of Fluids, Vol. 8, No.12, 1965
[9] Ferziger J.H and Peric M, (2002), Computational methods for fluid dynamics, 3rd edition, Springer.
[10] Braza, M., Chassaing, p., Ha Minh, H. (1986), Numerical study and physical analysis of the pressure and velocity fields in the near waked of a circular cylinder, journal of fluid mechanics, vol. 165, pp. 79-130.
[11] Meneghini, J. R., & Bearman, P. W. (1995). Numerical simulation of high amplitude oscillatory flow about a circular cylinder. *Journal of Fluids and Structures*, *9*(4), 435-455.
[12] Park, J., Kwon, K., Choi, H., (1998), Numerical solutions of flow past a circular cylinder at Reynolds numbers up to 160, KSME international journal, vol. 12, no, 6, pp.1200-1205.
[13] Lai, M.C., Peskin C.S. (2000), An immersed boundary method with formal second-order accuracy and reduced numerical viscosity, Journal of Computational. Phys, vol. 160, pp. 705–19.
[14] Kim, J., Kim, D., Choi, H. (2001), An immersed boundary finite volume method for simulation of flow in complex geometries, Journal of computational physics vol.171, pp.132-140.
[15] Borazjani, I., Ge, L., Sotiropoulos, F. (2008), Curvilinear immersed boundary method for simulating fluid structure interaction with complex 3D rigid bodies, Journal of Computational Physics vol. 227 pp. 7587–7620,
[16] Ahn, H.T. and Kallinderis, Y. (2006), strongly coupled flow/structure interactions with a geometrically conservative ALE scheme on general hybrid meshes, Journal of Computational Physics, 219(2), pp. 671-696.

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis(Eds)

A NEW VERY LARGE EDDY SIMULATION MODEL IN THE CONTEXT OF FLUID-STRUCTURE INTERACTION

A. KONDRATYUK*, M. SCHÄFER†

* Graduate School of Computational Engineering (GSC CE)
Technical University Darmstadt
Dolivostrae 15, 64293 Darmstadt, Germany
e-mail: kondratyuk@gsc.tu-darmstadt.de, web page: http://www.graduate-school-ce.de

† Institute of Numerical Methods in Mechanical Engineering
Technical University Darmstadt
Dolivostrae 15, 64293 Darmstadt, Germany
e-mail: schaefer@fnb.tu-darmstadt.de, web page: http://www.fnb.tu-darmstadt.de

Key words: Turbulent Flow, Very Large Eddy Simulation, Hybrid Methods, Fluid-Structure Interaction

Abstract. In the present work one of the hybrid VLES (Very Large Eddy Simulation) turbulence models is investigated in the context of fluid-structure interaction (FSI). Firstly, the formulation of the VLES model for two different RANS models ($k-\varepsilon$ and $\zeta-f$) is validated with a fully-developed channel flow at a turbulent Reynolds number of $Re = 395$. Then, this model is used to calculate the flow over an inclining plate in order to investigate the potential of VLES for moving structures. The results of simulations using two different background RANS models are compared to URANS and DDES results. In addition, the simulation results for different underlying RANS models are discussed.

1 INTRODUCTION

Direct Numerical Simulation (DNS) of turbulence can only be applied to flows with very simple geometries and small Reynolds numbers, because a complete resolution of the turbulent structures is necessary. Another possibility is Large Eddy Simulation (LES), where the big energy containing turbulent scales are resolved, while the small ones are modeled. In this case the computational costs increase very quickly with increasing Reynolds number. This is why in many industrial fields Reynolds-Average-Navier-Stokes (RANS) models are still the preferred method for the prediction of turbulent flows. In this approach all turbulent structures are modeled. Therefore, the results obtained with RANS models are often not satisfactory for many kind of flows, in particular, for massively separating flows.

In the last decade so-called hybrid turbulence models became increasingly popular. Compared to LES and DNS they deliver satisfactory results while demanding reduced computational costs. The underlying idea is to combine the advantages of different modeling approaches. These new models make it possible to solve also complex industrial problems.

The most popular hybrid turbulence model which has also been successfully used for many complex turbulent flow tasks is the Detached Eddy Simulation (DES), which was first proposed by Spalart [11]. It combines a RANS mode in the attached boundary layers with LES in separated regions and regions far from the wall. The complication in the applying of DES is the "gray area", in which an undefined modeling zone exists. In this area the solution is neither pure RANS nor pure LES [5].

Another kind of hybrid methodology, the so-called Very Large Eddy Simulation (VLES), was proposed by Speziale [12]. This model provides a seamless change from RANS to DNS depending on the numerical resolution. However, the original VLES model damped the Reynolds stress too much and required a fine mesh resolution. Therefore, modification were proposed in [5] or [3], with which the approach shows high efficiency and robustness in many applications [6], [5], [3].

In [4] it has been reported, that the predictive accuracy of VLES depends on the specific RANS turbulence model. This aspect is particularly observable for complex flows with movable or deformable objects. This especially occurs for flows with separations and thus justifies the use of more complex RANS models, like the $\zeta - f$ model.

The turbulent fluid-structure interaction (FSI) is currently not an established research object. However, especially in the case of FSI the problems become much more demanding in terms of computational cost owing to additional equations of motion for the structural part, which have to be solved together with the flow equations within a coupled solution procedure. Therefore, a reduction of computing times is especially important in the context of FSI. The potential of hybrid modeling employed to FSI is mentioned in [14], while some experiences are described in [2]. In these works the DES model has been used and investigated. Studies on the behavior of the VLES model in the context of FSI are rare.

In this paper $k - \varepsilon$ and $\zeta - f$ VLES models are investigated in the context of FSI. First, the VLES models are validated by computing the attached flow in a channel at a Reynolds number of $Re = 395$. Finally, the models are applied to investigate a flat plate which inclines at a constant angular velocity from $0°$ to $45°$ at $Re = 30000$. The focus lies on the investigation of the results from the VLES model with different basic RANS models and the comparison of the results to DES and URANS results.

2 GOVERNING EQUATIONS

2.1 Reynolds Averaged Navier Stokes equations

In the present study an incompressible Newtonian fluid with constant fluid properties is considered. The Navier-Stokes equations describe the conservation of mass and momentum for such fluids. Using the Reynolds decomposition and time averaging these equations lead to the so-called RANS equations with an additional term, which arises due to the averaging:

$$\frac{\partial \overline{u_i}}{\partial x_i} = 0 \,, \tag{1}$$

$$\frac{\partial \overline{u_i}}{\partial t} + \overline{u_j}\frac{\partial \overline{u_i}}{\partial x_j} = \frac{\partial}{\partial x_j}\left(\nu \frac{\partial \overline{u_i}}{\partial x_j} - \tau_{ij}\right) - \frac{1}{\rho}\frac{\partial \overline{p}}{\partial x_j}, \tag{2}$$

where $\overline{u_i}, \overline{p}$ and ν are mean velocity components, mean pressure and viscosity, respectively. τ_{ij} represents the Reynolds stress tensor, which can be represented by the Boussinesq's approximation as

$$\tau_{ij} = -\overline{u_i' u_j'} = 2\nu_t S_{ij} - \frac{2}{3}k\delta_{ij}, \quad S_{ij} = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right). \tag{3}$$

Here ν_t presents a turbulent viscosity, which has to be modeled. In this paper only the $\zeta - f$ model is presented, a detailed description of the $k - \varepsilon$ model can be found in [8].

2.2 $\zeta - f$ model

The $\zeta - f$ RANS model developed by Hinjalic et al. [7] uses a transport equation for the velocity scales ratio $\zeta = \overline{v^2}/k$ and the equation of the so-called elliptic relaxation function f, additionally to the equations for turbulent kinetic energy k and its dissipations rate ε:

$$\frac{\partial k}{\partial t} + u_j\frac{\partial k}{\partial x_j} = P_k - \varepsilon + \frac{\partial}{\partial x_j}\left[\left(\nu + \frac{\nu_t}{\sigma_k}\frac{\partial k}{\partial x_j}\right)\right], \tag{4}$$

$$\frac{\partial \varepsilon}{\partial t} + u_j\frac{\partial \varepsilon}{\partial x_j} = \frac{C_{\varepsilon_1}P - C_{\varepsilon_2}\varepsilon}{T} + \frac{\partial}{\partial x_j}\left[\left(\nu + \frac{\nu_t}{\sigma_\varepsilon}\frac{\partial \varepsilon}{\partial x_j}\right)\right], \tag{5}$$

$$\frac{\partial \zeta}{\partial t} + u_j\frac{\partial \zeta}{\partial x_j} = f - \frac{P}{k}\zeta + \frac{\partial}{\partial x_j}\left[\left(\nu + \frac{\nu_t}{\sigma_\zeta}\frac{\partial \zeta}{\partial x_j}\right)\right], \tag{6}$$

$$L^2\Delta f - f = \frac{1}{T}\left(C_1 + C_2\frac{P}{\varepsilon}\left(\zeta - \frac{2}{3}\right)\right). \tag{7}$$

The corresponding turbulence viscosity is defined as

$$\nu_t = C_\mu^\zeta \zeta k T, \tag{8}$$

where T is the turbulent time scale and C_μ^ζ is a model constant. The coefficients and a detailed description of this model can be found in [6].

This eddy-viscosity-based model yields better results in comparison to other RANS models for the wall-bounded flows [6]. Since the predictive accuracy of VLES depends on the specific RANS turbulence model [4], the application of the $\zeta - f$ model as a background RANS model for VLES appears to be promising.

3 VLES MODEL

The VLES approach switches from RANS to DNS depending on the numerical resolution. Between these two limits a LES will be recovered [3]. The switching is realized through the rescaling of the subscale stress resolution control function F_r that is introduced in this approach:

$$\tau_{ij}^{sub} = F_r \tau_{ij}^{RANS}. \tag{9}$$

F_r is a function of two length scales: the turbulent length scale L_c related to the spectral cut-off and the integral length scale L_i ($\propto k^{3/2}/\varepsilon$):

$$F_r = min\left[1, \left(\frac{L_c}{L_i}\right)^{\frac{3}{4}}\right]. \tag{10}$$

F_r gets a value between one and zero. When F_r approaches 0, then all scales are resolved and the VLES model behaves like a DNS. In the near-wall region $F_r \to 1$, because $L_c > L_i$ and the model works as a RANS model, what is similar to the DES concept. In [3] a detailed description of the VLES approach and the resolution control function F_r can be found.

The VLES model can be blended with any trusted RANS turbulence model. In this paper it was implemented with the standard $k-\varepsilon$ model [8] and with the $\zeta-f$ model [6] described above. Compared to basic RANS models the VLES modifies only the formulation of the turbulent viscosity. For example, for the $\zeta - f$ model the turbulent viscosity takes the form

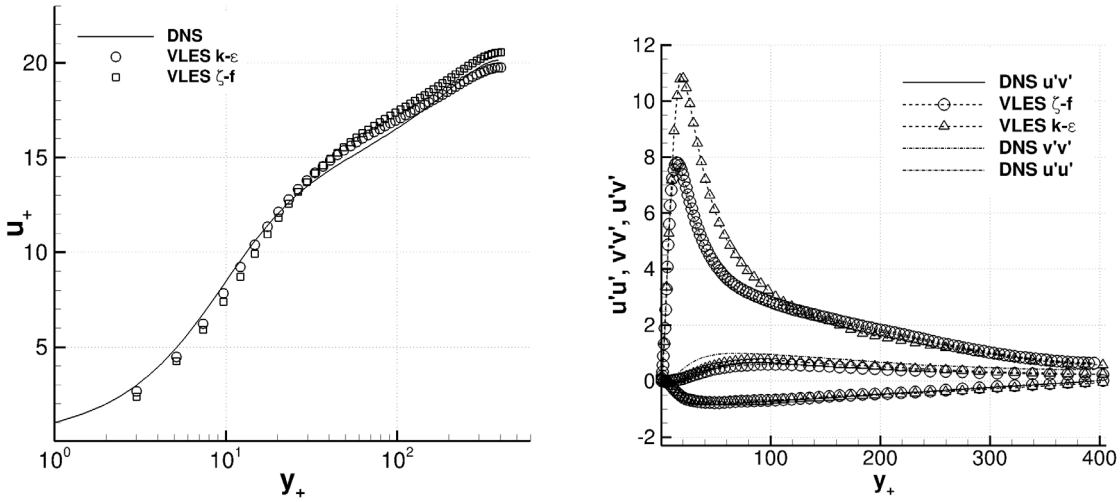
$$\nu_t = F_r C_\mu^\zeta \zeta k T. \tag{11}$$

In [5] and [3] it has been shown that the VLES approach is capable of achieving good predictions for a wide range of turbulent flows with less computational effort in comparison to LES.

4 VALIDATION

The VLES model described before is validated with a fully-developed channel flow at a turbulent Reynolds number of $Re_\tau = 395$, based on the friction velocity u_τ. This test case shows the feasibility of the VLES model to predict the attached boundary layer flows.

The size of the computational domain is given by $L_x = 2\pi, L_y = 2$ and $L_z = \pi$, for streamwise, wall-normal and spanwise direction, respectively. For the simulation a mesh with 64x100x40 grid points is used. The first grid node is located at the normalized wall distance $y^+ = 1$. Periodic boundary conditions are applied in streamwise and spanwise



(a) Mean streamwise velocity profiles (b) Reynolds stress component profiles

Figure 1: Comparison of results for channel flow

directions. The results of the calculation within the VLES model for two different basic RANS models, $k-\varepsilon$ and $\zeta-f$, are compared to the DNS data contributed by Moser [10].

Figure 1a shows the mean streamwise velocity given by different background RANS models. Some difference in the buffer layer and the log-low region can be seen. However, VLES shows good agreement to the reference for both overlying RANS models considering that the mesh is quite coarse. The RMS velocities are compared in Figure **??** (b). The VLES with the basis $\zeta-f$ model predicts the results very well in all of the three directions. The $k-\varepsilon$ shows good agreement in the $v'v'$ and $u'u'$ components, while the values for $u'v'$ are overpredicted in the buffer layer. It can be seen, that the results of the VLES models for the turbulent channel flow are in good agreement with the DNS prediction.

5 NUMERICAL METHODS AND COMPUTATIONAL DETAILS

Next, the flow around a flat plate, which is inclining from 0° to 45°, is investigated with the VLES model. The plate is mounted inside a plane channel. The computational domain is shown schematically in Figure 2. It is the same configuration as in [15]. The length of the computational domain corresponds to 2 m, the height and the depth of the channel are equal to 0.45 m. The chord length and the thickness of the plate is $c = 0.12\,\mathrm{m}$ and 0.006 m, respectively, the aspect ratio is $AR = 3.67$.

The plate is placed in the center of the channel and changes its angle of attack rapidly from 0° to 45° at an averaged rotational speed of

$$\alpha = \frac{10\pi}{3}\frac{\mathrm{rad}}{\mathrm{s}}. \tag{12}$$

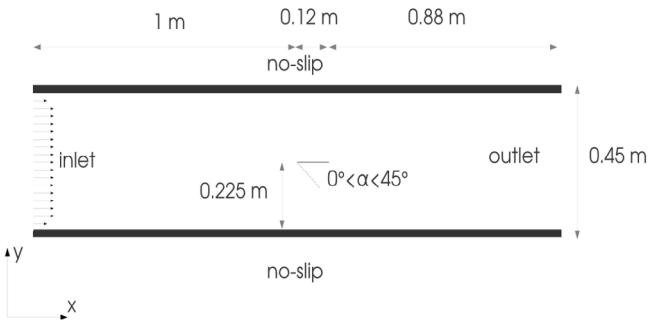


Figure 2: Sketch of computational domain

For the description of this motion the following equation is used:

$$\alpha(t) = \frac{\pi}{4}\left[\sin\left(\frac{\pi}{T}t\right)\right]^2, \tag{13}$$

where $T = 0.15\,\mathrm{s}$ equals one period. This expression makes it possible to avoid the discontinuities in the calculations. The reduced pitching frequency for a non periodic motion is defined as $k_{pitch} = \Delta\alpha c/(U_b \Delta t) = 0.168$, where $U_b = 3.75\,\mathrm{m/s}$ is the inlet bulk velocity.

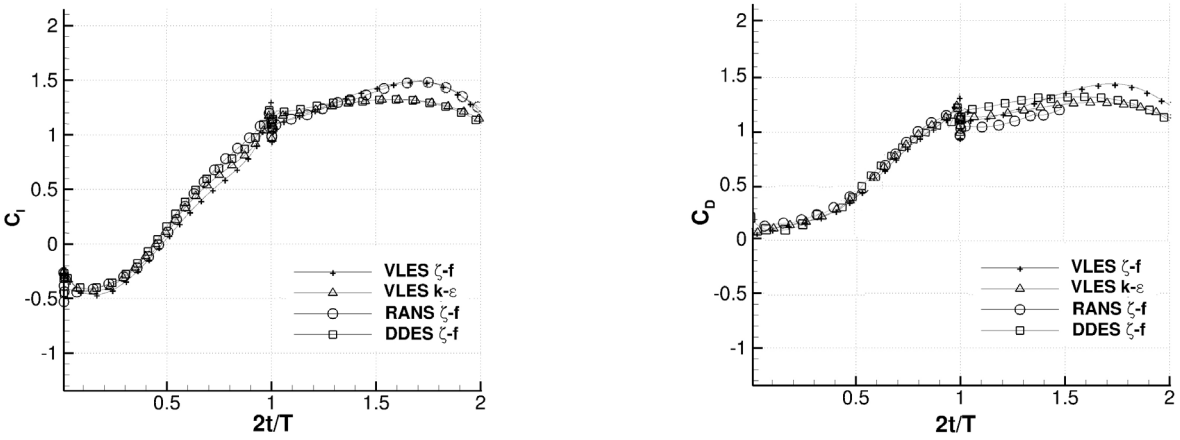
This simulation was carried out at a Reynolds number of $Re_c = 30000$. Non-slip boundary conditions are applied at the surface of the plate as well as at the top and the bottom of the computational domain. A profile of a fully developed turbulent channel flow without any perturbations is prescribed at the inlet. At the outlet, a zero gradient boundary condition is applied. Furthermore, periodic boundary conditions in the spanwise direction (z-axis) are assumed.

For all simulations the same grid with about 2.6 million CVs is applied, with 240 CVs in circumferential direction of the plate. The first node in normal grid direction is located at $y^+ = 1.0$. The mesh is clustered around the plate and this block rotates with the plate. This ensures a good grid quality in the region, where the vortex shedding and recirculation is expected. This allows to lower the computational cost, because no re-meshing and no additional grid generation methods are required in this region.

All simulations are carried out with the code FASTEST [13], which is based on the finite-volume method for block-structured grids. The parallelization in FASTEST is done via domain decomposition using MPI. For the approximation of convective and diffusive fluxes the central scheme of second-order accuracy is applied. A second-order backward differencing scheme is used for the time discretization. The coupling between pressure and velocity is done with the SIMPLE algorithm.

6 RESULTS AND DISCUSSION

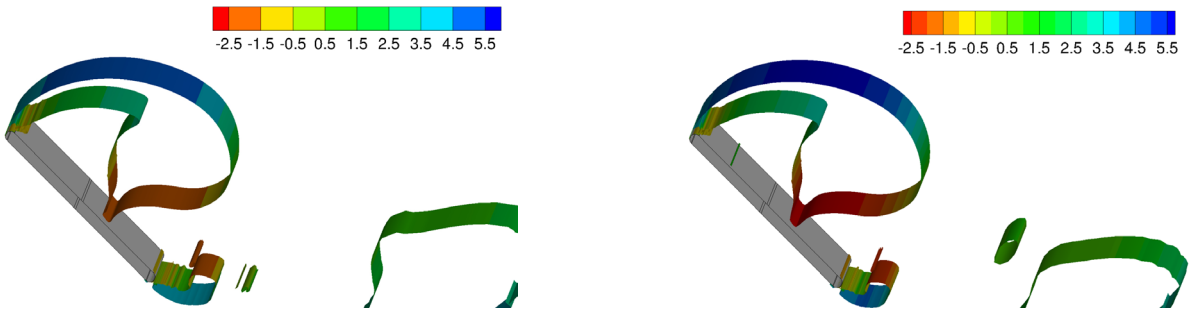
The flow over an inclined flat plate is simulated with the $\zeta - f$ and $k - \varepsilon$ VLES models. The character of the flow and the generation of the fluctuation for the quickly moving structures are investigated.



(a) Lift coefficient for $0 \leq t \leq T/2$ (b) Drag coefficient over $0 \leq t \leq T/2$

Figure 3: Comparison of lift and drag coefficients

In the beginning of the simulation the plate changes its angle of attack from 0° to 45°. In this upstroke phase no outstanding fluctuations are generated and the flow has a two-dimensional character. This effect corresponds to the results by Martian [9], who investigated the flow past a pitching NACA0012 airfoil. In this phase VLES, DDES and URANS predict similar values for the lift C_L and drag C_D coefficient (Figure 3).



(a) VLES $\zeta - f$ for $t = T/2$ (b) VLES $k - \varepsilon$ for $t = T/2$

Figure 4: Fluctuation visualized by Q-criterion at $Q = 5$, colored by streamwise velocity

Afterward, the plate remains at the constant angle of attack $\alpha = 45°$ and the C_L and C_D values deviate (see Figures 3). Between URANS and VLES $\zeta - f$ no significant

difference can be detected. In this flow phase fluctuations should start to be generated and become more dominant over time [15]. However, the VLES simulations further shows a two-dimensional character of the flow. Fluctuations have not been developed yet (Figures 4). The lack of fluctuations for hybrid turbulence models is also observed by Türk et al. [15]. There it is also shows, that the URANS approach can predict the three-dimensional character of this flow already in the beginning of the "remain"-phase.

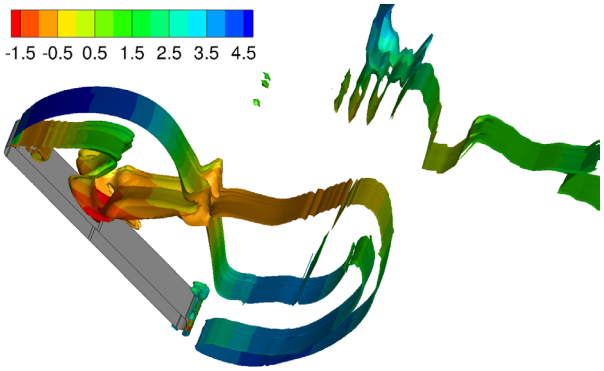


Figure 5: VLES $k-\varepsilon$ for $t = 3T$

Until $t > 3T$, the fluctuations start to form up and the flow shows a three-dimensional character (Figure 5). The values of C_L show a good agreement after this instant. C_L is underpredicted by the VLES $k-\varepsilon$ model in the beginning of the "remain"-phase, but it approaches the values of VLES $\zeta - f$ after a few iterations (Figure 6).

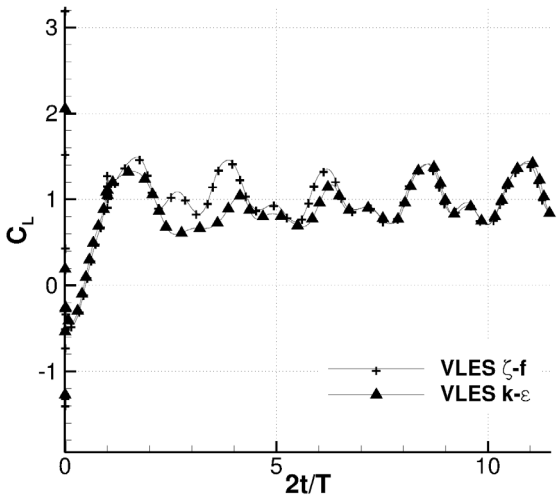
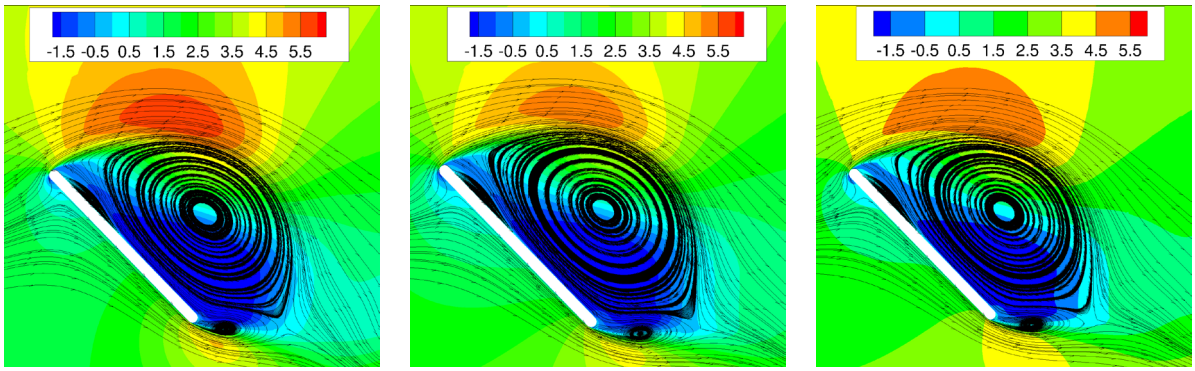


Figure 6: Comparison of lift coefficient for $0 \leq t \leq 6T$



(a) VLES $k-\varepsilon$ (b) VLES $\zeta-f$ (c) URANS $\zeta-f$

Figure 7: Contour and streamlines of streamwise velocity component for $2t/T = 1.8$

The character of the flow past an inclined plate is known from a study by Breuer et. al [1]. A large clockwise rotating vortex exists at the leeward side of the plate. This vortex forms from the flow separation at the leading edge. The small size clockwise recirculation originates at the trailing edge because of a roll-up of the shear layer. VLES simulations predict these overall flow features very well (Figure 7). An insignificant difference appears only in the location of the center of small vortices (Figure 8).

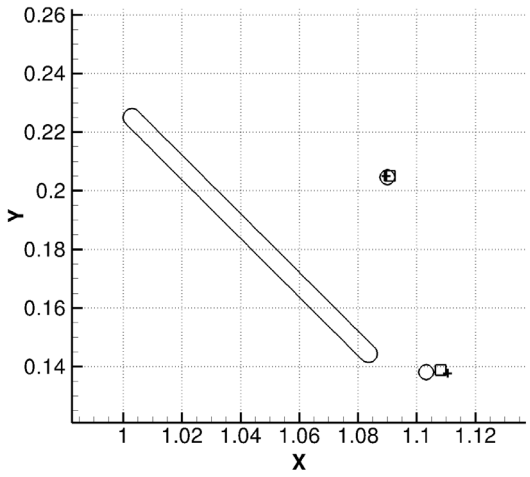


Figure 8: Location of the vortex cores, + VLES $\zeta-f$, ∘ VLES $k-\varepsilon$ and □ URANS

7 CONCLUSION

We have presented the new VLES $\zeta-f$ model. We validated this model with a turbulent channel flow at a Reynolds number of $Re = 395$. The validation result shows

good agreement with the reference DNS data. Then the flow over an inclined plate was investigated with the VLES approach and compared with URANS results. It was shown that the VLES model yields unsatisfactory results on moving structures. The VLES approach requires some computational time to predict the 3D character of the flow after the plate has come to a rest. However, the character of the flow over the fixed inclined plate was predicted very well by the VLES model.

This study is the first step to investigate the potential of VLES for moving structures. The VLES approach with a modified $\zeta - f$ model can lead to an acceleration in the development of fluctuation. Another factor that needs to be investigated is the influence of the filter width on the results within VLES model.

8 ACKNOWLEDGMENT

This work is supported by the 'Excellence Initiative' of the German Federal and State Governments within the Graduate School of Computational Engineering at Technische Universität Darmstadt.

REFERENCES

[1] Breuer, M., Jovičić, N. and Mazaev, K. Comparison of DES, RANS and LES for the separated flow around a flat plate at high incidence. *International journal for numerical methods in fluids*, 41(4), 357-388, 2003

[2] Bunge, U. Numerische Simulation turbulenter Strömungen im Kontext der Wechselwirkung zwischen Fluid und Struktur. *Univ.-Verlag der TU, Univ.-Bibliothek*, 2005

[3] Chang, C. Y., Jakirlic, S., Dietrich, K., Basara, B. and Tropea, C. Swirling flow in a tube with variably-shaped outlet orifices: An LES and VLES study. *International Journal of Heat and Fluid Flow*, 49 pp. 28-42. 2014

[4] Han, X., and Krajnović, S. "A new very large eddy simulation model for simulation of turbulent flow." *Progress in Hybrid RANS-LES Modeling*. Springer Berlin Heidelberg, pages 131-140, 2012.

[5] Han, X. and Krajnović, S. Validation of a novel very large eddy simulation method for simulation of turbulent separated flow. *International Journal for Numerical Methods in Fluids*, 73(5), pages 436-461, 2013

[6] Han, X. and Krajnović, S. An efficient very large eddy simulation model for simulation of turbulent flow. *International Journal for Numerical Methods in Fluids*, 71(11), pages 1341-1360, 2013

[7] Hanjalić, K., Popovac, M. and Hadžiabdić, M. A robust near-wall elliptic-relaxation eddy-viscosity turbulence model for CFD. *International Journal of Heat and Fluid Flow*, 25(6), pages 1047-1051, 2004

[8] Jones, W. P. and Launder, B. The prediction of laminarization with a two-equation model of turbulence. *International Journal of Heat and Mass Transfer*, 15(2), pages 301-314, 1972

[9] Martinat, G., Braza, M., Hoarau, Y. and Harran, G. Turbulence modeling of the flow past a pitching NACA0012 airfoil at 105 and 106 Reynolds numbers. *Journal of Fluids and Structures*, 24(8), 1294-1303, 2008

[10] Moser, R. D., Kim, J. and Mansour, N. N. Direct numerical simulation of turbulent channel flow up to Re= 590. *Phys. Fluids*, 11(4), 943-945, 1999.

[11] Spalart, P. R., Jou, W. H., Strelets, M., and Allmaras, S. R. Comments on the feasibility of LES for wings, and on a hybrid RANS/LES approach. *Advances in DNS/LES, 1*, pages 4-8, 1997

[12] Speziale, C. Turbulence modeling for time-dependent RANS and VLES: a review.*AIAA Journal*, 36(2), pages 173-184, 1998

[13] Sternel, D. C. FASTEST-Manual, 2005.

[14] Squires, K. D. Detached-eddy simulation: current status and perspectives. *In Direct and large-eddy simulation V*, pages 465-480. Springer Netherlands, 2004.

[15] Türk, S., Reimann, T., Sternel, D. C., and Schäfer, M. On the performance of detached-eddy simulation in fluid-structure interaction problems. *V International Conference on Computational Methods for Coupled Problems in Science and Engineering, Coupled Problems 2013*, , Barcelona, 2013. Proceedings.

[16] Wilcox, D. C. *Turbulence modeling for CFD* (Vol. 2, pp. 103-217). La Canada, CA: DCW industries, 1998.

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis(Eds)

A NUMERICAL INVESTIGATION OF THE ADDED MASS EFFECT DUE TO FLUID-STRUCTURE COUPLING IN A RECTANGULAR TANK – COUPLED PROBLEMS 2015

A. BRANDELY*, E. LEFRANCOIS†

* Roberval Laboratory CNRS UMR 6253
Univ. of Techn. of Compiègne
CS 60319, 60203 Compiègne, France
e-mail: anais.brandely@utc.fr, web page: http://www.utc.fr/

† Roberval Laboratory CNRS UMR 6253
Univ. of Techn. of Compiègne
CS 60319, 60203 Compiègne, France

Key words: Sloshing effect, partially filled tank, fluid-structure interaction, added mass, finite elements

Abstract. The objective of this paper is to present a numerical model for fluid-structure interactions (abbr. FSI). Indeed, emergence of hybrid vehicles, system such as Start & Stop and reduction of engine noises have contributed to the appearance of components previously inaudible. It is therefore necessary for fuel tank manufacturer to improve the vibro-acoustic of their product. Numerical simulation of the fluid-structure interaction is a way to identify the noise generated by fuel sloshing and to reduce costly experimental tests. Its purpose, within the context of sloshing effect in a movable partially-filled tank, is to improve understanding of interactions between fluid and the dynamics of the tank flexibly attached to the vehicles. The FSI investigation is based on an added mass corrected version of the classical strongly-coupled partitioned scheme exposed in [24]. A particular case of a closed and full tank is treated and underlines the fact that the added mass effect depends on the mass ratio between the fluid and the tank. Concerning the general case for a partially-filled tank submitted to a sudden braking, the added mass effect depends of the aspect ratio of the fluid domain. FSI calculations based on the non-corrected coupling-scheme have been conducted and show the convergence limit. For both cases the added mass term is integrated to the corrected staggered scheme and results show that the corrected version permits to systematically converge to the coupled solution. For rare cases where added mass effect is not so penalizing, it however permits to significantly reduce the iterations required.

1 INTRODUCTION

Sloshing can be defined as the motion of the free surface of the liquid in a partially filled, moving tank. When tank motion is excited with frequencies that are at some distance from the resonance frequencies of the sloshing, the response amplitude of the liquid free surface remains small and the sloshing is considered as linear. However, as the excitation period gets closer to the sloshing's natural period, its response becomes more violent and when it is coupled with the flexibility of the tank may be the source of undesired noise and critical mechanical strength (tank walls and attachments to the vehicle for example).

Numerical simulation of *fluid-structure interactions* (abbr. FSI) is a way to identify the noise resulting from fuel sloshing that can reduce the need for very expensive experimental tests. Numerical investigation of the FSI between the sloshing sollicitation and a flexible partially-filled tank, requires the use of a *partitioned* coupling: it consists in using a dedicated solver for each of the two physics. Exchanges take place regularly between the two solvers, through the use of a coupling scheme [6, 7, 8]. This coupled scheme is based on successive solutions produced by the fluid and structure solvers: the coupling is referred to as *loosely-coupled partitioned* if only one shot per time step is required for each field, and *strongly-coupled partitioned* if an iterative procedure is used to ensure convergence of the coupled solution [5].

Its major drawback is that where higher density fluids are involved, convergence is no longer guaranteed, and divergence will generally be observed regardless of the chosen time step for incompressible flows [9]. A number of approaches have been proposed to counter this drawback,[10, 11, 12], including semi-implicit discretization [13] and adaptive Aitken under-relaxation [14], but convergence is not always guaranteed, or may be slow in cases of high-density fluids such as blood or water. An attractive technique for new partitioned procedures based on Robin-type transmission conditions is proposed in [15], but this technique requires adjustable parameters to fit.

The objective of this paper is to apply the corrected FSI scheme presented in [24] to the case of a partially filled tank in two dimensions (2D). The flexibility here comes from the mountings attaching the tank to the vehicle, since the tank itself is assumed to be rigid. Particular attention is given to quantifying the effect of additional mass on convergence and on the coupling results. At the same time, this study allows predicting influence of the added mass effect according to the flexibility of attachement during an automotive braking.

2 DESCRIPTIONS OF THE DIFFERENT MODEL USED TO INVESTIGATED THE SLOSHING EFFECT IN A PARTIALLY FILLED TANK

2.1 Partially filled tank case

The tank considered here has dimensions corresponding to a small real full tank but with a simplified shape. It has been used precedently in [1] to study 2D fluid flow behavior for axial sloshing in parallelepipedic tanks. L, b are respectively length and width of the

tank. The tank of mass m is assumed rigid but flexibly attached to chassis of the moving vehicle. This flexible assembly equates to a spring of rigidity k fixed to one end to a fixed element. We denote respectively $u(t)$ and $\dot{u}(t)$ position and velocity of the tank with respect to its rest position.

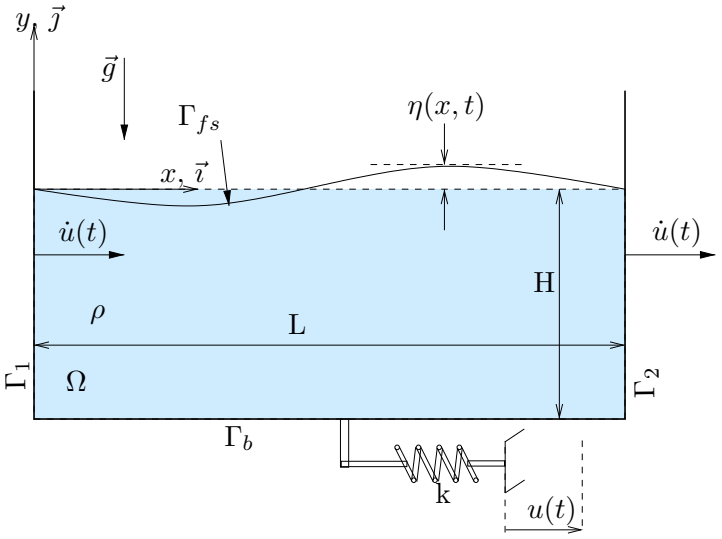


Figure 1: Definition of the moving, partially filled 2D tank with external rigidity

The tank is partially filled with a inviscid heavy fluid defined by its volumic mass ρ. Its free surface elevation is defined by $\eta(x,t)$.

2.2 Structure Model

The tank is governed by the *Fundamental Principle of Dynamics* (FPD). For a movable tank located at $u(t)$ with respect to its position at rest, its x-axis projected form may be written as:

$$m\frac{d^2u}{dt^2} + ku(t) = f_p(t), \quad \text{with} \quad u(0) = u^0 \text{ and } \dot{u}(0) = 0, \tag{1}$$

The f_p term, obtained from the integral of the pressure on the two internal faces Γ_1 and Γ_2, is the pressure effect acting on the tank's walls:

$$f_p(t) = f_{p,1}(t) + f_{p,2}(t) = b\int_0^H (p_2(y,t) - p_1(y,t))\,dy, \tag{2}$$

where b denotes the tank width.

The time resolution of the left term is here obtained using a Newmark-Wilson finite difference [22] scheme. It is based on the following time series expansions on u and $\dot{u}$ which allow the variation $\Delta u = u^{n+1} - u^n$ to be deduced from:

$$\left(\frac{2m}{b\Delta t^2}+k\right)\Delta u = f_p^n - ku^n + m\left(\frac{2}{b\Delta t}\dot{u}^n + \frac{(1-b)}{b}\ddot{u}^n\right). \tag{3}$$

It should be pointed out that the fluid load term f_p is here computed at time step n because of the partitioned nature of the considered coupling scheme.

2.3 Potential fluid model

Potential fluid flow theory is based on the assumptions such as incompressible and irrotational inviscid fluid flow. Under these assumptions it is possible [20][21] to define a velocity potential function ψ that explicitly takes the tank motion into account:

$$\vec{\nabla}\psi = \vec{V} + \dot{u}(t)\vec{i} = \vec{\nabla}\Phi + \dot{u}(t)\vec{\nabla}\psi_0, \text{ and then } \psi = \Phi + \dot{u}(t)\psi_0, \tag{4}$$

where $\vec{V}(x,y,t)$ denotes the local fluid flow velocity and is relative to the tank velocity $\dot{u}$. The potential function has been split into two components Φ and ψ_0, respectively related to the relative fluid velocity and to the wall tank velocity $\dot{u}$. Combining this with the two assumptions yields the classical Poisson equation:

$$\Delta\psi = 0 \text{ in } \Omega. \tag{5}$$

2.3.1 Free surface equation:

By combining a dynamic condition (pressure) and a kinematic condition (velocity), the free surface equation is obtained.

$$\ddot{\psi} + g\frac{\partial\psi}{\partial n} = 0 \text{ on } \Gamma_{fs}. \tag{6}$$

2.3.2 Boundary conditions on the walls:

The two vertical walls $\Gamma_{1,2}$ are assumed to be rigid but movable at velocity $\dot{u}(t)$:

$$\frac{\partial\psi}{\partial n} = -\dot{u}(t) \text{ on } \Gamma_1 \text{ and } \frac{\partial\psi}{\partial n} = \dot{u}(t) \text{ on } \Gamma_2, \tag{7}$$

where $\vec{n}$ always denotes a normal exterior vector. The boundary condition for the bottom part is:

$$\frac{\partial\psi}{\partial n} = 0 \text{ on } \Gamma_b. \tag{8}$$

The complete set of equations (5), (6), (7) and (8) governs the potential fluid flow problem.

2.3.3 Condensing the free surface problem:

In order to reduce significantly the size of the 2D problem to solve, it is condensed by projecting it on the 1D free surface. Condensing method used has been developed by Mottelet in [21].

2.4 Strongly coupled partitioned scheme

After validation and verification of the fluid model, the fluid-structure interaction equation (1) can be treated. The principle of a partitioned approach consists in allocating each of the terms to the left and right of the equal sign to its own dedicated solver (fluid and solid). The staggered nature of the coupling scheme means that both terms are successively solved in 4 steps, as illustrated in Figure 2.

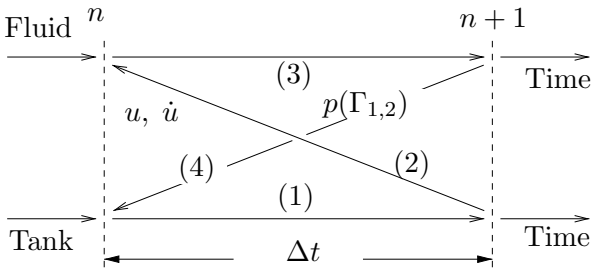


Figure 2: Partitioned and staggered coupling scheme

Indices n and $n+1$ correspond to time steps t and $t+\Delta t$ respectively. In order to reduce the discrepancy between a tank motion computed at $n+1$ from a pressure given at time n, and consequently to reinforce the coupling, an enclosed iterative procedure is then added inside the temporal loop.

2.4.1 Data passing:

The coupling scheme is based on data passing, so as to update variables common to the two physics. The tank transmits its position u and its velocity $\dot{u}$, whereas the fluid transmits its pressure field acting on the two vertical walls. We obtain:

$$p(y,t) = -\rho g y + \frac{1}{2}\rho \dot{u}^2 - \rho \frac{d\psi}{dt} \text{ on } \Gamma_{1,2}, \tag{9}$$

where $\dfrac{d\psi}{dt}$ is computed from a finite difference scheme for which the order of accuracy may have an impact on the convergence properties.

2.4.2 Energy considerations:

This energy conservation (in absence of a physical damping effect and no forced excitation, the sum of the total energy of the fluid and the tank remains constant over time and equal to the energy resulting from the initial conditions) is one of the key elements that ensure high quality FSI calculations.

The total energy of the fluid part results from the kinetic and potential (gravity) components:

$$\mathcal{E}_{FL}(t) \quad = \quad <\varphi> \frac{\rho b}{2g}[K_\varphi]\{\varphi\} + <\eta> \frac{\rho\, g\, b}{2}[M_\varphi]\{\eta\}. \tag{10}$$

The free surface elevation η is computed from dynamic free surface condition. The mechanical energy of the tank also results from two components:

$$\mathcal{E}_{TK}(t) = \frac{1}{2}m\dot{u}^2 + \frac{1}{2}ku^2.$$

3 LIMITS OF THE COUPLING SCHEME DUE TO THE ADDED MASS EFFECT AND CORRECTED SCHEME

3.1 Description of particular case of a closed full tank

In addition to the analytical solution provided, the particular case of a closed full tank represents an extreme case, that is the validation of an FSI where only the mass effect of the fluid interacts with the dynamics of the tank.

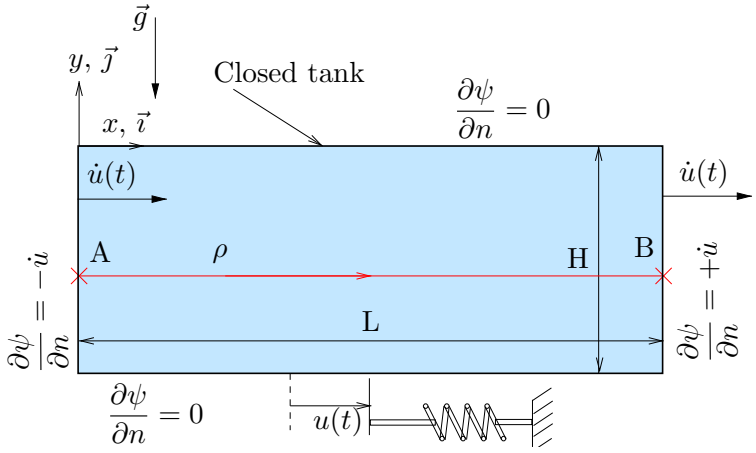


Figure 3: Particular case of a closed, full tank

The density of the considered fluid is equal to $\rho = 1000\, kg/m^3$. Naturally introducing the added mass term m_{add} that is identical to the fluid mass. Injecting it into Equation (1) yields the classical and predictable mass-spring form:

$$(m + m_{add})\frac{d^2u}{dt^2} + ku = 0, \quad \text{with} \quad u(0) = u_o,$$

whose exact solution is:

$$u(t) = u_o \cos(\omega_c t), \quad \text{with} \quad \omega_c^2 = \frac{k}{m + m_{add}}, \tag{11}$$

where ω_c denotes the pulsation of the coupled system.

3.2 Convergence analysis of the FSI scheme:

From equations (1) and (9), the iterative coupling process is then given by:

$$m\ddot{u}^{i+1} + ku^{i+1} = -m_f \left(\frac{\dot{u}^i - \dot{u}^n}{\Delta t} \right). \tag{12}$$

The term $d\psi/dt$ required for the pressure calculation is here discretized according to a first-order finite difference scheme. Convergence analysis requires it to be rewritten in the form:

$$u^{i+1} = f(u^i) \quad \text{with} \quad u^o,$$

for which convergence is ensured if $f(u)$ is continuously differentiable, $|G| = |df/du| < 1$ and u^o sufficiently close to the solution. When combined with Equation (3) taken at iterations $i+1$ and i, Equation (12) can be rewritten as:

$$u^{i+1} = -\frac{m_r}{2 + \Delta t^2 \omega^2/2} u_i + \dots$$

The term $m_r = m_f/m$ is the mass ratio and $\omega = k/m$ denotes the natural pulsation. The time step Δt can be rewritten as:

$$T = N\ \Delta t \quad \text{with} \quad T = \frac{2\pi}{\omega} \quad \text{and then} \quad \Delta t^2 \omega^2 = \frac{4\pi^2}{N^2}.$$

where N denotes the sample of the period T. We then identify:

$$f(u^i) = G\ u^i + ... \quad \text{with} \quad G = -\frac{m_r}{2 + 2\pi^2/N^2}. \tag{13}$$

The convergence is then ensured if:

$$\left| \frac{df}{du} \right| = |G| < 1 \quad \Rightarrow \quad m_r < 2 \left(1 + \frac{\pi^2}{N^2} \right). \tag{14}$$

We consider a 2nd order of accuracy for $d\psi/dt$, we can easily show that the corresponding critical values for m_r are respectively multiplied by 2/3.

$$\left. \frac{d\psi}{dt}^i \right|_{2nd} = \frac{\frac{3}{2}u^i - 2u^n + \frac{1}{2}u^{n-1}}{\Delta t} + \Delta t^2(\dots),$$

It is to be noticed that increasing the order of accuracy is extremely penalizing for convergence.

For particular full tank closed case presented in Table 1, predicted limit mass ratio m_r ensuring convergence according to a 2nd order accuracy for time discretization is equal to 1.24.

Parameter	$u_o\,[m]$	$m\,[kg]$	$k\,[N/m]$	$L\,[m]$	$H\,[m]$	$\Delta t\,[s]$	$nstep$	N
Value	0.01	50	10^4	0.5	0.225	10^{-2}	10	45

Table 1: Parameters for full, closed tank coupling

3.3 Corrected staggered scheme with added mass effect

The correction that we introduce to improve the convergence is similar to the one described by Lefrançois in [24]. It is based on the observation that increasing the inertial term or decreasing the force term will favor the iterative process. The main idea is to add an inertial term on both sides of the equation, function of the added mass, such that:

$$m_{add}\ddot{u}^{i+1} + \quad m\ddot{u}^{i+1} + ku^{i+1} \approx f_p^i \quad + m_{add}\ddot{u}^i. \tag{15}$$

where the contributions of the left- and right-hand terms are taken at iterations $i+1$ and i respectively. If convergence is reached, $\ddot{u}^i = \ddot{u}^{i+1}$ then the original expression (Equation (1)) is exactly satisfied. Analysis of the convergence is similar to [24], the criterion for convergence being:

$$G = \frac{m_r}{4(1+m_r)(1+\pi^2/N^2)} \quad \text{that always satisfies} \quad 0 < G < 1 \quad \forall\, m_r.$$

The convergence of the corrected scheme is therefore always ensured.

3.4 Results of non-corrected and corrected staggered scheme with added mass effect on the full closed tank case

To validate the predicted results of the effect of mass ratio on convergence, FSI calculations were carried out for the case illustrated in Figure 3 with different values of m_r. The number of iterations (mean value) required for convergence is given in Table 2. Each column corresponds to a mass ratio (function of ρ) and the line to 2nd order of accuracy for $d\psi/dt$. The '-' sign denotes a divergence result.
Results obtained with corrected staggered scheme with added mass effect are presented in the Table 2. Convergence is systematically reached whatever the mass ratio, which confirms the beneficial effect of the correction for the coupling scheme. The FEM solutions perfectly match (not shown here) the predicted coupled solution given by (11) in terms of pressure, displacement and coupled pulsation ω_c. Moreover, the number of iterations required are lower than for the classical case.

m_r	0.225	0.45	0.9	1.32	1.57	4.5
Non-corrected scheme	6	9	22	850	-	-
Corrected scheme	4	5	5	6	6	6

Table 2: Number of iterations for FSI convergence by using the non-corrected and corrected staggered coupling scheme at second accuracy order

4 RESULTS OBTAINED ON PARTIALLY FILLED TANKS SUBMITTED TO AN AUTOMOTIVE BREAKING

4.1 Comparison between numerical and experimental results when the tank is rigid

4.1.1 Added mass estimation

For general cases, as illustrated in Figure 2, the added mass effect is related to a matrix denoted by $[M_{add}]$ as detailed in [24]. Since we have here only one dof for the tank dynamics, the added mass term is given by a scalar. The calculation of this term m_{add} is based [25] on the assumptions that the fluid flow is inviscid and that the convective effects can be neglected with regard to the pressure gradient field.

4.2 Influence of the flexible tank attachment on sloshing and validation of beneficial effects of the corrected scheme for sloshing FSI simulation

4.2.1 Comparison between experimental and numerical forces result in case of rigid attachment:

Some experimental tests are in progress on a 3D tank model. The tank is made of 20 mm thick Plexiglas whose mass m is equal to 30 kg.

The test rig used for models tests is a specific test bench which reproduces real automotive braking and the cart deceleration is mesured by a capacitive accelerometer. The global forces measurement set-up (empty tank + liquid) is mainly based on four Kistler force sensors (type 9327A) that are mounted between two aluminum plates and can measure compression and tensile forces during tests without deformation (rigid case) of sensors or tank's attachment. This installation allows measuring dynamics and quasi-static forces engendered during sloshing with 1kH sampling rate. Experimental global force results, engendered only by sloshing, are compared to numerical results issued from two numerical codes. The first one is the linear model presented previously in section 2 and the second one is a commercial non-linear CFD code - Star CCM+. The imposed acceleration (Figure 4) is included in the numerical models and the term of rigidity k is considered as infinite.

As expected in Figure 5, during a sudden braking, amplitude of sloshing is important. Therefore, some discontinuities at free surfaces are observable and linear resolution are not enough accurate [1]. For example in Figure 5(b), the hydrodynamic impact of fluid

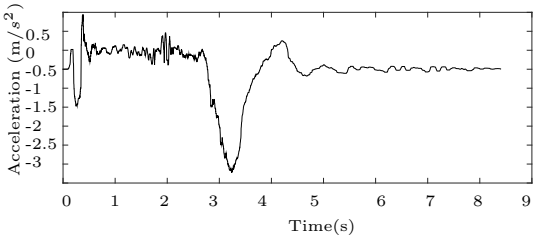
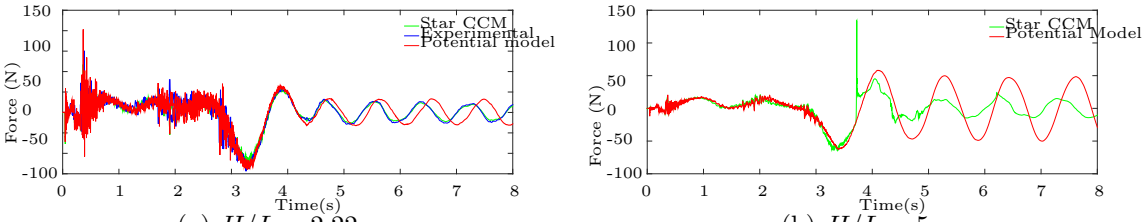


Figure 4: Profil of an automotive experimental acceleration



(a) $H/L = 2.22$ (b) $H/L = 5$

Figure 5: Global X-axis force time histories in case of rigid attachment

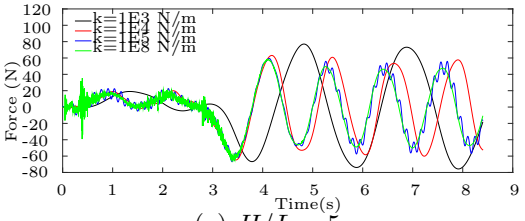
with air entrapment can not be taken into account by potential theory. However the linear code is kept on using in this study because it allows validating our coupling scheme at a lower cost.

4.2.2 Predicted sloshing force results according to the flexibility attachment and explaination by cinetic and beneficial effects of the corrected scheme for sloshing FSI simulation:

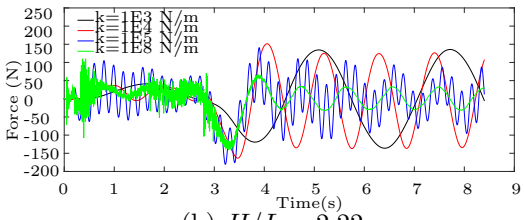
In this section, the aim is to predict the fluid force engendered by sloshing according to the flexibility of the mountings. The minimum and maximum values of k correspond respectively to a flexible and quasi-rigid mounting.

Results put in evidence that when the coupling frequency is closed to a low natural frequency of the sloshing phenomenon, global fluid x-axis force are maximum. For $L/H = 5$,maximal global fluid force appears when the rigidity of the attachment corresponds to $1e^3 N/m$ and for $L/H = 2.22$, it appears when k is assumed to be equal to $1e^4 N/m$.

In the same way, flexibility value impacts the number of iteration (Figure 7). Moreover, we can note that the slopes of the corrected coupling scheme case are lower than non-corrected one. It induces that the corrected coupling scheme allow limiting the number of iterations whatever the rigidity value, the filling level or the imposed motion.



(a) $H/L = 5$



(b) $H/L = 2.22$

Figure 6: Global X-axis force time histories in case of flexible attachment

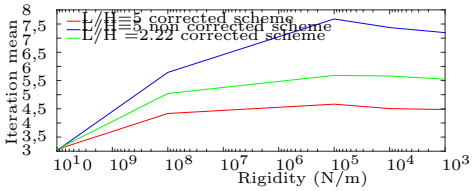


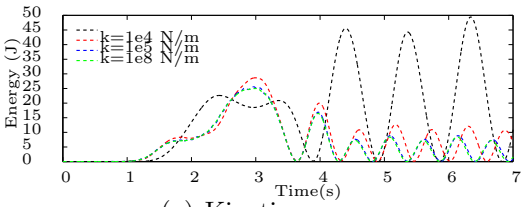
Figure 7: Number of iteration in function of rigidity k

H/L	2.22	5
Madd	51.76	10.75

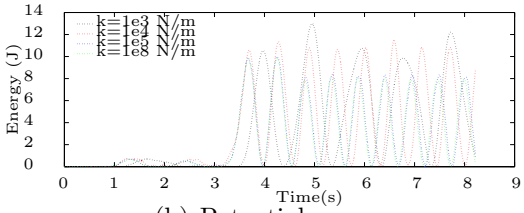
Figure 8: Aspect ratio and according added mass

4.2.3 Numerical investigation of the link between the added mass and the aspect ratio L/H:

We can observe a relation between the convergence property and the aspect ratio: the lower the aspect ratio is, the higher the added mass term is (Figure 8), and consequently the greater the difficulty for a non-corrected coupling scheme to converge is (Figure 7).



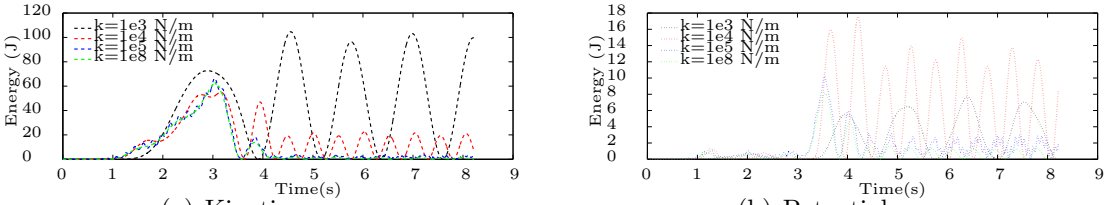
(a) Kinetic energy



(b) Potential energy

Figure 9: Fluid energy time histories, H/L=5

Unsurprisingly, given that low aspect ratio values react as a fluid column strongly interacting with the walls, higher values have a longer free surface over which energy can be spread, resulting in less interaction with the walls. In order to confirm this interpretation, time histories for fluid and tank energies are plotted in Figure 10(a) and Figure 9(a) respectively for $L/H = 2.22$ (hight filling) and 5. Kinetic and potential components are plotted as solid and dashed lines respectively whatever the rigidity. The most critical case for convergence is for $L/H = 2.22$ and it can clearly be seen that the level of energy exchange is lower than in the other case. Indeed, the potential component for the fluid is relativly lower regarding to the kinetic component (Figure 10(a) and (b)).



(a) Kinetic energy (b) Potential energy

Figure 10: Fluid energy time histories, H/L=2.22

5 CONCLUSION AND PERSPECTIVES

This paper presents a corrected version of a strongly coupled partitioned FSI scheme for studying the sloshing effect in a partially filled rectangular tank. It proves that in the particular case of a closed, full tank, the iterative convergence of a classical FSI partitioned scheme ceases to be guaranteed once the mass ratio (fluid mass divided by tank mass) exceeds a predictable value. Correcting the FSI scheme using an added mass effect, as detailed in [24], allows convergence to be ensured whatever the mass ratio. For the general case of partially filled rectangular tanks submitted to an automotive braking, this scheme significantly reduces the number of iterations required for convergence in low-density cases that already converge, and ensures convergence in cases where the classical FSI scheme fails to converge. This study allows predicting the effect of the flexibility attachment on the global sloshing forces whatever the aspect ratio.

Future work will focus on solving sloshing cases taking into account complex geometric tank shapes and flexible walls by using a non linear fluid code and by implementing the corrected scheme in a non-linear structure model. The goal is to obtain more accurate numerical data to compare with experimental data.

6 ACKNOWLEDGMENTS

Some of this work benefited from the support of PILCAM2 (Plateforme Inter-Laboratoires de CAlcul Distribués) computer resources. The authors wish to thank everyone at Plastic Omnium Auto Inergy Division for the support they provided.

REFERENCES

[1] A. Brandely, J.-S. Schotté, E. Lefrançois, B. Hagège and R. Ohayon, Simulation of the dynamic response of a sloshing liquid to horizontal movements of the tank, *ASME, Pressure Vessels & Piping Conference PVP2014, July 20-24, Anaheim, California, USA*, 2014

[2] Veldman, A. E. P., Gerrits, J., Luppes, R., Helder, J. A., and Vreeburg, J. P. B., The numerical simulation of liquid sloshing on board spacecraft. *Journal of Computational Physics, 224(6), pp. 82-99, 2007.*

[3] Michler C, van Brummelen E, Hulshoff S, de Borst R. The relevance of conservation for stability and accuracy of numerical methods for fluid-structure interaction. *Comput. Methods Appl. Mech. Engrg. 192, 4195-4215* 2003.

[4] van Brummelen E, Hulshoff S, de Borst R. Energy conservation under incompatibility for fluid-structure interaction problems. *Computer Methods in Applied Mechanics and Engineering, Volume 192, Issue 25, 2727-2748* 2003.

[5] Idelsohn SR, Del Pin F, Rossi R, Oñate E. Fluid-structure interaction problems with strong added-mass effect. *International Journal for Numerical Methods in Engineering* 2009; **80**(10):1261–1294.

[6] Piperno S, Farhat C, Larrouturou B. Partitioned procedures for the transient solution of coupled aeroelastic problems. part I: Model problem, theory and two-dimensional application. *Comput. Methods Appl. Mech. Engrg. 124, pages 79-112*, 1995.

[7] Farhat C, van der Zee KG, Geuzaine P. Provably second-order time-accurate loosely-coupled solution algorithms for transient nonlinear computational aeroelasticity. *Comput. Methods Appl. Mech. Engrg. 195, 1973-2001* 2006.

[8] Felippa CA, Park KC, Farhat C. Partitioned analysis of coupled mechanical systems. *Computer Methods in Applied Mechanics and Engineering, Vol. 190, Issues 24-25, 3247-3270* 2001.

[9] van Brummelen E.H. Added mass effects of compressible and incompressible flows in fluid-structure interaction, *J. Appl. Mech.* 76 (2009), 021206-7.

[10] Hughes TJR, Liu WK, Zimmermann TK. Lagrangian-eulerian finite element formulation for incompressible viscous flows. *Comput. Meths. Appl. Mech. Engrg., 29, pages 329-349* 1981.

[11] Fernández M, Gerbeau JF, Grandmont C. A projection algorithm for fluid-structure interaction problems with strong added-mass effect. *Comptes Rendus Académie des Sciences, Paris, Vol. 342, Issue 4, 279-284* 2006.

[12] Tallec PL, Gerbeau JF, Hauret P, Vidrascu M. Fluid structure interaction problems in large deformation. *Comptes Rendus Mecanique, Vol. 333, Issue 12, 910-922* 2005.

[13] Causin P, Gerbeau J, Nobile F. Added-mass effect in the design of partitioned algorithms for fluid-structure problems. *Computer Methods in Applied Mechanics and Engineering, Vol. 194, Issues 42-44, 4506-4527* 2005.

[14] Küttler U., and Wall W., Fixed-point fluid-structure interaction solvers with dynamic relaxation, *Computational Mechanics* 43 (2008), no. 1, 61-72.

[15] Badia S., Nobile F., Vergara C. Fluid-structure partitioned procedures based on Robin transmission conditions. *J. Comput. Phys.* 227 (2008), 7027-7051.

[16] Michler C., van Brummelen E.H., and de Borst R., An interface Newton-Krylov solver for fluid-structure interaction, *Int. J. Numer. Meth. Fluids* 47 (2005), 1189-1195

[17] Tezduyar T.E., Finite element methods for fluid dynamics with moving boundaries and interfaces, *Encyclopedia of Computational Mechanics* (E. Stein, R. de Borst, and T.J.R. Hughes, eds.), vol. 3, John Wiley and Sons, 2004.

[18] Tezduyar T.E., Sathe S., Keedy R., Stein K., Space-time finite element techniques for computation of fluid-structure interactions, *Comput. Methods Appl. Mech. Engrg.* 195 (2006), 2002-2027.

[19] Connell B.S.H., Yue D.K.P. Flapping dynamics of a flag in a uniform stream. *Journal of Fluid Mechanics*, Vol. 581:33-67, 2007

[20] S. Mottelet, Controllability and stabilization of liquid vibration in a container during transportation, *39th IEEE Conference on Decision and Control, Sydney, pp. 4641-4646, vol. 5*, 2000

[21] S. Mottelet, Controllability and stabilization of a canal with wave generators, *SIAM journal on control and optimization, 38, pp. 711-735*, 2000

[22] G. Dhatt, G. Touzot, E. Lefrançois, Finite Element Method, *Wiley-ISTE*, 2012

[23] S. Piperno, C. Farhat and B. Larrouturou. Partitioned procedures for the transient solution of coupled aeroelastic problems. Part I: Model problem, theory and two-dimensional application In *Comput. Methods Appl. Mech. Engrg.* **9**(124) :79–112

[24] M. Song, E. Lefrançois and M. Rachik. Development of a partitioned algorithm for fluid-structure coupling with no fluid density dependency. In *Computer & Fluids, DOI : 10.1016/j.compfluid.2013.05.022*, pp. 190-202, 2013

[25] Brennen C. A review of added mass and fluid inertial forces. *Technical Report*, Department of the Navy, Port Hueneme, CA 1982.

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis (Eds)

AERODYNAMIC EXCITATION FORCES IN AIR CONDITIONERS WITH ROTATING FAN-MOTOR SYSTEM

T. SATO [*], S. NITA[†1], H. OTA[†2] AND K. NAGAHASHI[†2]

[*] Department of Mechanical Engineering, School of Engineering, Tokyo Denki University
5 Senju-Asahi-cho, Adachi-ku, Tokyo, 120-8551, Japan
e-mail: taichi@mail.dendai.ac.jp

[†1] Graduate School of Engineering, Tokyo Denki University
5 Senju-Asahi-cho, Adachi-ku, Tokyo, 120-8551, Japan

[†2] Hitachi Appliances Inc., Shimizu Air Conditioning Works
390 Muramatsu, Shimizu-ku, Shizuoka-shi, Shizuoka, 424-0926, Japan

Key words: Vibration, Noise, Aerodynamic Excitation Force, Outdoor unit, Air conditioner, CFD simulation.

Abstract. Aerodynamic pressure pulsation generated by rotating blades was measured using an experimental device consisting of fan blades, a motor, and pressure gauges. The motor was installed on a high-stiffness block via load cells and the motor reaction forces were measured by load cells. The aerodynamic pressure pulsation was also calculated using CFD simulation software, and based on those CFD results, motor reaction forces were also derived. Calculation results and experimental results were compared for both pressure pulsation and motor reaction forces and the calculations were found to agree well with experimental results.

1 INTRODUCTION

Many air conditioners use fan systems for heat exchange. Vibratory motion and noise in the air conditioner occur by the aerodynamic excitation forces generated by the rotating fan blades. Calculating these aerodynamic excitation forces is therefore a useful solution for reducing vibration and noise.

In this report, we investigate aerodynamic excitation forces in the outdoor unit of air conditioners (Fig. 1). Aerodynamic excitation forces in air conditioners having a fan-motor system can be divided into the following two types:

(1) aerodynamic pressure pulsation acting on the cabinet (the thin metal plate) of the air conditioner (referred to below as “aerodynamic pressure pulsation”), and

(2) reaction forces acting on the base of the motor as a result of aerodynamic pressure pulsation (referred to below as “motor reaction forces”).

In terms of prior work, Ota et al. measured the aerodynamic pressure pulsation and researched techniques for predicting structural vibration caused by the pressure pulsation [1]. Sato et al. discussed the relationship between aerodynamic pressure pulsation and motor reaction force experimentally [2]. Furukawa et al. simulated the vertical flow in a propeller

fan using large eddy simulation [3]. Watanabe et al. discussed the aerodynamic and noise characteristics of a centrifugal fan [4].

In the study reported here, we measured the aerodynamic pressure pulsation of the rotating blades using an experimental device consisting of fan blades, a motor, pressure gauges, and load cells. The motor, which has four feet at its base, was installed on a high-stiffness block via load cells. Fan blades were attached to the motor axis by screwing them on through bosses. A measurement board was installed for receiving aerodynamic pressure pulsation, which was measured using a pressure sensor attached to the measurement board.

We also calculated the aerodynamic pressure pulsation on the measurement board and fan blades using CFD simulation software, and based on those CFD results, we also derived motor reaction forces. We compared calculation results and experimental results for both pressure pulsation and motor reaction forces and found that the calculations agreed well with the results obtained by experiment.

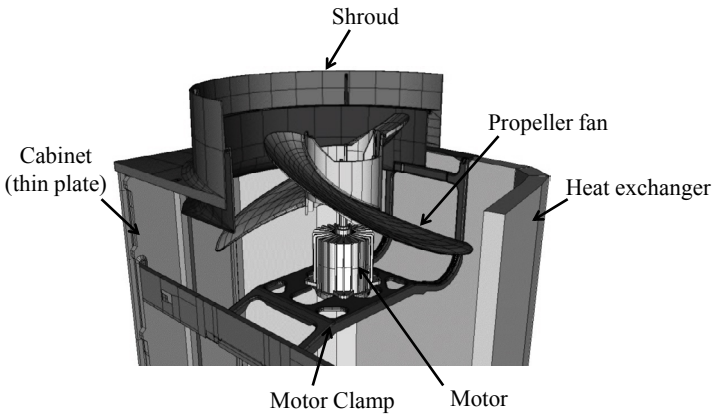
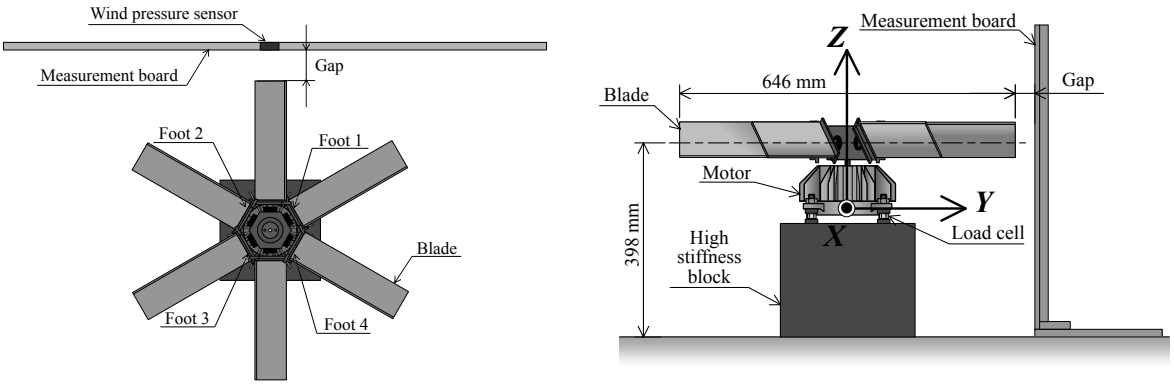


Figure 1: Outdoor unit of air conditioners



(a) Top view (b) Side view

Figure 2: Experimental setup

2 EXPERIMENTAL MEASUREMENTS AND CALCULATION OF AERODYNAMIC PRESSURE PULSATION

A top view and side view of the experimental setup are shown in Figs. 2 (a) and (b), respectively. The motor has four feet at its base. The right-front foot facing the measurement board is called “foot 1” while the other feet are numbered as “foot 2,” “foot 3,” and “foot 4” in a counterclockwise direction.

The motor is installed on a high-stiffness block via load cells. Six fan blades are screwed on to the motor axis at equal intervals through bosses, and each fan blade is installed with a 60-degree tilt from the horizontal. The direction of this tilt is the direction of upward aerodynamic flow generated by the fan.

The measurement board shown in Fig. 2 receives the aerodynamic pressure pulsation of the rotating fan blades. It corresponds to one of the thin metal plates making up the cabinet of an outdoor air conditioning unit. Here, the distance of closest approach between the measurement board and the tip of a fan blade is called the “gap.” In this experiment, we fix this gap to 10 mm and the rotating speed of the motor to 225 rpm.

A wind pressure sensor is installed on the measurement board at the position shown in Fig. 3. It is installed at a height of 398 mm corresponding to the heightwise center of the fan blades. Since the fan blades pass in front of the measurement board from right to left, the coordinate system of the measurement board is arranged accordingly with numerical signs running from minus to plus. The point corresponding to the closest approach of the fan blades (the center of the measurement board in the horizontal direction) is taken to be the origin of this coordinate system.

The results of measuring the aerodynamic pressure pulsation on the measurement board for the three measuring points shown in Fig. 3 are shown in Fig. 4 (solid line). These experimental results represent average values over time and therefore exclude the random components characteristic of aerodynamic excitation forces. We also performed a computational fluid dynamics (CFD) analysis using the standard k-ε model in fluid flow analysis software PHOENICS developed by CHAM Limited. The calculation results of this analysis (dashed line) are shown in the figure superposed on the experimental results. It can be seen that the calculations represent the experimental results well in terms of magnitude and waveform.

The results of performing a frequency analysis on the results obtained for aerodynamic pressure pulsation in Fig. 4 are shown in Fig. 5. A major peak can be observed at 22.5 Hz for both experimental results and calculations. This frequency, called the blade passing frequency (BPF), is given by the product of the rotating speed of the motor and the number of fan blades. Calculations agree well with experimental results for both the BFF (22.5 Hz) component and its higher frequency components.

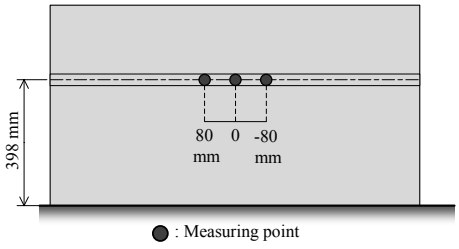
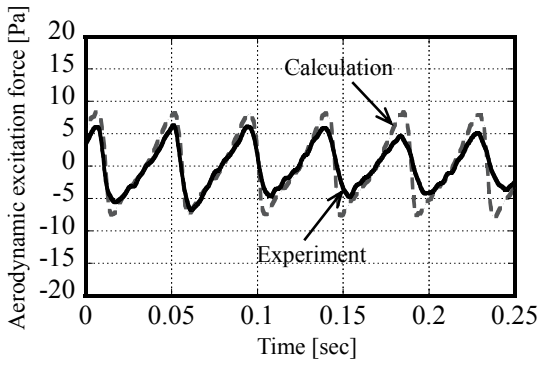
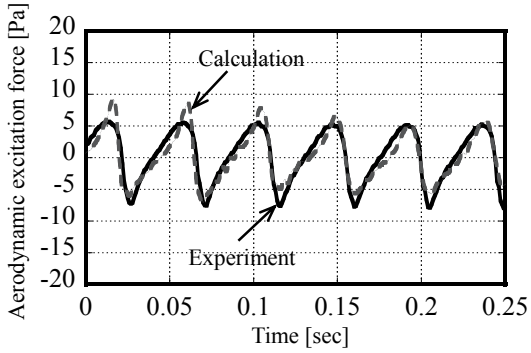


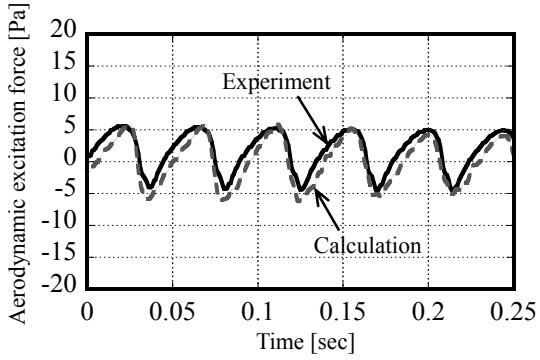
Figure 3: Three measuring points on measurement board



(a) -80mm

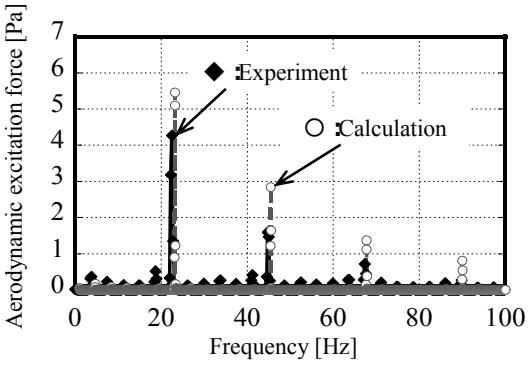


(b) 0mm

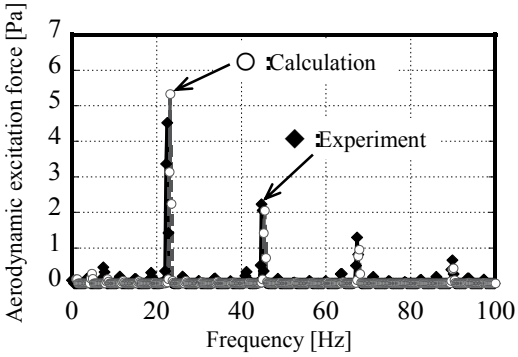


(c) 80mm

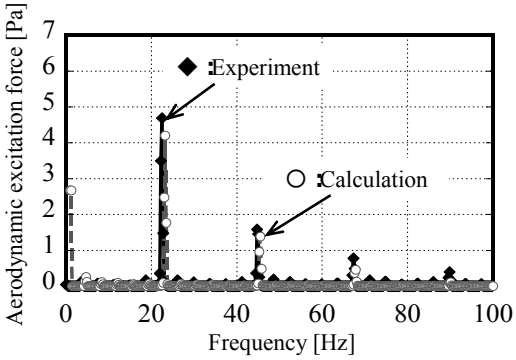
Figure 4: Time history response of aerodynamic pressure pulsation



(a) -80mm



(b) 0mm



(c) 80mm

Figure 5: FFT results of aerodynamic pressure pulsation

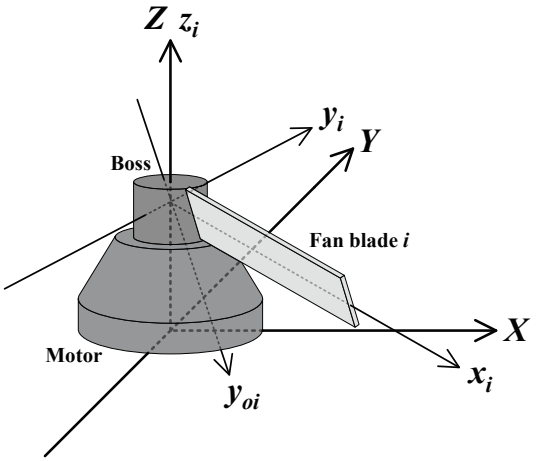


Figure 6: Coordinate system

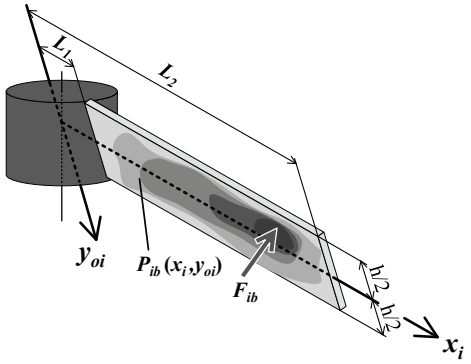


Figure 7: Pressure distribution and concentrated load

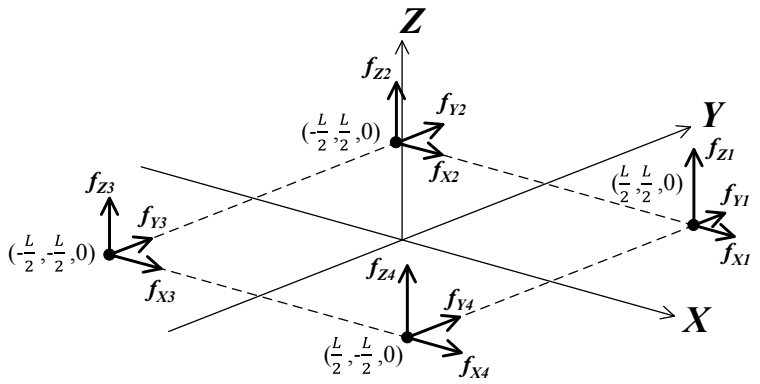


Figure 8: Force received at the motor's foot-support points

3 CALCULATION PROCEDURE AND CALCULATION/EXPERIMENTAL RESULTS OF MOTOR REACTION FORCES

We also calculated the pressure acting on the fan-blade surface in the CFD analysis described above and used those results to calculate the motor reaction forces by the procedure described below.

As shown in Fig. 6, the fan-motor system is installed in the *X*-*Y*-*Z* coordinate system of Fig. 2 fixed in space. The four motor feet lie on the *X*-*Y* plane. The origin of this *X*-*Y*-*Z* coordinate system is the point where the plane formed by the motor feet (*X*-*Y* plane) intersects the motor's axis.

The measurement board shown in Fig. 4 constitutes the plane that intersects the *Y* axis at the *Y* coordinate equal to "fan radius + gap."

In addition, the x_i-y_i-z_i coordinate system is fixed to the *i*th fan blade (*i*=1–*N*, where *N* is the number of fan blades; *N*=6 in this study), which means that it rotates with the blade. As shown in Fig. 6, the x_i-y_i plane runs parallel to the *X*-*Y* plane, but the wind-receiving surface

of fan blade i tilted with respect to the horizontal plane is separately specified as the x_i - y_{oi} plane.

Next, as shown in Fig. 7, pressure $P_{ia}(x_i, y_{oi})$ and $P_{ib}(x_i, y_{oi})$ acting on the surface of the ith fan blade can be replaced by concentrated loads F_{ia} (acting point (x_{ia}, y_{ia}, z_{ia})) and F_{ib} (acting point (x_{ib}, y_{ib}, z_{ib})) applied to the fan blade. Here, suffixes a and b denote the front surface facing the direction of movement (pressure surface) and the rear surface (suction surface), respectively. We point out here that the surface acted on by pressure $P_{ia}(x_i, y_{oi})$ is on the side of the blade not shown in the figure, so $P_{ia}(x_i, y_{oi})$ and F_{ia} do not explicitly appear in the figure.

In addition, we convert concentrated loads F_{ia} and F_{ib} to component forces (F_{Xia}, F_{Yia}, F_{Zia}) and (F_{Xib}, F_{Yib}, F_{Zib}), respectively, corresponding to the axes in the X-Y-Z coordinate system, and likewise convert concentrated-load acting points (x_{ia}, y_{ia}, z_{ia}) and (x_{ib}, y_{ib}, z_{ib}) to coordinates (X_{ia}, Y_{ia}, Z_{ia}) and (X_{ib}, Y_{ib}, Z_{ib}), respectively, in the X-Y-Z coordinate system.

As a result, we can replace the pressure acting on the fan-blade surface with the load acting at the origin of the X-Y-Z coordinate system and the moments about each axis in the X-Y-Z coordinate system as given by Eqs. (1) and (2).

$$\begin{pmatrix} F_X \\ F_Y \\ F_Z \end{pmatrix} = \sum_{i=1}^{N} \left\{ \begin{pmatrix} F_{Xia} \\ F_{Yia} \\ F_{Zia} \end{pmatrix} + \begin{pmatrix} F_{Xib} \\ F_{Yib} \\ F_{Zib} \end{pmatrix} \right\} \tag{1}$$

$$\begin{pmatrix} M_X \\ M_Y \\ M_Z \end{pmatrix} = \sum_{i=1}^{N} \left\{ \begin{pmatrix} X_{ia} \\ Y_{ia} \\ Z_{ia} \end{pmatrix} \times \begin{pmatrix} F_{Xia} \\ F_{Yia} \\ F_{Zia} \end{pmatrix} + \begin{pmatrix} X_{ib} \\ Y_{ib} \\ Z_{ib} \end{pmatrix} \times \begin{pmatrix} F_{Xib} \\ F_{Yib} \\ F_{Zib} \end{pmatrix} \right\} \tag{2}$$

Now, if we denote the force received at the motor's foot-support points from foot 1 to foot 4 as shown in Fig. 8, Eqs. (3) and (4) hold based on the balance of forces and balance of moments, respectively, on the motor.

$$\begin{pmatrix} F_X \\ F_Y \\ F_Z \end{pmatrix} = \begin{pmatrix} -f_{X1} - f_{X2} - f_{X3} - f_{X4} \\ -f_{Y1} - f_{Y2} - f_{Y3} - f_{Y4} \\ -f_{Z1} - f_{Z2} - f_{Z3} - f_{Z4} \end{pmatrix} \tag{3}$$

$$\begin{pmatrix} M_X \\ M_Y \\ M_Z \end{pmatrix} = \frac{L}{2} \begin{pmatrix} -f_{Z1} - f_{Z2} + f_{Z3} + f_{Z4} \\ f_{Z1} - f_{Z2} - f_{Z3} + f_{Z4} \\ (f_{X1} - f_{Y1}) + (f_{X2} + f_{Y2}) - (f_{X3} - f_{Y3}) - (f_{X4} + f_{Y4}) \end{pmatrix} \tag{4}$$

Here, L denotes the interval from one foot to the other.

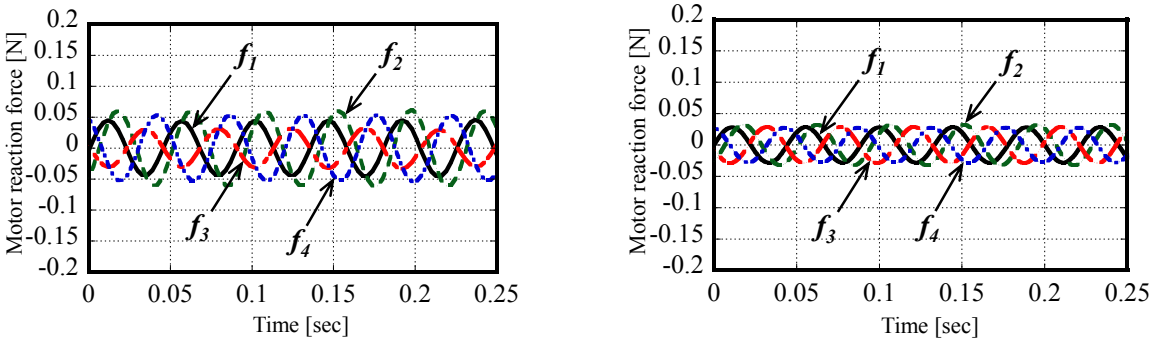
A fan-motor system in an outdoor air conditioning unit is generally attached to the thin-plate support structure. It therefore becomes necessary to apply an out-of-plane load when such a system is causing the thin-plate support structure to vibrate. In other words, F_Z and M_X, and M_Y in Eqs. (3) and (4) take on particular importance at this time.

Assuming that the stiffness of the motor is sufficiently high compared with the stiffness of the structure supporting the motor, and using the geometrical relationship of the four motor feet positioned on the same plane, motor reaction forces $f_{Z1} - f_{Z4}$ can be given by Eq. (5) using F_Z and M_X, and M_Y determined from Eqs. (3) and (4).

$$
\begin{aligned}
f_{z1} &= -\frac{F_z}{4} - \frac{M_X - M_Y}{2L} \\
f_{z2} &= -\frac{F_z}{4} - \frac{M_X + M_Y}{2L} \\
f_{z3} &= -\frac{F_z}{4} + \frac{M_X - M_Y}{2L} \\
f_{z4} &= -\frac{F_z}{4} + \frac{M_X + M_Y}{2L}
\end{aligned}
\tag{5}
$$

Based on the flow described above, we calculated the motor reaction forces. We found that the main frequency component of motor reaction forces was also the same BPF component (22.5 Hz) as that for aerodynamic pressure pulsation described in the previous section.

On the basis of this result, we decided to compare the results obtained by calculations with those by experiment for the BPF component of motor reaction forces for all four feet. The resulting temporal waveforms for the BPF component are shown in Fig. 9. It can be seen that calculation results reproduce well the experimental results in terms of magnitude and phase difference for the motor feet. This result demonstrates that motor reaction forces can be calculated with sufficient accuracy by the sequence of calculations presented in this report.



(a) Calculation (b) Experiment

Figure 9: BPF component of motor reaction forces

4 CONCLUSIONS

- In this paper, we investigated aerodynamic pressure pulsation and motor reaction forces as aerodynamic excitation forces generated by a rotating fan-motor system. We showed that pressure pulsation on a measurement board as determined by experiment could be represented well by the results of CFD analysis. We also

determined the pressure pulsation on the fan-blade surface by CFD analysis and used that result to also derive a formula for calculating the motor reaction forces. We found that the results of calculating motor reaction forces with that formula agreed well with experimental results, which demonstrated that motor reaction forces could be calculated with sufficient accuracy.

REFERENCES

[1] Ota et al., Direct measurement of aerodynamic excitation force generated by rotating-blade fan, Noise Control Eng. J. 57(4), July-Aug 2009, pp. 310-317, 2009.
[2] Sato et al., Aerodynamic excitation force generated by rotating-blade fan and its reaction force, MIPE2012, pp. 40-42, 2012.
[3] Furukawa, et al., Analysis of unsteady pressure field in a propeller fan using large eddy simulation, Proceedings of JSCFD, B07-2, 1999 (in Japanese).
[4] Watanabe et al., Prediction of aerodynamic noise of fans, Trans. Jpn. Soc. Mech. Eng., Vol. 66, No. 642, pp. 453-459, 2000 (in Japanese).

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis (Eds)

COMPARISON OF COUPLED EULER-LAGRANGE AND SMOOTHED PARTICLE HYDRODYNAMICS IN FLUID-STRUCTURE INTERACTION

CHRISTIAN ZEHETNER, MARKUS SCHÖRGENHUMER, FRANZ HAMMELMÜLLER AND ALEXANDER HUMER

Linz Center of Mechatronics GmbH
Altenbergerstraße 69
4040 Linz, Austria
e-mail: christian.zehetner@lcm.at, web page: www.lcm.at

Key words: Fluid-Structure Interaction, Smoothed Particle Hydrodynamics (SPH), Coupled Euler-Lagrange Finite Element Method (CEL)

Abstract. In this paper a fluid-structure interaction problem is investigated, in which fluid flow and flexible deformations of structures are coupled. Exemplarily, the collision of a moving deformable water-filled container with a rigid wall is considered. Two simulation methods are compared to analyze the impact: the Coupled Euler-Lagrange Finite Element Method (CEL) and Smoothed Particle Hydrodynamics Analysis (SPH). On the other hand, the solutions of two software packages are compared, the commercial Finite Element code Abaqus (CEL, SPH) and the open source package HOTINT/LIGGGHTS (SPH). Goal is to find the various advantages and disadvantages of the two simulation methods and the two software codes.

1 INTRODUCTION

In fluid-structure interaction problems, fluid flow and flexible deformations of structures are coupled. Examples are obstacles in fluid flow or the dynamics of fluid-filled structures. In this paper, two methods are compared for modelling and simulation of such processes.

First, the Coupled Euler Lagrange approach (CEL) is considered [1], where solid bodies and structures are modelled by Lagrangian finite elements in which the material is fixed to the element. On the other hand, the fluid is modelled by Eulerian finite elements which are fixed in space, and the material moves through the elements. The coupling is established by contact interactions, implemented as a penalty contact method.

In contrast, Smoothed Particle Hydrodynamics (SPH) [2] is a mesh-free Lagrangian method where the fluid is divided into discrete elements represented by particles. The interaction with solid structures, modelled by Lagrangian finite elements, is a crucial issue in today's SPH approaches. In the present work, penalty-based formulations are used to model the wall boundaries.

In the following computational study, a fluid-filled, flexible container colliding with an obstacle is considered. The resulting motion of the fluid, the deformation of the container, and the contact forces are investigated.

Three solutions are compared: (i) The multibody code HOTINT is coupled with the particle simulator LIGGGHTS which includes an SPH implementation [3]. Both packages are available as open-source software. The results of this co-simulation are compared with two formulations implemented in the commercial software Abaqus, i.e., the SPH formulation (ii) and the CEL approach (iii).

In section 2 the investigated problem is stated. In section 3 the simulation models are described in detail. Section 4 shows a comparison of the simulation results and a discussion of the advantages and disadvantages of the investigated simulation methods.

2 EXAMPLE PROBLEM

In the following we consider the example problem as shown in Figure 1. A water filled container is horizontally moving towards a rigid wall with constant speed v_0, which is defined as initial condition. Friction between container and floor is neglected, and the graviation g directed towards the floor is acting on both the container and the fluid. The impact due to the collision of the container with the wall causes an elastic deformation of the container and flow of the water. The deformation is assumed to be elastic, but large deformations are considered. Goal is to investigate the contact force between container and wall during impact, the deformation of the container, and the distribution and motion of the fluid after the impact.

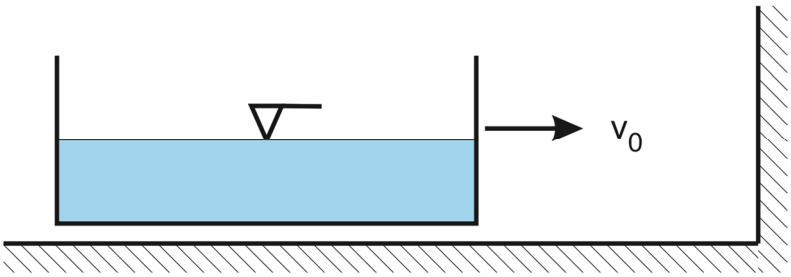


Figure 1: Water filled container colliding with a wall

Concerning the container, a linear elastic behavior is assumed, for which the material properties of aluminium in terms of Young's modulus E and Poisson's ratio ν are given below. The geometric dimensions are length L, width B, height H and wall thickness t. The container is filled with water up to a height of H/2. The properties of water are the dynamic viscosity η, speed of sound c_0 and density ρ. For the numerical investigations the parameters are chosen as

v_0 = 10e3 mm/s	L = 1e3 mm	η = 1e-9 Ns/mm^2	g = 9.81e3 mm/s^2
E = 70e3 N/mm²	B = 0.5e3 mm	c_0 = 1.45e5 mm/s	
ν = 0.3	H = 0.5e3 mm	ρ = 1e-6 kg/mm^3	
	t = 1 mm		

Note that c_0 is a factor of 10 smaller than the actual physical speed of sound in water, which leads to a higher compressibility of the simulated fluid. However, the density variations in the present example remain in the range of 1%. Therefore, the relevant behavior of water is still closely approximated, while the computational effort is significantly decreased.

In section 4, the simulation models are specified in more detail. The models in HOTINT and Abaqus are differing in several respects because of different implementation of the shell elements of the container, the contact formulations and the SPH formulations. In section 5, the solutions of the two software tools and the two simulation methods are compared. As mentioned above, the goal is to find the advantages and disadvantages of the simulation models and to verify the results.

3 SIMULATION MODELS

In this section three simulation models are compared. Care has been taken to develop coinciding models with respect to physical and numerical parameters. However there are some differences in the implementations in the two software tools Abaqus and HOTINT, as well as in the two simulation methods SPH and CEL in Abaqus. Especially, the contact algorithms are differing in the three investigated formulations. In the following sections the implemented models are described in detail.

3.1 HOTINT/LIGGGHTS – SPH

HOTINT is an open source (http://www.hotint.org), multi-purpose simulation software for both research and industrial applications [4]. The core functionality of HOTINT focuses on a detailed and efficient analysis of complex flexible multibody systems. In addition to conventional features of multibody codes, i.e., rigid bodies, simple flexible members and a large variety of kinematic pairs, HOTINT offers several elements for large deformation problems. The latter range from advanced beam, plate and shell elements to classical non-linear solid finite elements. In the present investigations, a four-noded large deformation plate element is utilized for the representation of the solid container. The plate element [5] employs Kirchhoff's kinematic assumption of shear deformation being negligible for thin structures. Concerning the numerical formulation, the element belongs to the family of elements based on the absolute nodal coordinate formulation (ANCF) [6]. The key idea of ANCF is to avoid rotational degrees of freedom in favour of slope vectors, i.e., the two gradient vectors of the mid-surface's position in the present case. Each node consequently has nine degrees of freedom requiring third-order polynomials for the interpolation. Due to the high order of convergence, a relatively small number of elements are required as compared to the linear elements used in Abaqus. Each side of the container is discretized separately first; afterwards, the sidewalls are stitched together by means of kinematic constraints that prohibit the relative translation and rotation at the nodal positions on the edges. A rigid body fixed to the ground represents the obstacle. For the contact between the container and that obstacle, a penalty formulation is utilized in the simulations.

Regarding time integration, a range of implicit high-order Runge-Kutta methods are available in HOTINT. For the present investigations, a third-order Radau IIA scheme with a constant step size is utilized.

The fluid part is not directly included in HOTINT; instead, a co-simulation with the open-source software LIGGGHTS (http://www.liggghts.com) is performed for this purpose. LIGGGHTS is a massively parallel solver originally developed for large-scale particle simulations in the field of molecular dynamics, on top of which the SPH functionality and a fluid-structure interface have been implemented [3].

As opposed to the solid structure, the explicit time integration of the SPH formulation typically requires much smaller time-steps. Accounting for that, the two simulators HOTINT and LIGGGHTS are coupled in the sense of a weakly-coupled force-displacement scheme of Jacobi type (see, e.g., [7]). Performing the time-integration, several explicit time-steps are computed for the fluid during a single implicit step on the solid side. After each implicit step, the updated forces of the fluid on the structure and the position and velocity data of the structure's surface (which forms the moving boundary for the fluid) are exchanged mutually by means of a TCP/IP interface. During the fluid substeps, the velocity field of the boundary is assumed constant, and the positions are obtained by explicit integration. Further details on the underlying penalty-based boundary formulation for SPH, the coupling scheme, and the implementation can be found in [3]; for advanced features and applications, the reader is referred to [8].

3.2 Abaqus – SPH and CEL

The container is modelled by linear shell elements of type S4R, the floor and wall by rigid elements of type R3D3. In the SPH-model, the fluid is initially modelled by 3D continuum elements C3D8R and automatically converted to particles by the solver in the first simulation step. This conversion procedure enables a more comfortable way for defining the particles. On the other hand, in the CEL-model the fluid is represented by Eulerian elements of type EC3D8R. For modelling the interactions of the parts, a general contact formulation with soft behavior in normal direction is defined. Thus, contact is considered between fluid and container, as well as container and wall. Due to the differing implementations of SPH and Eulerian elements, there are slight differences in the contact model. Explicit time integration is performed considering large deformations. Further information about the implemented elements and algorithms can be found in Abaqus Manual [9].

4 SIMULATION RESULTS

4.1 Displacements in the Container without Fluid

In the first step, the displacements of the container in the static case are investigated. The container is clamped on the back wall. On the front wall a pressure of 4e-4 N/mm² is applied. The results of a geometrically nonlinear analysis for the magnitude of the displacements are shown in Figure 2. The maximum displacements in longitudinal (x) and transversal (z) direction obtained by HOTINT and Abaqus are compared in Table 1.

Table 1: Maximum longitudinal and transversal displacements in a static simulation

Model	u_x (mm)	u_z (mm)
HOTINT (8x4x4)	20.51	11.39
HOTINT (16x8x8)	21.15	11.93
Abaqus	21.73	11.38

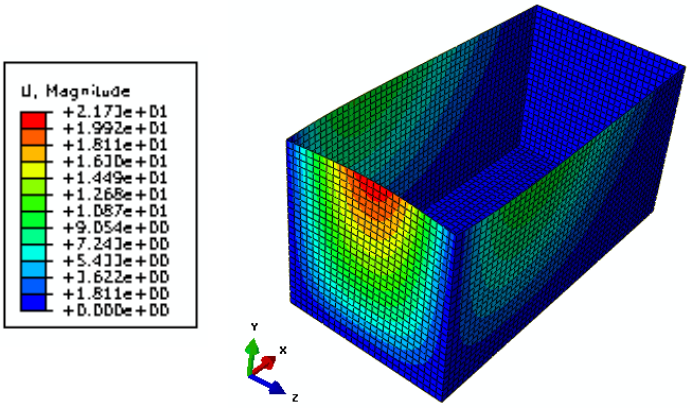


Figure 2: Magnitude of displacements in a static simulation

In the next step, a modal analysis has been performed to obtain the eigenfrequencies of the container. The results obtained by HOTINT (three mesh sizes) and Abaqus (one mesh size) are compared in Table 2, Figure 3 shows the first two symmetrical modes.

Table 2: Eigenfrequencies

	Frequency (Hz)			
Mode Nr.	HOTINT (4x2x2)	HOTINT (8x4x4)	HOTINT (16x8x8)	Abaqus
1	3.370	3.378	3.380	3.373
2	5.388	5.444	5.459	5.454
3	7.680	7.698	7.706	7.696
4	12.233	12.346	12.364	12.356
5	16.075	15.890	15.963	15.962
6	16.599	16.156	16.133	16.129

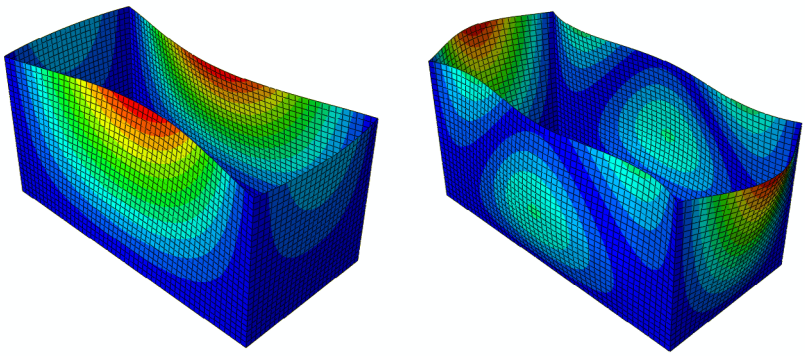


Figure 3: Symmetrical modes (a) mode nr. 2 and (b) mode nr. 5

In general these results show a good coincidence for the two simulation codes. The slight differences are caused by the different element types in the HOTINT and Abaqus model.

4.2 Motion and Pressure of the Fluid

Figures 4 and 5 show the situation 50 ms after the collision of the container and the wall. The figures show the deformed container and the motion of the fluid, by a plane cut through the scene. The contour plot refers to the pressure in the fluid. Good mutual agreement was obtained between the three results for both the displacement and the pressure field in the fluid. The Abaqus SPH-solution for the pressure exhibits much more noise than the SPH-solution of HOTINT/LIGGGHTS.

However, the deformation field in the three solutions is differing significantly at the same time. The mechanical behaviour of the container has been investigated in section 4.1, showing a very good coincidence of the static results and the modal analysis. Therefore, we can assume that the discrepancies of the displacement field of the container in Figures 4 and 5 are caused by the different contact formulations in the simulation models. The specific model of the impact has significant influence on the amplitude and phase of the excited modes of the container.

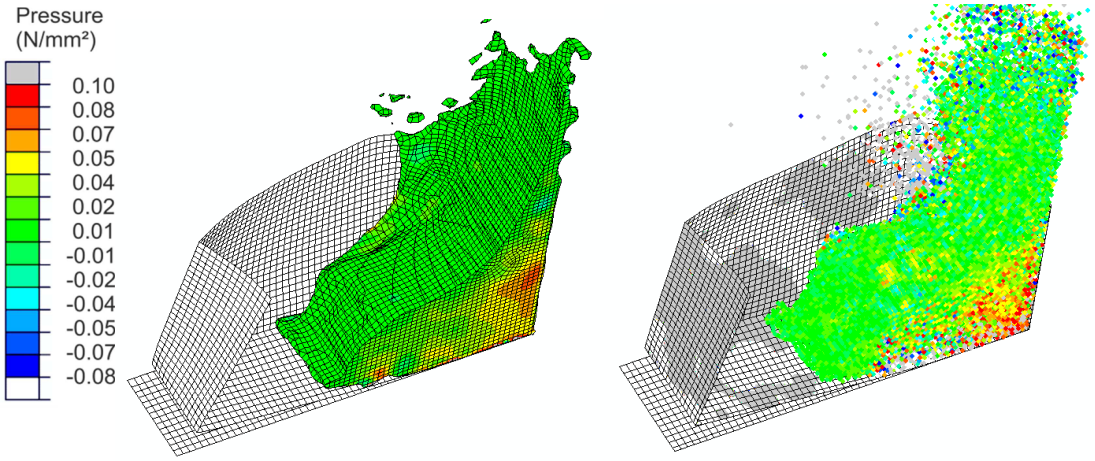


Figure 4: Pressure in the Fluid 50 ms after impact, (a) Abaqus CEL, (b) Abaqus SPH

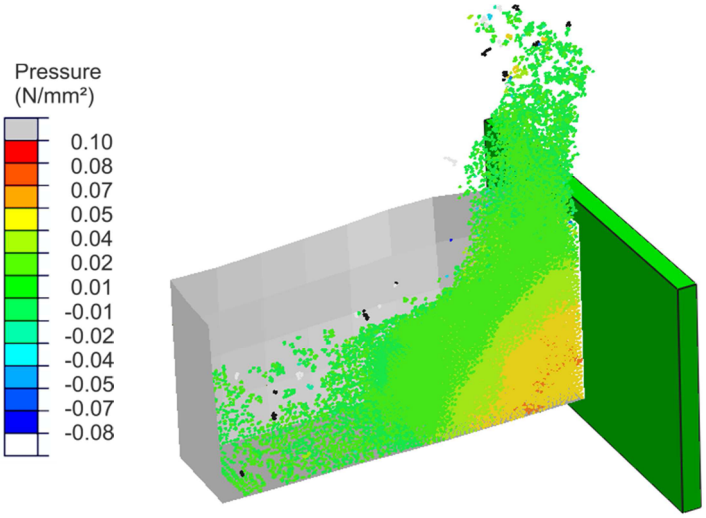


Figure 5: Pressure in the Fluid 50 ms after impact, HOTINT SPH

4.3 Stress in the container

In the next step, the distribution of the von Mises stress in the container is compared. Corresponding to the results in section 4.1, the von Mises stress field is compared for points in time with coinciding deformation of the container. Thus, Figure 6 shows the Abaqus results 38 ms after the impact and Figure 7 the HOTINT results 28 ms after impact. Comparison of the three simulation models shows a good coincidence for the maximum von Mises stress in the container of about 400 N/mm².

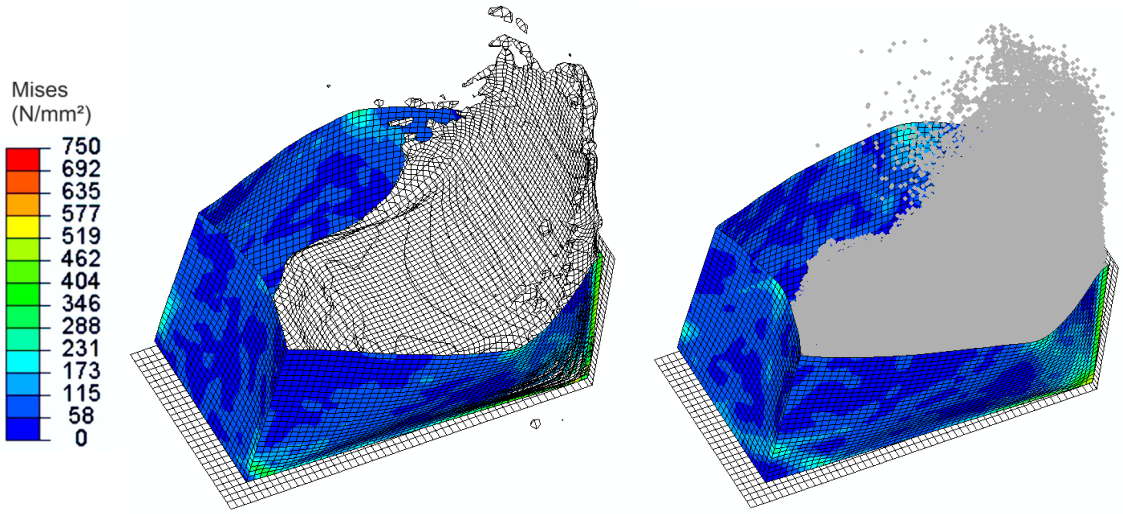


Figure 6: von Mises Stress in the container 38 ms after impact (a) Abaqus CEL (b) Abaqus SPH

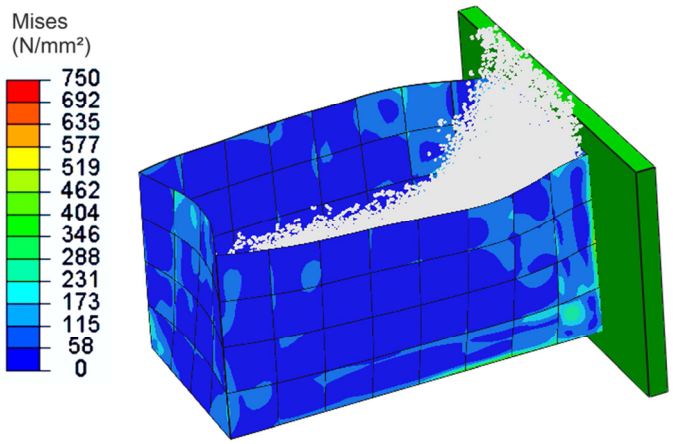


Figure 7: von Mises Stress in the Container 28 ms after impact, HOTINT

4.4 Contact Force

Finally, the resulting contact force between container and wall is compared for the three simulation models in Figure 8, starting 2 ms after the impact. Due to differing contact formulations it is not reasonable to compare the contact force during the impact. This is also the reason for the phase shift a short time after the impact, and thus for the differences in the displacement field in section 4.2. For t>40 ms after impact the contact force shows a very good coincidence.

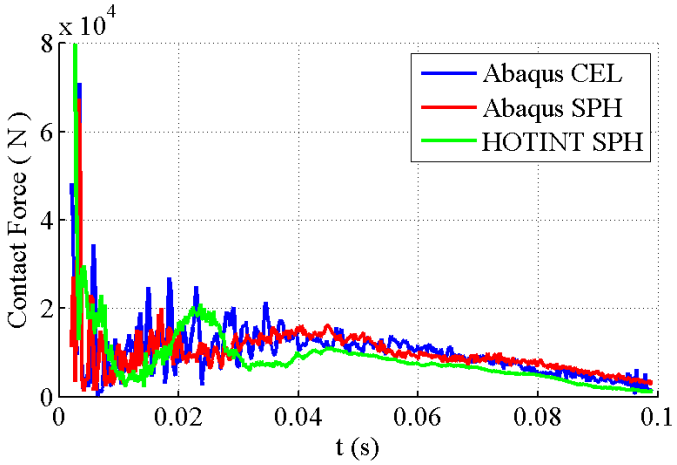


Figure 8: Contact force 2 ms after impact

4.4 Notes on the Computational Efficiency

All three coupled fluid-structure simulations were performed with the same spatial resolution for the fluid domain (in the range of 80000 particles or elements, respectively), and comparable accuracy for the solid parts. As to the latter, the spatial resolution for the plate elements of the container is, as already mention, significantly lower in case of the high-order elements in HOTINT (cf. section 3.1). The different resolutions can be compared Figure 6 and Figure 7.

Regarding time integration, Abaqus exclusively uses an explicit integrator, while the HOTINT/LIGGGHTS approach is based on a coupling between high-order implicit and fast explicit integration (see again section 3.1).

As a reference point, on 4 desktop CPUs, the average wall clock times for the whole impact simulation (i.e., 100 ms simulated physical time) are summarized in Table 3. We observed significantly better performance for the HOTINT/LIGGGHTS approach compared to the Abaqus simulations. Mainly, this is due to the highly efficient SPH implementation within LIGGGHTS. Note that the difference increases with increasing spatial resolution for the fluid domain.

Table 3: Wall clock times T for the three simulations

Simulation	T
Abaqus CEL	~ 3h 30min
Abaqus SPH	~1h 45min
HOTINT/LIGGGHTS – SPH	~40 min

5 CONCLUSIONS

In this paper two simulation methods and tools for modelling fluid-structure interaction have been compared. The results have shown a good coincidence for the displacement and pressure field of the fluid, the maximum von Mises stress in the container and the time response for the contact force. Due to differing contact formulations in the three models, it is not reasonable to compare the results during the impact. These differences also cause a phase shift in the results of the contact force and displacment field of the container. However, comparing the maximum displacements and von Mises stresses in the container for corresponding deformation shapes yields satisfactory agreement.

The presented results show that all three methods are appropriate for simulating fluid-structure interaction. Appropriate results can be expected for the motion of the fluid, the pressure field in the fluid, the contact force between container and wall, as well as the maximum displacements and stresses in the structure.

REFERENCES

[1] Benson, D.J. Computational Methods in Lagrangian and Eulerian Hydrocodes, *Comput. Methods Appl. Mech. Engrg.* (1992) **99**: 235–394

[2] Monaghan J.J. Smoothed particle hydrodynamics, *Rep. Prog. Phys.* (2005) **68**: 1703–1759

[3] Schörgenhumer, M., Gruber, P. and Gerstmayr J. Interaction of flexible multibody systems with fluids analyzed by means of smoothed particle hydrodynamics, *Multibody Syst. Dyn.* (2013) **30**: 53–76.

[4] Gerstmayr, J. Dorninger, A., Eder, R., Gruber, P., Reischl, D., Saxinger, M., Humer, A., Nachbagauer, K., Pechstein, A.S. and Vetyukov, Y. HOTINT – a script language based framework for the simulation of multibody dynamics systems. *In Proc. ASME 2013 IDETC/CIE*, Portland (OR), USA, 2013.

[5] Dmitrochenko, O.N., and Pogorelov, D.Yu. Generalization of plate finite elements for absolute nodal coordinate formulation, *Multibody Syst. Dyn*. (2003) **10**: 17–43.

[6] Gerstmayr, J., Sugiyama, H., Mikkola, A. Review on the Absolute Nodal Coordinate Formulation for Large Deformation Analysis of Multibody Systems. *J. Comput. Nonlinear. Dynam.* (2013) **8** (3): 031016 (12 pages).

[7] Busch, M. Zur Effizienten Kopplung von Simulationsprogrammen, PhD Thesis, University Kassel, Germany, 2012

[8] Schörgenhumer, M., Humer, A., and Gerstmayr, J. Efficient fluid-structure interaction based on modally reduced multibody systems and smoothed particle hydrodynamics. *In Proc. WCCM IX*, Barcelona, Spain, 2014.

[9] Abaqus 6.12 Online Documentation, Dassault Systèmes, 2012

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis(Eds)

EFFICIENT FSI CODES COUPLING WITH POSSIBLE LARGE ADDED MASS EFFECTS: APPLICATIONS TO RIGID AND ELONGATED FLEXIBLE BODIES IN THE MARITIME FIELD.

A. LEROYER†, C. YVIN*, E. GUILMINEAU†, M. VISONNEAU† AND P. QUEUTEY†

† LHEEA, UMR-CNRS 6598, Ecole Centrale Nantes
1 rue de la Noë, 44321 Nantes Cedex 3, France
e-mail: alban.leroyer@ec-nantes.fr

* DCNS Research,
1 rue de la Noë, 44321 Nantes Cedex 3, France
e-mail: camille.yvin@sirehna.com

Key words: Fluid-Structure Interaction, added-mass, hydrodynamics, codes coupling

Abstract. Co-simulation, which involves codes coupling, is the most popular technique in an industrial context to deal with multi-physics applications. This is mainly due to its modular nature and the use of specialized solvers which have the ability to integrate the most advanced numerical techniques and physical models in each scientific field. However, in many configurations, the development of coupling algorithms, easy to implement, leading to a stable, accurate and efficient tool is generally not straightforward. For Fluid-Structure Interaction (FSI) configuration involving hydrodynamics, it is well-known that added-mass effect tends to distabilize classical coupling algorithms, such as the Block-Gauss-Seidel algorithm (often denoted by Dirichlet-Neumann decomposition too). Here, some modifications of this algorithm are proposed to reach a weak-intrusive stable coupling method for rigid and elongated beam-like bodies. Efficiency is discussed and some applications are shown to demonstrate the capabilities of such a coupling.

1 INTRODUCTION

Fluid Structure Interaction (FSI) problems are commonly encountered in naval architecture. Even if this can be done through a monolithic approach within a single solver, a partitioned approach is most commonly used through codes coupling, especially well fitted to the resolution of complex FSI problems. Indeed, with this technique, complex models for both the fluid and the structure can be used because each solver can be numerically adapted and dedicated to its own physics. In this work, two different general solvers are used: ISIS-CFD and MBDyn, for the fluid and the structural part, respectively.

ISIS-CFD is a Navier-Stokes solver developed in the LHEEA Laboratory at Ecole Centrale de Nantes. MBDyn is an open-source solver intended to solve multi-disciplinary problems including non-linear dynamics of rigid and flexible bodies subjected to kinematic constraints, along with active controls. The combination of these two solvers makes possible the study of a broad spectrum of applications in the marine field which cannot be solved with a unique solver. After describing how the code coupling is handled and how the added-mass effects are tackled to reach a robust and efficient algorithm, some applications among those studied are shown. Especially, a test-cases of a ship with active control of appendages (roll damping) is described, where good agreement with the experimental data was obtained. Other applications with flexible slender bodies are also presented to demonstrate the capabilities of such a coupling.

2 DESCRIPTION OF THE TWO SOLVERS

2.1 The fluid part: ISIS-CFD

ISIS-CFD is available as a part of the FINE^{TM}/Marine computing suite which is dedicated to marine applications. This is an incompressible unsteady Reynolds-averaged Navier-Stokes (RANS) solver developed by the DSPM group of the *LHEEA Lab. of Ecole Centrale Nantes, UMR-CNRS 6598* . This solver is based on a fully unstructured finite-volume method to build the spatial discretisation of the conservation equations. Pressure-velocity coupling is obtained through a Rhie & Chow SIMPLE-type method : in each time step, the velocity update comes from the momentum equations and the pressure is given by the mass conservation, transformed into a pressure equation. An Arbitrary Lagrangian Eulerian (ALE) formulation is used to take into account modification of the fluid spatial domain [1]. It is associated with robust and fast grid deformation techniques [2, 3]. The temporal discretisation scheme is the Backward Difference Formula of order 2 (BDF2) when dealing with unsteady configurations. For each time step, an inner loop (denoted by non-linear loop) associated to a Picard linearisation is used to solve the non-linearities of the system and to converge all the sequential coupled equations.

Free-surface flow is addressed with an interface capturing method, by solving a convection equation for the volume fraction of water, which is discretised with specific compressive discretisation schemes [4]. The code is fully parallel using the MPI (Message Passing Interface) protocol. An automatic adaptive grid refinement technique [5] and a sliding grid method are also included .

2.2 The structure part: MBDyn

MBDyn (Multi-Body Dynamics), is an open-source solver under the GNU GPL license developed at the *Dipartimento di Ingegneria Aerospaziale* of the *Politecnico di Milano*. It is aimed at the modelling of complex multi-bodies systems and muti-discipline problems including non-linear dynamics, aero-servo-elasticity, smart piezo-structural components and electric and hydraulic components [6] To solve the kinematic laws of a multi-body

mechanical system, the Redundant Coordinate Set (RCS) formulation is used. This means that every inertial body has six rigid body Degrees of Freedom (DOF) even if they are constrained by joints for instance. Additional holonomic or nonholonomic constraint equations are added which introduce algebraic unknowns that are analogous to the Lagrange multipliers and directly represent the reaction forces and couples [7]. All these equations are written in the form of a set of first order Algebraic Differential Equations (ADE). A direct resolution of these multi-body ADE is made possible through a combination of a variable substitution of the algebraic unknowns, the scale of constraint equations by the time step and a A-stable multi-step time integration scheme to damp numerical oscillations. In this work, the BDF2 scheme, similar to the fluid solver, is used.

The ability to take into account multidisciplinary complex systems and the simplicity of implementation are the main advantages of this formulation.

3 CODES COUPLING AND FSI

3.1 General algorithm

FSI problems can simply be expressed as two continuum materials (structural and fluid, denoted by an index s and f, respectively) sharing a common interface $\Gamma_s = \Gamma_f = \Gamma$, where the two following conditions operate:

$$\text{kinematic condition:} \quad \delta_s = \delta_f \text{ on } \Gamma\,, \quad (\delta \text{ refers to the position}) \tag{1}$$

$$\text{dynamic condition:} \quad \sigma_s + \sigma_f = 0 \text{ on } \Gamma\,, \quad (\sigma \text{ refers to the stress vector}) \tag{2}$$

Using a domain decomposition point of view, let's introduce the Steklov-Poincaré operator $\mathcal{S}$ and its inverse $\mathcal{S}^{-1}$, which can be seen as the transfer function at the interface of the fluid and structure solver (subscript d represents the considered domain):

$$\mathcal{S}_d(\delta_d) = \sigma_d \qquad \mathcal{S}_d^{-1}(\sigma_d) = \delta_d \quad , \tag{3}$$

The conditions (1) and (2) can be expressed with the operators previously defined in different formulations which provide different resolution strategies ([8]). In particular, the classical fixed-point algorithm, leading to an implicit Block Gauss-Seidel (BGS) approach (also called Dirichlet-Neumann algorithm) can be simply represented with the following expression:

$$\sigma_f = -\mathcal{S}_f \circ \mathcal{S}_s^{-1}\left(\sigma_f\right) \tag{4}$$

Equation (4) is a non-linear time dependant equation and has to be fully solved at each time step. Explicit coupling schemes in time, which perform only one iteration to solve equation (4) using predictor and corrector, can produce or dissipate energy at the interface and leads to unphysical results if the problem is highly coupled. Moreover, they are very sensitive to the destabilizing added-mass effect [9]. Hence, in this work, only implicit time integration is considered. It means that equation (4) needs to be solved through an iterative procedure, indexed by i. The simplest implicit BGS algorithm to solve time iteration $n+1$ can be represented by equation (5):

$$\sigma_f|_{n+1}^{i+1} = -\mathcal{S}_f \circ \underbrace{\mathcal{S}_s^{-1}\left(\sigma_f|_{n+1}^{i}\right)}_{\delta|_{n+1}^{i+1}} \tag{5}$$

At each coupling iteration, the structure problem is solved using the current fluid loads $\sigma_f|_{n+1}^{i}$ at time step $n+1$ coming from the last iteration. This provides a new position $\delta|_{n+1}^{i+1}$, which is used as an input data for the fluid solver to compute new fluid variables at the same time step $n+1$, leading to an updated fluid stress field at the interface $\sigma_f|_{n+1}^{i+1}$.

This algorithm is quite easy to implement since it only operates at the temporal loop level, but needs a lot of coupling iterations before reaching convergence. In order to reduce the simulation time, Aitken Δ^2 relaxation technique is often used [10]. This technique is based on a geometrical approximation (tangent method) of the coupled problem. To be efficient, the solvers have to produce physical results, and consequently, each operator has to be solved accurately. This means that for an implicit coupling algorithm, the fluid problem has to be solved several times at each time step which is not acceptable in term of CPU time. In this work, the classical Steklov-Poincaré operator for the fluid problem is then modified to reduce the CPU time of this part. This new operator, denoted as $\mathcal{S}_f^*$ does not represent a global fluid resolution any more but only one Picard iteration of the fluid solver as already considered in [9]. It can be seen as a fluid linearised version of the Steklov-Poincaré operator. In a few words, the fluid problem converges at the same time as the coupled problem. In term of implementation, it means integrating the resolution of the structure solver inside the non-linear iteration of the fluid which becomes mixed up with the FSI coupling loop. To be efficient, the CPU time related to the structural solver has to be weak compared with that requested for the fluid part. A second requirement is to have available fast grid deformation techniques since such procedure is now called at each non-linear iteration to update the fluid mesh with the new boundary nodal positions related to the solved structures. The final algorithm (including a relaxation operator which is going to be introduced in section 3.2) is displayed on figure 1.

3.2 Added-mass effect and stabilization

When dealing with large fluid density or light structure, added-mass effects become an important physical phenomenum for the coupled problem. At the discrete level, this inertia effect which leads to fluid loads dependant on the acceleration of the structure makes all the code coupling algorithms previously described unstable. This numerical instability has been clearly highlighted in [11, 12]. When implicit algorithm is used, a popular technique to tackle this diverging behaviour is to modify the structure equation to decrease the acceleration dependancy of the source term. This modification can be interpreted as an approximated (and then iterative) resolution of a block-LU factorization of the monolithic system. Let's represent the formal monolithic system as:

$$\begin{bmatrix} \mathbf{F} & \mathbf{C}_{SF} \\ \mathbf{C}_{FS} & \mathbf{S} \end{bmatrix} \begin{bmatrix} x_F \\ x_S \end{bmatrix} = \begin{bmatrix} s_F \\ s_S \end{bmatrix} \tag{6}$$

x_F and x_S refer to the fluid and structure variables, respectively. s_F and s_S include the source term of the fluid and structure operators ($\mathbf{F}$ and $\mathbf{S}$, respectively). $\mathbf{C}_{SF}$ and $\mathbf{C}_{FS}$ define the data transfer operator at the interface between the two domains. A block-LU factorization leads to the following system:

$$(\mathbf{S} - \mathbf{C}_{FS}\mathbf{F}^{-1}\mathbf{C}_{SF})x_S = s_S - \mathbf{C}_{FS}\mathbf{F}^{-1}s_F \tag{7}$$

$$\mathbf{F}x_F = s_F - \mathbf{C}_{SF}x_S \tag{8}$$

This resolution exhibits operator composed with F^{-1}, which is not acceptable to build in the context of co-simulation. However, by reducing the Jacobian of $\mathbf{C}_{FS}\mathbf{F}^{-1}\mathbf{C}_{SF}$ only to its inertial effect (i.e. by the added mass operator -$\mathbf{M_a}$, approximation all the more valid as the time step is small [13, 14, 3]), it is possible to transform the direct resolution of (7) in an iterative procedure without any operator composed with F^{-1}. It comes indeed:

$$(\mathbf{S} + \mathbf{M_a})\, x_S^{k+1} = s_S^k - \mathbf{C}_{FS}x_F^k + \mathbf{M_a}\, x_S^k \tag{9}$$

$$\mathbf{F}x_F^{k+1} = s_F^k - \mathbf{C}_{SF}x_S^{k+1} \tag{10}$$

Whereas equation (10) refers to a classical resolution of the fluid solver, equation (9) is a modified version of the structure solver in which the right-hand side term is far less dependent on acceleration. In a code coupling context, this procedure is not suitable since it requires to modify the structural solver. To avoid this, it can be shown that the equation (9) is almost equivalent to a classical resolution of the structure solver added to a second step of relaxation with an operator $\mathbf{R}$, depending on the mass matrix of the structure and the added-mass operator (or an approximation of it). In the context of code coupling, the two approaches slightly differ. As a matter of fact, the evaluation of s_S and the non-linearities due to the rotation motion are not solved in the same manner, since for the second case, the relaxation occurs after the global resolution of the structure. The final algorithm is then simply described by equation (11) and is illustrated with a graph in figure 1.

$$\sigma_f|_{n+1}^{i+1} = -\mathcal{S}_f^* \circ \mathbf{R} \circ \mathcal{S}_s^{-1}\left(\sigma_f|_{n+1}^{i}\right) \tag{11}$$

For a body with one degree of freedom (DOF), it can be shown that the relaxation value for the acceleration is equal to $1/(1 + \frac{\tilde{m_a}}{m})$, where $\tilde{m_a}$ is an approximation of the added mass. In case of a six DOF rigid body, the interface displacement can be replaced by the generalized position (position and orientation) of the body denoted as δ and the added mass effect can be represented by a symmetric matrix of rank six denoted as $\mathbf{M}_a$. The corresponding approximated added mass matrix is denoted as $\widetilde{\mathbf{MA}}$. Thus, the acceleration relaxation step is defined by equation (12). $\tilde{\ddot{\delta}}$ is the direct result coming from the unmodified structure solver and $\mathbf{R}_a$ is the acceleration relaxation operator.

$$\ddot{\delta}\Big|_{n+1}^{i+1} = \ddot{\delta}\Big|_{n+1}^{i} + \mathbf{R}_a\left(\tilde{\ddot{\delta}}\Big|_{n+1}^{i+1} - \ddot{\delta}\Big|_{n+1}^{i}\right) \quad \text{, with} \quad \mathbf{R}_a = \left(\mathbf{I}_d + \mathbf{M}^{-1}\widetilde{\mathbf{MA}}\right)^{-1} \tag{12}$$

After the acceleration relaxation step, velocities and generalized positions are reconstructed according to the time integration scheme used. When using the implicit algorithm described in figure 1, the operator $\widetilde{\mathbf{MA}}$ does not need to accurately represent the physical added-mass operator $\mathbf{MA}$ to reach the converged solution. It is also worth noting that this converged solution does not depend on $\widetilde{\mathbf{MA}}$. However, it was shown in [15] that the number of iterations to reach a given gain is noticeably smaller when this operator is close to the physical added-mass operator. In this case, it was found that the number of iteration is even similar to an unsteady configuration without coupled FSI (for example, a simulation with an imposed body motion), which can be seen as an optimal efficiency. Tests have been carried out not only with classical hydrodynamics cases but also with configurations with extreme added mass effects. Even if the relaxation operator can lead to very small value, it does not compromise the efficiency of the coupling since this low relaxation has a physical origin. Let's now describe how the operator $\widetilde{\mathbf{MA}}$ is computed for a rigid body and how it is generalized for elongated beam bodies.

3.3 Computation and approximation of the added-mass operator

This physical added-mass operator, which is based on the instantaneous response to acceleration, generalizes the infinite-frequency added-mass operator deduced from classical linear potential flow approaches. But to avoid any recourse to such an external tool, we developed inside the CFD solver an integrated computation of the added-mass operator, which contributes to the originality of the present work. It is based on the resolution of the pressure field due to a brutal variation of the rigid body velocity (i.e. an acceleration step). Due to the different time scales of the Navier-Stokes equation, this perturbation denoted $\tilde{\mathbf{u}}$ is guided by the following equation, where $\tilde{p}$ refers to the additional pressure field due to this perturbation [11, 15].

$$\frac{\partial \tilde{\mathbf{u}}}{\partial t} == -\frac{1}{\rho}\boldsymbol{\nabla}\tilde{p} \tag{13}$$

By taking the divergence of equation (13) and using the Green-Ostrogradski theorem, it comes the following finite-volume form since $\tilde{\mathbf{u}}$ is divergence free:

$$\oiint_{\mathcal{S}} -\frac{1}{\rho}\boldsymbol{\nabla}\tilde{p}\cdot\mathbf{n}\,dS = 0 \tag{14}$$

This equation has to be solved with the Neumann boundary conditions for $\tilde{p}$ with a unit acceleration $\boldsymbol{\gamma}_b$ imposed at the surface of the considered body by a given DOF (one equation to solve for each one) and with zero for the possible other bodies or walls.

$$-\frac{1}{\rho}\boldsymbol{\nabla}\tilde{p}\cdot\mathbf{n} = \frac{\partial \tilde{\mathbf{u}}_b}{\partial t}\cdot\mathbf{n} = \boldsymbol{\gamma}_b\cdot\mathbf{n} \;\text{ at the wall of the considered body} \tag{15}$$

Equation (14) is finally similar to the pressure equation used to solve the real flow. The free surface shape is naturally taken into account by the spatial variation of the density ρ

when a multi- phase flow is considered. No additional equation has to be solved to take into account the free-surface position. This is not the case when a classical potential flow solver is used because the kinematic and the dynamic boundary conditions at the free surface have also to be respected. In addition to the easy-to-use feature due to its integration within the CFD solver (no need to generate another mesh, possible update during the computation), this method does not suffer from any limitations of linear potential solvers. For example, contrary to the latter, fine meshes can be used so all the details of the geometry can be entirely respected, and a larger amount of space configurations, like complex free surface position (wave breaking) or important confinement (shallow water, interactions with close bodies) is naturally taken into account and can be easily updated during the computation, if the conditions are strongly modified.

3.3.1 Case of rigid body

The added-mass operator is here restricted to a 6×6 matrix, which can be computed as follows: after resolution of $\tilde{p}$ for a given DOF index i, the integration of $\tilde{p}$ on the body surface provides the six coefficients of the column i of the added-mass matrix from the three forces and three moments. Some validations and comparisons with results coming from potential solver are available in [15].

3.3.2 Case of flexible elongated body

Assuming the structure of such a body is described with a beam structure, the number of DOF may be quite large compared to a rigid body, since each node of a classical Euler-Bernoulli beam gets at least three DOF of interest for the added-mass effects (the two transversal translation and the rotation of the section along the neutral line). As a consequence, the computation of each pressure field for each DOF becomes too CPU time consuming. Taking advantage of the elongated property of the body, the different added-mass coefficients associated to a given section (and then a given node) can be deduced considering the 2D potential flow in a plane normal to the neutral line as soon as the curvature effects are small ([16]). This reasonable approximation, resulting of the so-called elongated body theory is well suited for a body described with analytical definitions. However, for the sake of simplicity, this approach, which requires to extract 2D sections and to build a 2D mesh around them from a possible complex 3D body shape was not retained in this way in the context of a 3D CFD solver. Alternatively, a slightly similar approximation was done to avoid computing this series of 2D computations section by section : it consists here in the application of a unit acceleration corresponding to the same DOF with respect to the neutral line on all nodes of the beam simultaneously. The pressure field obtained within a single 3D computation is then concatenated towards the different nodes of the beam with the same loads transfer used during the FSI computation. The resulting nodal forces and moments which corresponds to the loads on the section surrounding the beam node, give access to the added-mass coefficients for each node. This

procedure can be applied for the three main DOF of interest and for the longitudinal translation along the neutral line too if it is needed. As a result, only a maximum of four resolution of Poisson operator is required to deduce a sufficient approximation of the added-mass operator for the entire DOF of each flexible beam. It can be noticed that this approach removes the possible coupling between the different nodes, by assuming the acceleration between the nearest neighbours is similar. This greatly simplifies the computation of the relaxation operator (12), since $\mathbf{I}_d + \mathbf{M}^{-1}\widetilde{\mathbf{MA}}$ remains block diagonal.

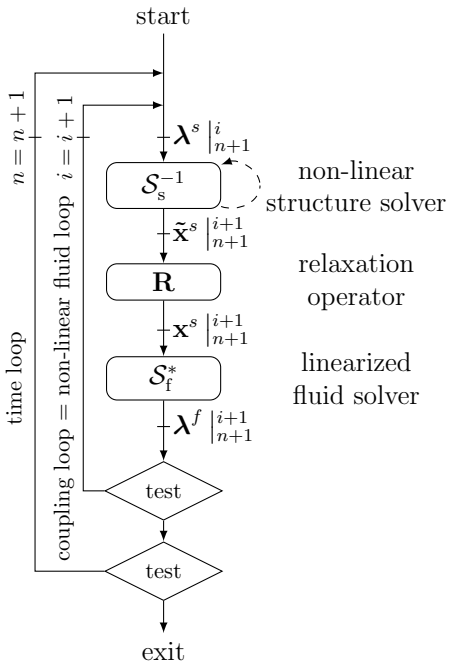


Figure 1: coupling algorithm

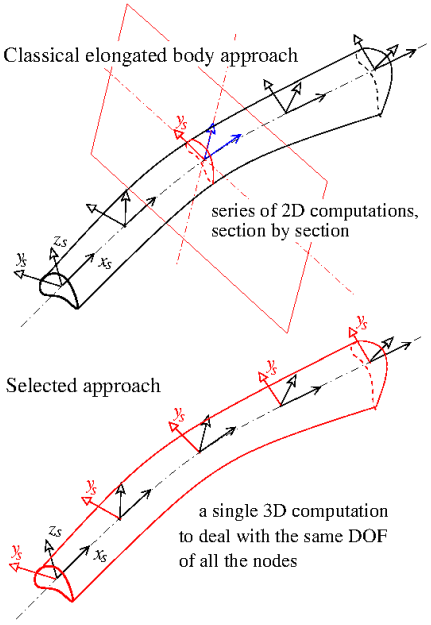


Figure 2: computation of the added-mass operator for an elongated body

4 APPLICATIONS

4.1 Roll damping with active control of appendages

This application is dedicated to the study of the roll decay of a frigate with active control of fins (see figure 3). Passive fins are also studied as a reference point. The numerical results are compared to experimental tests. The roll motion of the ship and the relative rotations of the fins are taken into account through a sliding grid technique. A combination of rigid motion and deformation algorithms are also used to take over the other degrees of freedom. The mesh is made up of 31 millions cells. The scenario is divided into three parts. First, the ship velocity is imposed (others degrees of freedom are released) until a steady state is obtained. Then, the roll angle is imposed to perturb the ship in a close way compared with the experimental procedure. In figures 4 and 5, this ramp is applied between t_2 and t_3. Finally, the roll motion is released and its decay is analysed. Comparisons between experimental tests and simulations are shown in figures

4, 5 and 6. One can see that both global quantities (roll angle) and local effects (drag and lift forces on fins) are well captured in spite of the difficulty to simulate this kind of complex experiments. It also can be seen in figure 4 that the roll damping is much more important when the fins are active than when they are passive.

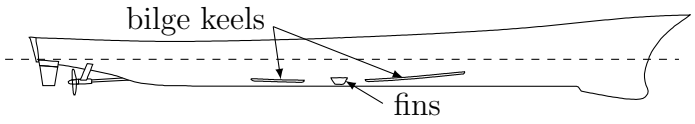


Figure 3: DTMB with active fins – Geometry

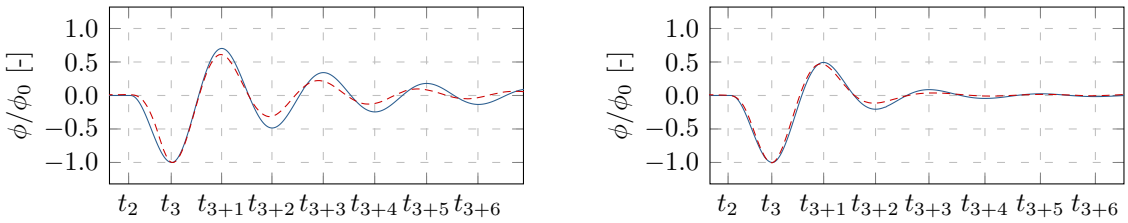


Figure 4: Roll angle – Passive fins (left) and active fins (right) : ---- experimental tests, —— present work

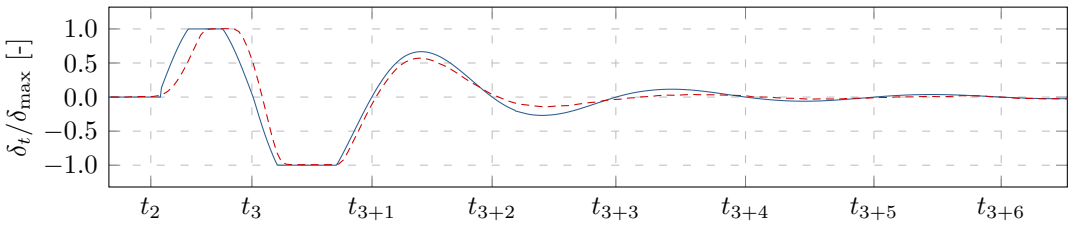
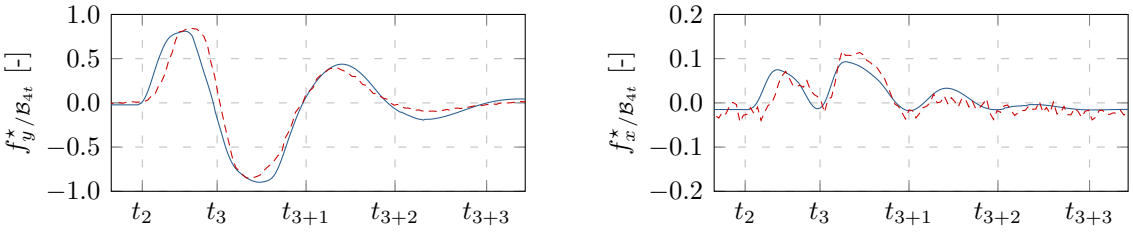


Figure 5: Starboard fin angle – Active fins : ---- experimental tests, —— present work



(a) Starboard fin – Lift force (b) Starboard fin – Drag force

Figure 6: Hydrodynamic forces on fins – Active fins : ---- experimental tests, —— present work

4.2 2D flexible membrane

This application is a typical validation case for fluid structure interaction problems dealing with flexible beams. A flexible beam is attached to the upstream face of a fixed rigid square. Due to the vortex shedding created by the square, the flexible beam is exited and presents strong oscillations which are taken into account with a deformable mesh technique. The settings of this numerical two-dimensional case and others references can be found in [17]. The results obtained are shown in figure 7. They are between the classical values which can be found in literature. Visualisations of the velocity fields are also proposed in figure 8.

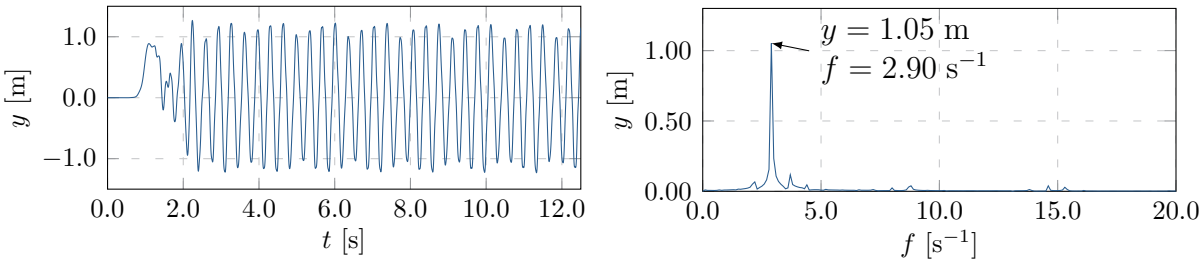
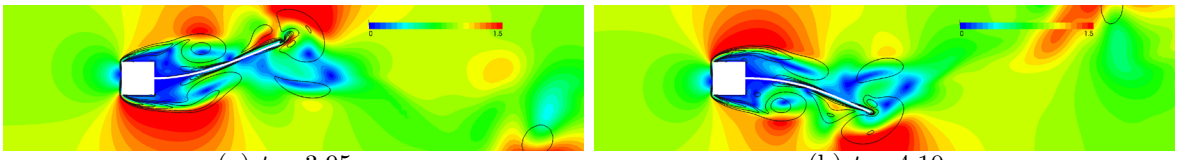


Figure 7: 2D flexible membrane – Vertical displacement of the beam end (left) and FFT (right)



(a) $t = 3.95$ s (b) $t = 4.10$ s

Figure 8: 2D flexible membrane – Dimensionless velocity field and iso-lines of vorticity

4.3 3D flexible barge

The current work is dedicated to the study of a very flexible barge in waves. The results could be compared to experimental tests [18] and numerical simulations using a simplified model [19]. The deformations of the barge for the first simulation with ISIS-CFD and MBDyn can be seen at two different instants in figure 9.

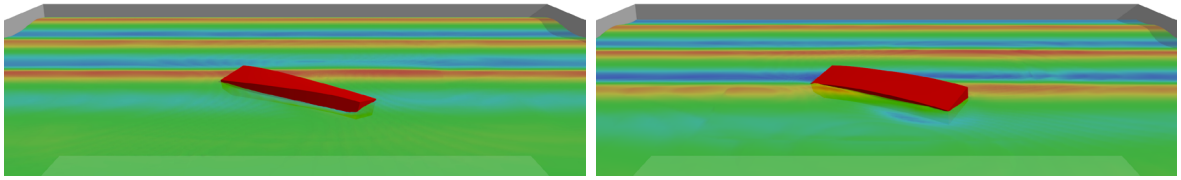


Figure 9: 3D flexible barge – Barge at two different times

5 CONCLUSIONS

This paper describes an efficient and robust algorithm, able to handle configuration with severe added mass effects, while remaining efficient. The coupling iteration occurs during the non-linear iterations of the fluid solver because it is the most costly part in our applications. To tackle the destabilising added-mass effects, a relaxation technique of the structural kinematics is used. The latter can be viewed as a modification of the artificial added mass method and does not need any intrusive modification of the structure solver. The relaxation operator is naturally related to the artificial added mass operator. Efficiency of the coupling is optimal when the artificial added mass is close to the physical one. The computation of the approximated or exact added mass operator for rigid and beam-like bodies is carried out through an on-line resolution of an original pressure-like equation integrated in the fluid solver ([11, 15]). Compared to a non-FSI simulation (unsteady simulation with imposed motion), the number of non-linear iterations to converge the fluid part is similar, even when large added mass effects occur. The additional cost of the FSI cases is then reduced to the resolution of the structural part and to the mesh deformation technique. This computation chain consisting here of ISIS-CFD and MBDyn has been validated numerically for 6 DOF rigid bodies with strong added mass effects [15] as well as for flexible bodies. Among these studies, some of them are briefly described here and comparisons with experimental data or other numerical results are shown.

REFERENCES

[1] Leroyer, A., Barré, S., Kobus, J., and Visonneau, M. *Experimental and numerical investigations of the flow around an oar blade. Journal of Marine Science and Technology* **13**(1), 1–15 (2008).

[2] Leroyer, A. and Visonneau, M. *Numerical methods for RANSE simulations of a self-propelled fish-like body. Journal of Fluids and Structures* **20**(7), 975–991 (2005).

[3] Durand, M. *Interaction fluide-structure souple et légère, application aux voiliers.* PhD thesis, Ecole Centrale Nantes, (2012).

[4] Queutey, P. and Visonneau, M. *An interface capturing method for free-surface hydrodynamic flows. Computers & fluids* **36**(9), 1481–1510 (2007).

[5] Wackers, J., Deng, G., Leroyer, A., Queutey, P., and Visonneau, M. *Adaptive grid refinement for hydrodynamic flows. Computers & Fluids* **55**(0), 85–100 (2012).

[6] Masarati, P. *MBDyn Theory and Developer's Manual Version 1.5.6.* Technical report, Dipartimento di Ingegneria Aerospaziale Politecnico di Milano, (2014).

[7] Masarati, P. *Comprehensive multibody AeroServoElastic analysis of integrated rotorcraft active controls.* PhD thesis, Ph. D. Thesis, Dipartimento di Ingegneria Aerospaziale, Politecnico di Milano, (2000).

[8] Deparis, S., Discacciati, M., and Quarteroni, A. *A domain decomposition framework for fluid-structure interaction problems. Computational Fluid Dynamics 2004* , 41–58 (2006).

[9] Badia, S. and Codina, R. *On some fluid–structure iterative algorithms using pressure segregation methods. Application to aeroelasticity. International Journal for Numerical Methods in Engineering* **72**(1), 46–71 (2007).

[10] Küttler, U. and Wall, W. *Fixed-point fluid–structure interaction solvers with dynamic relaxation. Computational Mechanics* **43**(1), 61–72 (2008).

[11] Söding, H. *How to integrate free motions of solids in fluids.* In *4th Numerical Towing Tank Symposium, Hamburg*, (2001).

[12] Förster, C., Wall, W. A., and Ramm, E. *Artificial added mass instabilities in sequential staggered coupling of nonlinear structures and incompressible viscous flows. Computer Methods in Applied Mechanics and Engineering* **196**(7), 1278 – 1293 (2007).

[13] Badia, S., Quaini, A., and Quarteroni, A. *Splitting methods based on algebraic factorization for fluid-structure interaction, added-mass operator (dt -¿ 0). SIAM Journal on Scientific Computing* **30**(4), 1778–1805 (2008).

[14] Joosten, M. M., Dettmer, W. G., and Perić, D. *Analysis of the block Gauss–Seidel solution procedure for a strongly coupled model problem with reference to fluid–structure interaction. International Journal for Numerical Methods in Engineering* **78**(7), 757–778 (2009).

[15] Yvin, C. *Interaction fluide-structure pour des configurations multi-corps. Applications aux liaisons complexes, lois de commande d'actionneur et systèmes souples dans le domaine maritime.* PhD thesis, Ecole Centrale de Nantes, (2014).

[16] Candelier, F., Boyer, F., and Leroyer, A. *Three-dimensional extension of Lighthill's large-amplitude elongated-body theory of fish locomotion. Journal of Fluid Mechanics* **674**(1), 196–226 (2011).

[17] Dettmer, W. and Perić, D. *A fully implicit computational strategy for strongly coupled fluid–solid interaction. Archives of Computational Methods in Engineering* **14**(3), 205–247 (2007).

[18] Malenica, Š., Molin, B., Remy, F., and Senjanović, I. *Hydroelastic response of a barge to impulsive and non-impulsive wave loads.* In *Proceedings of the 3rd International Conference on Hydroelasticity in Marine Technology*, 107–115, (2003).

[19] Senjanović, I., Malenica, Š., and Tomasĕvić, S. *Investigation of ship hydroelasticity. Ocean Engineering* **35**(5–6), 523 – 535 (2008).

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis (Eds)

FLOW PASSIVE CONTROL ON THE NACA AIRFOILS EXPERIMENTAL AND NUMERICAL STUDY

Hocine Tebbiche and Mohammed S. Boutoudj

Laboratoire d'Energétique, Mécanique et Matériaux –LEMM , Université Mouloud Mammeri, 15 000 Tizi-Ouzou, Algérie
e-mails: tebbichehocine@yahoo.fr, boutoudj_ms@yahoo.fr - web page: http://www.ummto.dz

Key words*:* Airfoil, Boundary layer, Vortex Generators, Lift, Drag, CFD

Summary*: Passive flow control by vortex generators (VGs) is the simplest solution to delay or eliminate the separation of the fluid from the wall. The present work concerns the effectiveness of a new geometry of VGs shaped delta wings implanted on the upper surface of two airfoils NACA (4412 and 0015). The experimental measurements in a wind tunnel at two Reynolds numbers show a significant aerodynamic performances improvement leading to an increase in the lift of 27% and a decrease in the drag of about 30% after stall. The results are in a good agreement with those given by a three dimensional simulation undertaken with a CFD commercial code (Fluent).*

1 INTRODUCTION

The flow on the upper surface airfoil is subject to an adverse pressure gradient when the incidence increases. This leads to the boundary layer separation which causes looses in the aerodynamic performances (lift decrease and drag increase). It is well known that the lift around an airfoil is rather created by the suction on the upper surface than the overpressure on the lower one. The flow control aims to delay or eliminate the fluid separation and its undesirable effects like vibrations and aerodynamic noise.

In the aircraft industries the flow control takes on capital importance for the reduction of the energy overconsumption and the aerodynamic noise. Moreover, during take-off or landing, the speed is low and the attack angle needs to be high for the lift enhancement.

Different techniques of the control exist. Methods such as suction, blowing, moving surface, vortex generators [1-5] have been rigorously investigated and also used in practice with a miscellaneous degree of success. Active methods are however difficult to implement and they cause problems of congestion.

Passive vortex generators (VGs) are simple use and known to bring momentum in the boundary layer which leads to the delay or suppression of the flow separation. Their interest is all the greater as they also have the advantage to be retractable when not in use.

Vortex generators enhance the aerodynamic performances and the most efficiency are the Lin's ones V-shaped when their height is less than the boundary layer thickness [2]. The boundary layer control by vortex generators can also be associated with another passive form such as riblets or roughness [6] to influence the transition from laminar to turbulent flow.

The present work focuses on the control of incipient separation on the upper NACA airfoils (4412 and 0015) by means of VGs with delta shape to improve the aerodynamic performances. The choice of the delta wing configuration type to control the separated flow is linked to the formation of two strong vortices at its leading edges. These vortices are contra rotating ones which are the most efficiency for the flow control. Furthermore, they break down only at high incidences.

The experimental results show the control influence on the lift and the drag of the airfoils as well as the importance of the apex angle of the VGs. These results are compared to those given by a three dimensional numerical simulation with a CFD code (Fluent).

2 GLOBAL CHARACTERISTICS OF THE FLOW AND PRELIMINARY TWO-DIMENSIONAL SIMULATION

The interaction between the fluid and an airfoil results in two forces, lift and drag, which are commonly given by the two respective aerodynamic coefficients:

$$C_L = \frac{F_y}{\frac{1}{2}\rho U_\infty^2 S} \tag{1}$$

And

$$C_d = \frac{F_x}{\frac{1}{2}\rho U_\infty^2 S} \tag{2}$$

F_y and F_x are respectively the lift and the drag, ρ is the volumic weight, S the surface airfoil and U_∞ the upstream velocity.

The pressure coefficient is defined by the following expression:

$$Cp = \frac{P - P_0}{\frac{1}{2}\rho U_\infty^2} \tag{3}$$

P is the local pressure and P_0 the upstream reference pressure.

The boundary layer which develops around the profile is such that its thickness is equal to the distance from the wall to which the speed reaches 99% of the external velocity. Two-dimensional simulations of the flow around the airfoils were undertaken using a CFD code (Fluent) in order to evaluate the boundary layer thickness from the velocity distribution. The

VGs height is generally expressed as a function of this parameter and directly affects the aerodynamic forces.

To ensure that the use of finer or coarser mesh does not affect the quality of the simulation results, it is necessary to properly positioning the wall parameter y^+. This parameter is defined only at the first adjacent cell to the wall; it depends on the grid resolution and the flow velocity. The value of y^+ dictates how the shear wall is calculated.
The size of the first cell is approached by using expression (4) below:

$$y^+ = \frac{yU_\infty \sqrt{\overline{C}_f/2}}{\nu} \tag{4}$$

y is the normal distance to the profile, $\overline{C_f}$ the skin friction coefficient, ν the kinematic viscosity.

By assimilating the airfoil to a flat plate, the skin friction coefficient can be estimated from the following empiric relation [7]:

$$\overline{C}_f/2 \approx 0.037\,\mathrm{Re}_L^{-0.2} \tag{5}$$

where Re_L is the Reynolds number related to the chord length.

Fig. 1 represents the dimensionless velocity obtained at 16% of the airfoil chord versus the distance from the wall, for an incidence of 12 degrees and a Reynolds number with equals 2.5 10^5. This velocity evolution allows an estimation of the boundary layer to 3 mm in these conditions. The numerical simulation validation was made by superposition of the wall pressure results with experimental data. The comparison in Fig. 2 between our computational results and those given by Pinkerton [8] experiment on the NACA 4412 airfoil shows a good agreement. A similar comparison was carried out in a previous work on the symmetrical profile NACA 0015 [9] and the boundary layer thickness calculated. Its value was found to be equal to 10 mm at 18% of the airfoil chord for the same Reynolds number and an incidence of 13 degrees.

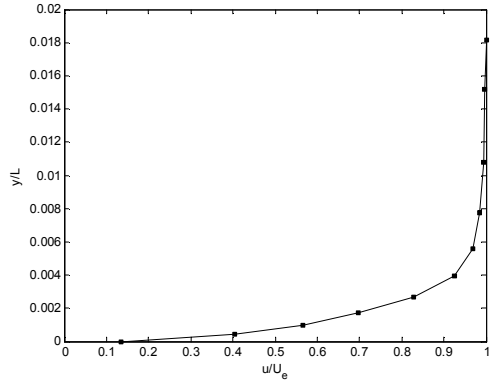


Fig. 1: Velocity profile at X/L=0.16, NACA 4412 airfoil, U_e is the external velocity.

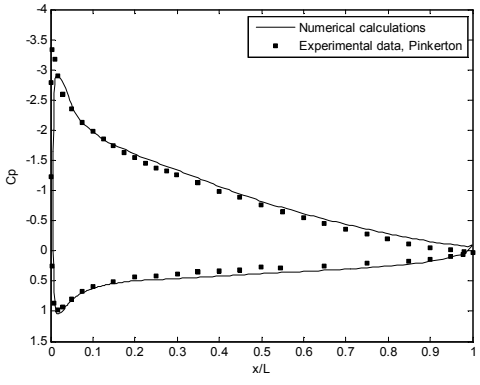


Fig. 2: Distribution of the pressure coefficient around the NACA 4412 profile, α=12°.

3 EXPERIMENTAL CONDITIONS

3.1 Wind tunnel and acquisition chain

The studied models are the NACA 4412 and the NACA 0015 profiles; the chord length is 150 *mm* and the depth is equal to 200 *mm*. Both airfoils are equipped with pressure taps laid out with the suction face for the measurement of the pressure field. Lift and drag forces were measured by way of an aerodynamic balance connected to an acquisition chain. Each test realized was repeated three times and averaged. The time of acquisition was 60 *s* with 500 *Hz* frequency.

All the experiments were performed in a Deltalab™ type open circuit. Maximum speed is higher than 45 *m/s*. The turbulence rate is fixed by a grid at the entry of 5×5 mm^2. The length and the wind tunnel section are respectively 100 *cm* and 30×30cm^2.

3.2 Passive Vortex Generators

The studied vortex generators are small delta wings placed on line on the suction face of the airfoils, at different locations from the leading edge. The choice of the delta wing configuration type resides in the fact that the delta wing generates two strong vortices at the leading edges which bring momentum in the flow. These delta wings are set at weak relative incidences in order to avoid their unhooking and the vortices bursting. Fig. 3 shows these VGs on the airfoil.

The geometrical parameters of the VGs such as the height H, the spacing λ, the apex angle β, the relative incidence angle of the VGs γ as well as their position according to the chord length affect the flow control efficiency. The optimal values of H, λ, γ and the position were estimated in a previous work [10] as $H/\delta = 1.8$, $\lambda/H = 3.8$ and $\gamma = 10°$ when they are placed at 10% of the chord on the upper surface of a NACA 4412 airfoil, for $\beta = 45°$. The apex angle of the VGs is examined in the present study and varied in the interval from 30° to 74.6° while the other parameters are kept equal to their optimal values.

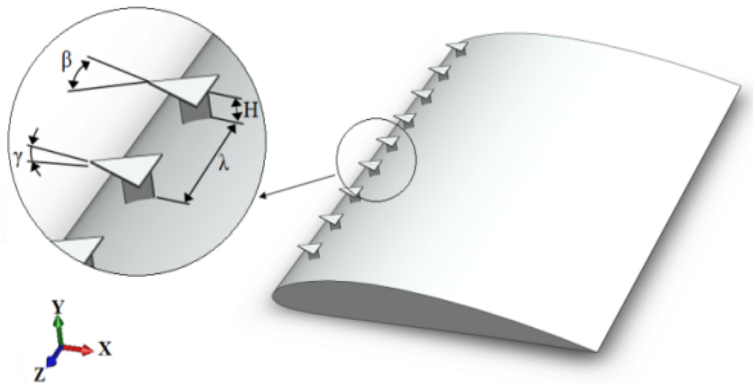


Fig. 3: Shape and disposition of the Delta VGs

4 EXPERIMENTAL RESULTS AND DISCUSSION

4.1 Effect of the flow control on the lift and drag coefficients

The objective of the control is focused on the aerodynamic lift increase and the drag reduction which are widely affected by the flow separation. The vortices generated by the small delta wing are expected to bring momentum in the boundary layer in order to eliminate or delay its separation. Several VGs with different angles of aperture where considered. The experiments were conducted in a range of angles from $\beta = 30°$ until $\beta = 74.6°$ for the study of the apex angles effects. All used angles are privileged ones that are known to create high intensity vortices compared with non-privileged ones [11]. Fig. 4 shows the maximum lift profit for the controlled flow on the NACA 4412 versus the apex angles. A peak value is distinguished for $\beta = 45°$ which produce a maximum with a lift profit that can reach approximately 16.5%.

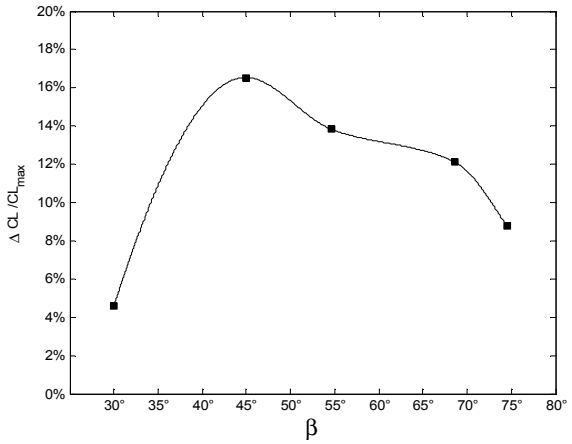


Fig. 4. Apex angle β effect of the Δ-VGs on the maximum lift (8 VGs, Re = 2.5 10^5)

Fig. 5 and Fig. 6 represent respectively the lift and the drag for the most efficiency VGs configuration ($\beta = 45°$) versus the attack angle. Fig. 5 shows a lift gain from 10° increasing to a peak of about 16% in the stall, which is also slightly delayed by one degree in the case of controlled flow.

In terms of drag, we see in Fig. 6 that the flow control introduces an increase from the incidence of 10° to 18° when the tendency reverses. When the incidence is greater than 20°, the drag decreases significantly from 30% in the controlled case. This means that the pressure drag drops drastically following a partial reattachment of the boundary layer on the upper surface of the profile.

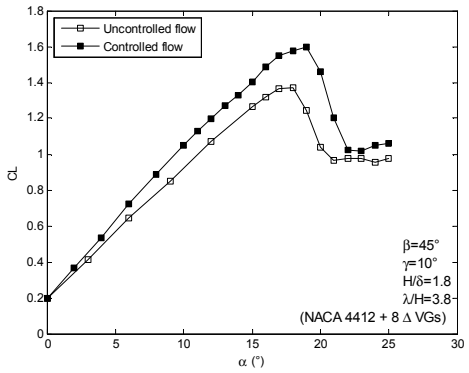


Fig. 5 : Lift coefficient, NACA 4412, Re=2.5 10^5

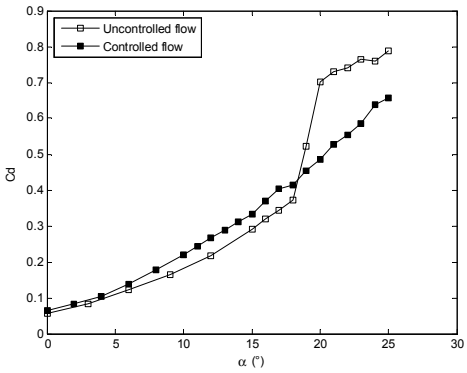


Fig. 6 : Drag coefficient, NACA 4412, Re=2.5 10^5

Fig. 7a and Fig. 7b represent the lift coefficient when we have added two supplemantary VGs, respectively for Re = 2.5 10^5 and Re = 3 10^5. One can see on this figures better performances for the VG as the maximum lift enhancement reaches 20%. On the other hand, the stall occurring at α = 18° is delayed to α = 20° for Re = 3 10^5.

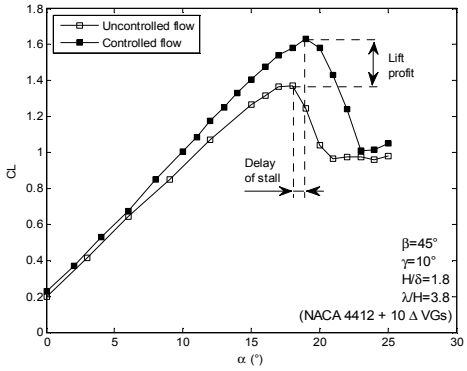


Fig. 7a : Lift coefficient, NACA 4412, Re=2.5 10^5

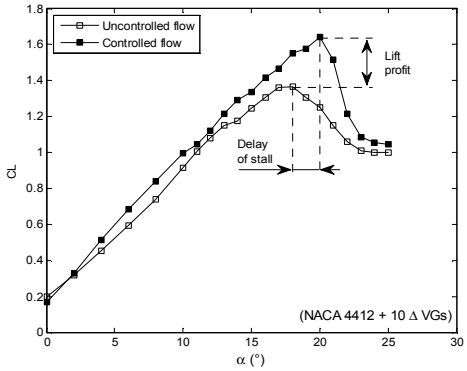


Fig. 7b : Lift coefficient, NACA 4412, Re=3 10^5

The Delta-VGs has also been tested and confirmed expectations for their effectiveness in improving the lift in the case of the NACA 0015 airfoil. Fig. 8 and Fig. 9 show the lift versus the incidence angle obtained with 10 optimized VGs placed at 10% of the chord for two Reynolds numbers.

For both speeds studied, lift increase is noticed. At Reynolds number equal to 2.5 10^5, control has achieved 16% increase in lift (Fig. 8). For the Reynolds number of 3 10^5, the stall occurring at α=14° is postponed until α=16° with a considerable profit in the lift which reaches approximately 27% (Fig. 9).

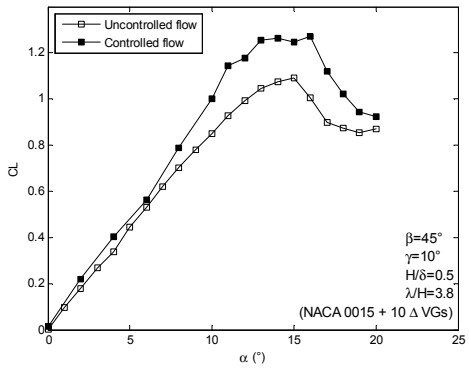


Fig. 8: Lift coefficient, NACA 0015, Re=2.5 10^5

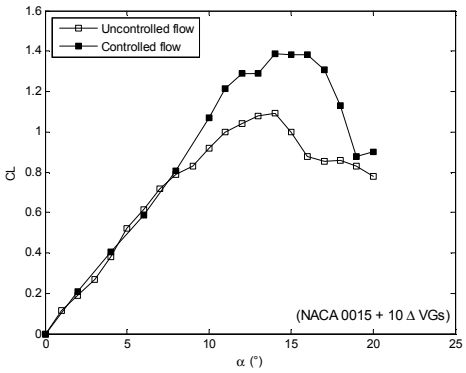


Fig. 9: Drag coefficient, NACA 0015, Re=2.5 10^5

4.2 Pressure field

Experimental pressure fields data related to the controlled case represented on maps of Iso-values (Fig. 10a and Fig. 10b) are given for two velocities and attack angles for the NACA 4412 airfoil. The figures show the formation of two symmetrical low pressure pockets which are spread through 30% of the chord. We notice in Fig. 10c and Fig. 10d a negative pressure peak at the location of the generators between *x/L = 0.1* and *x/L = 0.25*, which reaches a minimum value of about -4.

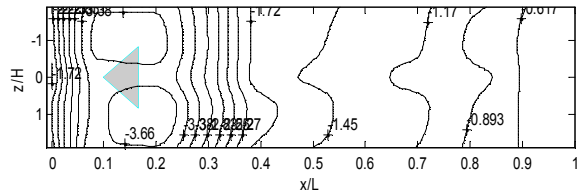


Fig. 10a: Iso-values of the pressure coefficient Re=2.5 10^5, α=19°

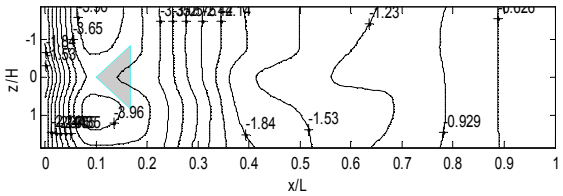


Fig. 10b: Iso-values of the pressure coefficient Re=3 10^5, α=20°

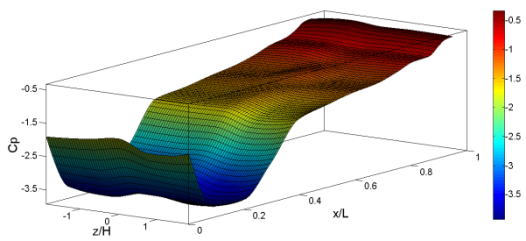


Fig. 10c: 3D pressure coefficient, Re=2.5 10^5, α=19°.

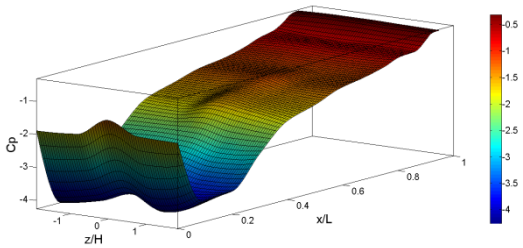


Fig. 10d : 3D pressure coefficient, Re=3 10^5, α=20°.

5 THREE DIMENSIONAL SIMULATION

5.1 Methodology

The turbulent flow around the profiles is calculated in the case of a three-dimensional incompressible steady averaged Navier-Stokes (RANS) using the Fluent commercial calculation code. The turbulence model used is the *kω-SST*, the size of the first mesh is maintained at $y^+ \approx 2.5$ at $Re = 2.5\ 10^5$. In order to reduce the computation time and faster convergence to the desired solution, we chose to simulate only one vortex generator in the middle of the spanwise airfoil by adopting the geometrical parameters of the Δ-VGs defined in the section 3.2 and corresponding to the experimental conditions.

The calculation box is modeled by a volume of 100 cm length and a vortex generator wide (18.9 mm). The generation of the 3D meshing is a very important step in a CFD analysis as this parameter influences the computed solution. The generated meshing is hybrid; its interest is in the fact that it combines the privileges of structured meshes and those of unstructured grid by reducing errors. Close to the profile's walls, the generated meshes are hexahedral followed by relaxed prismatic meshes. As for the vortex generators, due to the complexity of the showed geometry, the VG is delimited by a sub-domain and meshed by cells of kinds, hexahedral tetrahedral and pyramidal (Fig. 11).

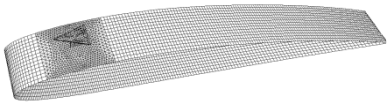


Fig. 11: Meshing of the NACA 4412 profile with a Delta-VG

5.2 Aerodynamic loads

Fig. 12 shows the experimental and the numerical data obtained for the lift and the drag coefficients versus the attack angle for the NACA 4412 airfoil while Fig. 13 is relative to the NACA 0015. One can see on these figures a very good agreement for the lift coefficient. However, the numerical values are lower than the experimental ones for the drag coefficient.

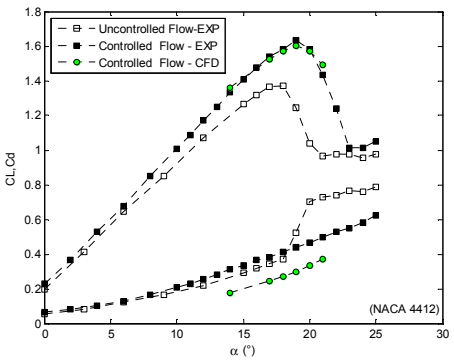


Fig. 12: Lift and drag curves, experimental and numerical results, Re=2.5 10^5.

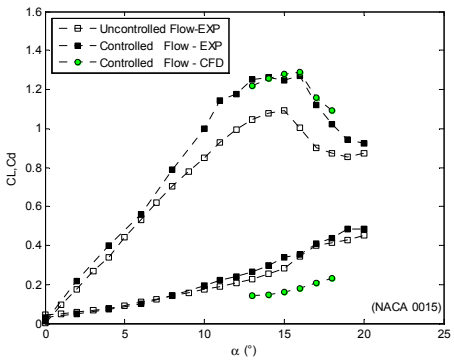


Fig. 13: Lift and drag curves, experimental and numerical results, Re=2.5 10^5.

6 CONCLUSION

Delta wings shaped vortex generators have been proposed for passive flow control on the upper surface of the NACA 4412 and the NACA 0015 profiles.
The experimental study in a wind tunnel highlighted appreciable improvement in their aerodynamic performances. Thereby, the flow control showed a lift increase in order of 27% for the symmetrical airfoil, 20% for the cambered one accompanied by a drag decrease after stall. The effect of the VGs apex angle value has also been highlighted and the most effectiveness of the control was obtained with the apex of 45°.

Furthermore, the wall pressure field measurement shows a peak of depression at the location of the vortex generators.

The numerical values obtained with the three dimensional simulation for the lift coefficient are in good agreement with the experimental data. However, the numerical values are lower for the drag coefficient.

REFERENCES

[1] Taylor, H. D, *The elimination of diffuser separation by vortex generators,* United Aircraft Corporation Connecticut, Research Department Report No R-4012-3, (1947).
[2] Lin, J. C., *Control of turbulent boundary-layer separation using micro-vortex generators*, AIAA paper, (99-3404), (1999)
[3] Godard, G. and Stanislas, M., *Control of a decelerating boundary layer*; *Part 1: Optimization of passive vortex generators,* Aerospace Science and Technology, 10(3), (pp. 181-191), (2006)
[4] Bak, C., Fuglsang, P., Johansen, J. and Antoniou, I., *Wind tunnel test of the NACA 63-415 and a modified NACA 63-415 Airfoil*, Riso R-1193, Riso National Laboratory, Roskilde, Denmark, (2000).
[5] Tebbiche, H. and Boutoudj, M.S., *Vortex generators contribution to the enhancement of the aerodynamic performances*, Advanced Materials Research ,Vol. 950, pp. 268-274, (2014).
[6] Tebbiche, H. and Boutoudj, M.S., *Aerodynamic drag reduction by turbulent flow control with vortex generators*, International Symposium on Aircraft Materials, Marrakech, (2014)
[7] Gerasimov, A., *Modeling Turbulent Flows with FLUENT*, Europe, ANSYS, Inc, (2006).
[8] Pinkerton, R. M., *The variation with Reynolds number of pressure distribution over an airfoil section*, Vol. 613, pp. 65-84, National Advisory Committee for Aeronautics, (1938).
[9] Tebbiche, H. and Boutoudj, M. S.; *Optimized Vortex Generators in the Flow Separation Control around a NACA 0015 Profile*, 9th International Conference on Structural Dynamics, Porto, Portugal, pp. 3219-3226, (2014).
[10] Tebbiche, H. and Boutoudj, M. S., *Contribution de générateurs de tourbillons en forme d'ailes delta au contrôle du décollement de la couche limite à l'extrados du profil NACA 4412*, Colloque International « Caractérisation et Modélisation des Matériaux et Structures » CMMS'14, UMMTO, Tizi-Ouzou, Algérie, (2014).
[11] Le Ray, M., Deroyon, M J., Deroyon, J.P. and Minair, C., *Critères angulaires de stabilité d'un tourbillon hélicoïdal, rôle des angles privilégiés dans l'optimisation des ailes, voiles, coques des avions et navires*, Association technique et maritime, 1985.

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis (Eds)

HIGHER-ORDER FEM FOR NONLINEAR HYDROELASTIC ANALYSIS OF A FLOATING ELASTIC STRIP IN SHALLOW-WATER CONDITIONS

ANGELIKI E. KARPERAKI[*1], KOSTAS A. BELIBASSAKIS[*2], THEODOSIOS K. PAPATHANASIOU[†3] AND STILIANOS I. MARKOLEFAS[‡4]

[*]School of Naval Architecture and Marine Engineering
National Technical University of Athens
Heroon Polytechniou 9, Zografos 15773, Athens, Greece
e-mail: [1]karperaki.ang@gmail.com, [2]kbel@fluid.mech.ntua.gr, http://arion.naval.ntua.gr/~kbel/

[†] School of Applied Mathematical and Physical Science
National Technical University of Athens
Heroon Polytechniou 9, Zografos 15773, Athens, Greece
email: [3]papathth@gmail.com

[‡] Department of Mechanical Engineering
Technological Educational Institute of Central Greece
Psachna, Evia 34400
e-mail: [4]markos34@gmail.com

Key words: Higher-order FEM, hydroelasticity, Shallow Water

Abstract. The hydroelastic response of a thin, nonlinear, elastic strip floating in shalow-water environment is studied by means of a special higher order finite element scheme. Considering non-negligible stress variation in lateral direction, the nonlinear beam model, developed by Gao, is used for the simulation of large flexural displacement. Full hydroelastic coupling between the floating strip and incident waves is assumed. The derived set of equations is intended to serve as a simplified model for tsunami impact on Very Large Floating Structures (VLFS) or ice floes. The proposed finite element method incorporates Hermite polynomials of fifth degree for the approximation of the beam deflection/upper surface elevation in the hydroelastic coupling region and 5-node Lagrange finite elements for the simulation of the velocity potential in the water region. The resulting second order ordinary differential equation system is converted into a first order one and integrated with respect to time with the Crank-Nicolson method. Two distinct cases of long wave forcing, namely an elevation pulse and an N-wave pulse, are considered. Comparisons against the respective results of the standard, linear Euler-Bernoulli floating beam model are performed and the effect of large displacement in the beam response is studied.

1 INTRODUCTION

The hydroelastic interaction problem of free surface gravity waves with large, floating

bodies is found in numerous applications, ranging from the design and construction of marine structures to the response analysis of ice formations. Advances is marine technology, along with the growing need for commercial space in the condensed coastal areas has led to the rise of Very Large Floating Structures (or VLFS). Nowadays, VLFS are used, both near shore and in the open ocean, as energy plants, docking stations, storage facilities and even as floating airports and recreational amenities [1]. The analysis of such marine structures, is an important stepping stone towards robust design and construction [1-3].

Hydroelastic analysis is also relevant in the case of large ice floes under ocean wave excitation [4]. The continuous oscillatory, flexural motion undergone by large ice formations lead to their eventual splitting and disintegration. The demise of the Marginal Ice Zone (MIZ), the boundary between ice shelves and the open ocean occupied predominantly by ice floes, is linked with wave forcing [5]. In absence of a dense MIZ sums of wave energy reach the ice shelves leading to collapse events with profound environmental impact.

As already mentioned, both of the above problems have their foundations set in hydroelasticity [1-4]. Their large horizontal dimensions compared to thickness, make hydroelastic effects dominant. Large floating structures are most commonly modelled in the literature as plates with zero or non-zero draft. The Kirchhoff thin plate theory is employed in the majority of works [6-8], while some consider the Reissner-Mindlin and Von Karman plate models accounting for shear deformation and large deflection effects [9-114].

For the hydrodynamic modelling, typically the linearised wave theory is utilised. When dealing with harmonic excitation eigenfunction expansion methods [12], Galerkin schemes [13] and Green functions have been employed for the solution of the hydroelastic problem in the frequency domain. However, the consideration of irregular loading dictates time domain analysis tools, such as direct integration schemes [14] and Fourier transforms [15]. Considering long wave excitation, Sturova [16], developed an eigenfunction expansion technique for the calculation of the dynamic response of a floating, thin, elastic plate of variable thickness over shallow bathymetry regions. Following the same line work, Papathanasiou et al. [17] consider a higher order finite element scheme for the solution of the transient hydroelastic problem posed by a thin, elastic, heterogeneous beam floating over shallow waters.

Since large floating bodies are expected to span over great distance, the effects of variable bathymetry must also be taken into account. Belibassakis and Athanassoulis [18] derived a consistent coupled mode method for the hystroelastic analysis of a thin floating body over general bathymetry, exhibiting continuous variation. In [19] and [20] the authors extended previous work in order to account for weakly non-linear wave excitation and shear deformable bodies over general bathymetries.

In the present work, the finite element method is employed for the solution of the 1D hydroelastic problem of a uniform, elastic strip floating over an uneven bottom, under shallow water conditions. The employed shallow water assumption allows for the study of tsunami impact on large floating bodies, like VLFS. The Gao beam theory [21], accounting for large defections but infinitesimal strains, is used for the approximation of the elastic strip response while the non-linear shallow water theory is chosen for the hydrodynamic model. In Section 2 of the paper the governing equations of the model in question are presented. Subsequently, in Section 3, the equivalent variational problem is derived and the proposed finite element scheme implementation is presented. In Section 4, a series of numerical results are presented

for a given configuration and variable steepness for the long wave excitation, in the form of an elevation pulse. Finally, results of the hydroelastic code are provided for the case of an N-wave excitation.

2 GOVERNING EQUATIONS

In this section, the hydroelastic problem of a thin, elastic strip floating over shallow waters is presented. At the area where hydroelastic coupling is present the following equations hold,

$$m\partial_t^2\eta - I_r\partial_t^2\partial_x^2\eta + D\partial_x^4\eta - s(\partial_x\eta)^2\partial_x^2\eta + \rho_w g\eta + \rho_w\partial_t\phi + \frac{\rho_w}{2}\left(\partial_x\phi\right)^2 = -q(x,t) \tag{1}$$

$$\partial_t\eta + \partial_x\left([b(x)+\eta]\partial_x\phi\right) = 0, \tag{2}$$

where a superimposed dot denotes differentiation with respect to time. In the above equations ρ_w, g are the water density and gravity acceleration respectively, $q(x,t)$ denotes an external load applied to the beam and $b(x)$ is the, possibly varying, bathymetry. Assuming that the density of the beam is ρ_e, its Elastic Modulus E and Poisson's coefficient ν, while its thickness is τ, the constants appearing in Eq. (1) are $m = \rho_e\tau$ the mass per width, $I_r = \rho_e\tau^3/12$ the rotary inertia per width and $D = E\tau^3(1-\nu)(1+\nu)^{-1}(1-2\nu)^{-1}/12$ is the flexural rigidity per width. Finally, $s = 3E\tau(1-\nu^2)^{-1}/2$ is the coefficient of the nonlinear term in the Gao beam model per width [21]. Compared to the classical Euler – Bernoulli beam, the above model incorporates the effects of rotary inertia, introduced by Lord Rayleigh and the nonlinear term $-S(\partial_x\eta)^2\partial_x^2\eta$ derived by Gao for the large deflection of a thin beam, when non-negligible stress variation in the lateral direction is considered. In addition, the pressure forcing terms $\rho_w g\eta + \rho_w\dot{\phi} + \rho_w\left(\partial_x\phi\right)^2/2$ appearing in Eq. (1), include the nonlinear term $\rho_w\left(\partial_x\phi\right)^2/2$ which is significant when the velocity $u = \partial_x\phi$ becomes large.

In the regions where no floating structure is present, the Shallow Water Equations (SWE) are considered for the long wave propagation simulation. The equations read

$$\partial_t u + u\partial_x u + g\partial_x\eta = 0, \tag{3}$$

$$\partial_t\eta + \partial_x\left([b(x)+\eta]\partial_x\phi\right) = 0, \tag{4}$$

In the following, assuming sufficient regularity and introducing the velocity potential ϕ such that $u = \partial_x\phi$, we will express equations (3) and (4) as a single evolution equation for ϕ. For that, let us differentiate Eq. (3) with respect to t and equation (4) with respect to x, to get

$$\partial_t^2 u + \frac{1}{2}\partial_{xt}^2\left(u^2\right) + g\partial_{xt}^2\eta = 0, \tag{5}$$

$$\partial_{xt}^2\eta + \partial_x^2\left([b(x)+\eta]u\right) = 0, \tag{6}$$

Substituting the term $\partial_{xtt}^2\eta$ of equation (5) using (6) and setting $u=\partial_x\phi$,

$$\partial_{xtt}^3\phi+\frac{1}{2}\partial_{xt}^2\left(\partial_x\phi\right)^2-g\partial_x^2\left([b(x)+\eta]\partial_x\phi\right)=0 \tag{7}$$

Integrating (7) with respect to x leads to,

$$\partial_{xtt}^3\phi+\frac{1}{2}\partial_{xt}^2\left(\partial_x\phi\right)^2-g\partial_x^2\left([b(x)+\eta]\partial_x\phi\right)=0 \tag{8}$$

Select the constant $C(t)=0$. Setting $u=\partial_x\phi$ directly in equation (3) and integrating with respect to x, we get

$$\partial_t\phi+\frac{1}{2}\left(\partial_x\phi\right)^2+g\partial_x\eta=c(t) \tag{9}$$

Setting $c(t)=0$ equation (9) becomes,

$$\eta=-\frac{1}{2g}\left(\partial_x\phi\right)^2-\frac{1}{g}\partial_t\phi \tag{10}$$

Finally, eliminating η from Eq. (8) using Eq. (10)

$$\partial_t^2\phi+\frac{1}{2}\partial_t\left(\partial_x\phi\right)^2-g\partial_x\left(b(x)\partial_x\phi\right)+\frac{1}{2}\partial_x\left(\partial_x\phi\right)^3+\partial_x\left(\partial_t\phi\partial_x\phi\right)=0 \tag{11}$$

Introducing the nondimensional quantities $\tilde{x}=L^{-1}x$, $\tilde{t}=g^{1/2}L^{-1/2}t$, $\tilde{\eta}=L^{-1}\eta$, $\phi=g^{-1/2}L^{-3/2}\phi$, where L denotes the Length of the beam, equations (1), (2) and (11) become, after dropping tildes

$$M\varepsilon\partial_t^2\eta-I_R\varepsilon^3\partial_t^2\partial_x^2\eta+\varepsilon^3K\partial_x^4\eta-\varepsilon S(\partial_x\eta)^2\partial_x^2\eta+\eta+\partial_t\phi+2^{-1}\left(\partial_x\phi\right)^2=Q(x,t), \tag{12}$$

$$\partial_t\eta+\partial_x\left([B(x)+\eta]\partial_x\phi\right)=0, \tag{13}$$

$$\partial_t^2\phi+2^{-1}\partial_t\left(\partial_x\phi\right)^2-\partial_x\left(B(x)\partial_x\phi\right)+2^{-1}\partial_x\left(\partial_x\phi\right)^3+\partial_x\left(\partial_t\phi\partial_x\phi\right)=0, \tag{14}$$

where $M=\rho_e/\rho_w$, $I_R=\frac{\rho_e}{12\rho_w}$, $K=\frac{E(1-v)}{12(1+v)(1-2v)\rho_w gL}$, $S=\frac{3E}{2(1-v^2)\rho_w gL}$, $B(x)=\frac{b(x)}{L}$ $Q(x,t)=\frac{-q(x,t)}{\rho_w gL}$, and $\varepsilon=\tau/L\ll 1$.

The bending moment and shear force inside the beam are

$$M_b=\varepsilon^3K\partial_x^2\eta \text{ and} \tag{15}$$

$$V=\varepsilon^3K\partial_x^3\eta-\varepsilon^3I_R\partial_t^2\partial_x\eta-\varepsilon S3^{-1}(\partial_x\eta)^3, \tag{16}$$

In equation (12), there appear four constants, namely $M\varepsilon$, $I_R\varepsilon^3$, $K\varepsilon^3$ and $S\varepsilon$. Observe that two of them scale as $O(\varepsilon)$, while the other two are $O(\varepsilon^3)$. In addition, constants K and S are inversely proportional to L, the length of the beam.

2.1 Initial-Boundary Value Problem formulation

In order to formulate the initial-boundary value problem of a freely-floating strip interacting with a surface wave over shallow water conditions, let us define the three non-overlapping sets $\Omega_1 \equiv (-\infty, 0)$, $\Omega_0 \equiv (0, L)$ and $\Omega_2 \equiv (L, \infty)$ (see Fig. 1).

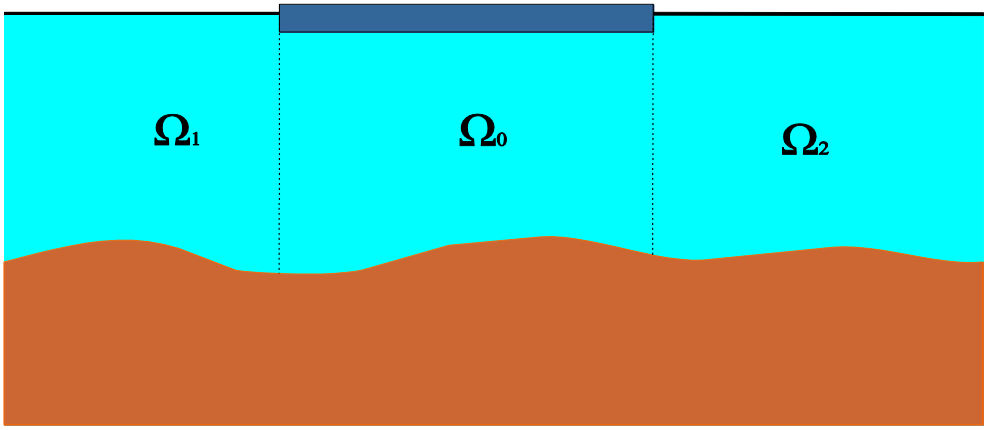


Figure 1: The initial boundary value problem configuration

Considering the non-dimensional Eqs, (12)-(14), derived in the previous section, the IBV problem is,

Find $\eta_0 : \Omega_0 \to \mathbb{R}$, $\phi_i : \Omega_i \to \mathbb{R}$, $i = 0,1,2$, *such that*

$$\partial_t^2\phi_1 + 2^{-1}\partial_t\left(\partial_x\phi_1\right)^2 - g\partial_x\left(b(x)\partial_x\phi_1\right) + 2^{-1}\partial_x\left(\partial_x\phi_1\right)^3 + \partial_x\left(\partial_t\phi_1\partial_x\phi_1\right) = 0 \text{, in } \Omega_1\times(0,T] \tag{17}$$

$$\begin{aligned} M\varepsilon\partial_t^2\eta_0 - I_R\varepsilon^3\partial_t^2\partial_x^2\eta_0 + \varepsilon^3 K\partial_x^4\eta_0 - \varepsilon S(\partial_x\eta_0)^2\partial_x^2\eta_0 \\ + \eta_0 + \partial_t\phi_0 + 2^{-1}\left(\partial_x\phi_0\right)^2 = Q(x,t) \end{aligned} \tag{18}$$

$$\text{and } \partial_t\eta + \partial_x\left([B(x)+\eta]\partial_x\phi\right) = 0 \text{ , in } \Omega_0\times(0,T], \tag{19}$$

$$\partial_t^2\phi_2 + 2^{-1}\partial_t\left(\partial_x\phi_2\right)^2 - g\partial_x\left(b(x)\partial_x\phi_2\right) + 2^{-1}\partial_x\left(\partial_x\phi_2\right)^3 + \partial_x\left(\partial_t\phi_2\partial_x\phi_2\right) = 0 \text{, in } \Omega_2\times(0,T]. \tag{20}$$

with $\eta_i = -\partial_t\phi_i - 2^{-1}\left(\partial_x\phi_i\right)^2$, $i = 1,2$.

The equations above are supplemented with the following boundary conditions,

$$\partial_x\phi_1(|x|\to\infty,t) = \partial_x\phi_2(|x|\to\infty,t) = 0 \quad t\in(0,T] \text{ and} \tag{21}$$

$$M_b(0,t)=V(0,t)=M_b(1,t)=V(1,t)=0\,,t\in(0,T].$$

Appropriate interface conditions expressing mass and momentum conservation at the interfaces are

$$\left(B(0^-)+\eta_1(0^-,t)\right)\partial_x\phi_1\big|_{x=0^-}=\left(B(0^+)+\eta_0(0^+,t)\right)\partial_x\phi_0\big|_{x=0^+},\ \partial_t\phi_1\big|_{x=0^-}=\partial_t\phi_0\big|_{x=0^+},t\in(0,T] \quad (22)$$

$$\left(B(1^-)+\eta_0(1^-,t)\right)\partial_x\phi_0\big|_{x=1^-}=\left(B(1^+)+\eta_2(1^+,t)\right)\partial_x\phi_2\big|_{x=1^+},\partial_t\phi_0\big|_{x=1^-}=\partial_t\phi_2\big|_{x=1^+},t\in(0,T] \quad (23)$$

The initial state for $t=0$ describing still water conditions and zero upper surface elevation for regions Ω_1 and Ω_0 , while imposing an initial upper surface elevation located at a subdomain of Ω_2 are

$$\eta_0(x,0)=\phi_0(x,0)=0\,,\ \text{in }\Omega_0\,, \quad (24)$$

$$\phi_1(x,0)=\partial_t\phi_1(x,0)=0\,,\ \text{in }\Omega_1\,, \quad (25)$$

$$\phi_2(x,0)=0,\ \partial_t\phi_2(x,0)=-G(x)\,,\ \text{in }\Omega_2\,, \quad (26)$$

3 VARIATIONAL FORMULATION

In the present section the variational formulation of the hydroelastic problem defined in Section 2.1 will be derived. Multiply Eqs. (17), (20) with $w_1\in H^1(\Omega_1)$ and $w_2\in H^1(\Omega_2)$ respectively. Multiply Eq. (18) with $v\in H^2(\Omega_0)$ (function v is not to be confused with Poisson's ratio, a constant) and Eq. (19) with $-w_0\in H^1(\Omega_0)$. Performing integration by parts it is,

$$\begin{aligned}&\int_{-\infty}^{0}w_1\partial_t^2\phi_1dx+\int_{-\infty}^{0}w_1\partial_x\phi_1\partial_{tx}^2\phi_1dx-\left[w_1\left(B(x)+\eta_1\right)\partial_x\phi_1\right]_{-\infty}^{0}\\&+\int_{-\infty}^{0}\partial_xw_1B(x)\partial_x\phi_1dx-2^{-1}\int_{-\infty}^{0}\partial_xw_1\left(\partial_x\phi_1\right)^3dx-\int_{-\infty}^{0}\partial_xw_1\partial_t\phi_1\partial_x\phi_1dx=0\end{aligned}\ ,\ \text{for every } w_1\,,$$

$$\begin{aligned}&M\varepsilon\int_0^L v\partial_t^2\eta_0dx+I_R\varepsilon^3\int_0^L\partial_xv\partial_t^2\partial_x\eta_0dx+\varepsilon^3K\int_0^L\partial_x^2v\partial_x^2\eta_0dx\\&+\varepsilon S3^{-1}\int_0^L\partial_xv(\partial_x\eta_0)^3dx+\left[v\left(\varepsilon^3K\partial_x^3\eta_0-I_R\varepsilon^3\partial_t^2\partial_x\eta_0-\varepsilon S3^{-1}(\partial_x\eta_0)^3\right)\right]_0^L\\&-\varepsilon^3K\left[\partial_xv\partial_x^2\eta_0\right]_0^L+\int_0^L v\eta_0dx+\int_0^L v\partial_t\phi_0dx+2^{-1}\int_0^L v\left(\partial_x\phi_0\right)^2dx=\int_0^L vQ(x,t)dx\end{aligned}\ ,\ \text{for every } v\,,$$

$$-\int_0^L w_0\partial_t\eta_0dx+\int_0^L\partial_xw_0[B(x)+\eta_0]\partial_x\phi_0dx-\left[w_0\left(B(x)+\eta_0\right)\partial_x\phi_0\right]_0^L=0\,,\ \text{for every } w_0\,,$$

$$\begin{aligned}&\int_L^{+\infty}w_2\partial_t^2\phi_2dx+\int_L^{+\infty}w_2\partial_x\phi_2\partial_{tx}^2\phi_2dx-\left[w_2\left(B(x)+\eta_2\right)\partial_x\phi_2\right]_L^{+\infty}\\&+\int_L^{+\infty}\partial_xw_2B(x)\partial_x\phi_2dx-2^{-1}\int_L^{+\infty}\partial_xw_2\left(\partial_x\phi_2\right)^3dx-\int_L^{+\infty}\partial_xw_2\partial_t\phi_2\partial_x\phi_2dx=0\end{aligned}\ ,\ \text{for every } w_2\,.$$

Using the boundary and interface conditions, described by Eqs. (21), (22), (23) the variational problem becomes,

Find η_0 *and* ϕ_i, $i=0,1,2$, *such that for every* $w_i \in H^1(\Omega_i)$, $i=0,1,2$ *and* $v \in H^2(\Omega_0)$ *it is*

$$
\begin{aligned}
&\int_{-\infty}^{0} w_1 \partial_t^2 \phi_1 dx + \int_{-\infty}^{0} w_1 \partial_x \phi_1 \partial_{tx}^2 \phi_1 dx + \int_{-\infty}^{0} \partial_x w_1 B(x) \partial_x \phi_1 dx \\
&-2^{-1} \int_{-\infty}^{0} \partial_x w_1 \left(\partial_x \phi_1\right)^3 dx - \int_{-\infty}^{0} \partial_x w_1 \partial_t \phi_1 \partial_x \phi_1 dx \\
&+M\varepsilon \int_0^L v \partial_t^2 \eta_0 dx + I_R \varepsilon^3 \int_0^L \partial_x v \partial_t^2 \partial_x \eta_0 dx + \varepsilon^3 K \int_0^L \partial_x^2 v \partial_x^2 \eta_0 dx + \varepsilon S 3^{-1} \int_0^L \partial_x v (\partial_x \eta_0)^3 dx \\
&+\int_0^L v \eta_0 dx + \int_0^L v \partial_t \phi_0 dx + 2^{-1} \int_0^L v \left(\partial_x \phi_0\right)^2 dx - \int_0^L w_0 \partial_t \eta_0 dx + \int_0^L \partial_x w_0 [B(x) + \eta_0] \partial_x \phi_0 dx \\
&\int_L^{+\infty} w_2 \partial_t^2 \phi_2 dx + \int_L^{+\infty} w_2 \partial_x \phi_2 \partial_{tx}^2 \phi_2 dx + \int_L^{+\infty} \partial_x w_2 B(x) \partial_x \phi_2 dx \\
&-2^{-1} \int_L^{+\infty} \partial_x w_2 \left(\partial_x \phi_2\right)^3 dx - \int_L^{+\infty} \partial_x w_2 \partial_t \phi_2 \partial_x \phi_2 dx = \int_0^L v Q(x,t) dx
\end{aligned}
\tag{27}
$$

and $(\phi_1(x,0), w_1)_1 = (\phi_0(x,0), w_0)_0 = (\phi_2(x,0), w_2)_2 = 0$, $(\eta_0(x,0), w_0)_0 = (\partial_t \varphi_1(x,0), w_1)_1 = 0$, $\left(\partial_t \phi_2(x,0), w_2\right)_2 = -\left(G(x), w_2\right)_2$, $(\ ,\ )_i$, $i=0,1,2$ being the L^2-inner product in region Ω_i.

3.1 Finite Element Implementation

The numerical solution of the variational problem described in Eq. (27) is derived by means of the finite element method. The free water surface regions are approximated by quadratic Lagrange elements while a special element is introduced for the hydroelasticity dominated region. The reader is directed to the work of Papathanasiou *et al.* [17] for a more in depth analysis. The hydroelastic element incorporates fifth order Hermite polynomials for the interpolation of the beam deflection/upper surface elevation in the domain of the hydroelastic coupling and fourth order Lagrange polynomials for the interpolation of the velocity potential (see Fig. 2). The straightforward discretization of Eq. (27), and the substitution of the approximate solutions results in the following system of nonlinear ordinary differential equations,

$$\mathbf{M}\ddot{u} + \mathbf{C}(u)\dot{u} + \mathbf{K}(u)u = 0 \ . \tag{28}$$

After setting $\dot{u} = y$ and taking $\mathbf{z} = [u \ y]^{\mathrm{T}}$, Eq. (28) is reduced to the first order system of nonlinear equations, $\mathbf{A}\dot{z} + \mathbf{B}(z)z = 0$. This last equation is integrated with respect to time using the Crank-Nicolson method.

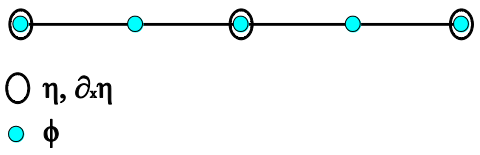


Figure 2: Schematic of the special hydroelastic finite element [17].

4 NUMERICAL RESULTS

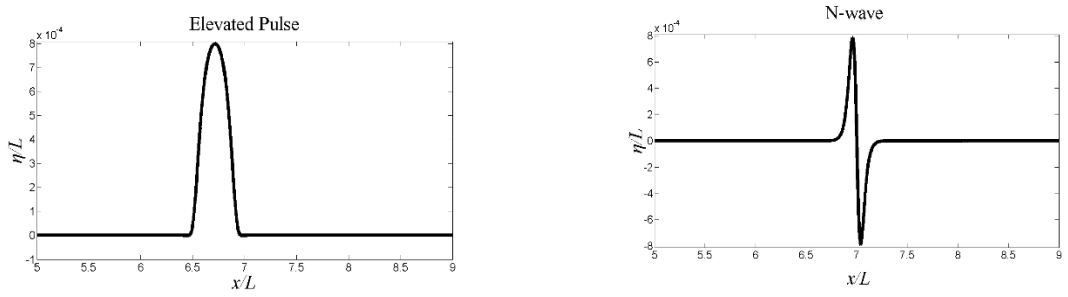
Numerical results for the proposed, higher-order finite element methodology will be presented in this section. The cases of an elevated and an N-wave pulse, typical in long wave modelling, are considered. The elevated pulse is given by

$$G(x) = A\exp(-x^{-\mu(x_0+w^2)} - x^{-\mu(x-x_0+w)(x-x_0-w)}), \tag{29}$$

where A is the amplitude, x_0 is the point of origin, w is the wavelength and μ is a positive parameter controlling the smoothness of the initial pulse. For the isosceles N-wave profile, following Tadepalli and Synolakis [22], initial excitation is given by,

$$G(x) = Ad(x-x_0)\mathrm{sech}^2(\sqrt{\frac{3A}{4d^3}}(x-x_0)), \tag{30}$$

where d is the local depth at the origin. Figure 3, shows the corresponding initial upper surface disturbance, for each of the considered cases, that is allowed to propagate in the free surface region Ω_2.



(a) (b)

Figure 3: (a) Initial excitation in the form of an elevated pulse with $A = 0.4m$ and wavelength $w = 265m$ (b) Initial excitation in the form of an N-wave with $A = 0.4m$ and $L = 265m$

The bathymetric profile considered in the following examples is kept flat underneath the strip, at a depth of $5m$. At a distance, equal to strip length, from the right edge of the floating body the depth is allowed to increase linearly until it is kept constant at 15 m for the rest of the Ω_2 subdomain. Furthermore, the strip thickness is assumed uniform at $2m$, while its length is taken as $500m$. Finally, the material constants selected are material density $\rho_e = 922.5\ kg/m^3$, water density $\rho_w = 1025\ kg/m^3$, Young's modulus $E = 5\cdot 10^9\ Pa$ and Poisson's ratio $v = 0.3$. The acceleration of gravity is $g = 10\ m/\sec^2$.

4.1 Elevation pulse

For the following analysis 100 special hydroelastic elements (Ω_0) and 10000 time steps were used for the calculation of the transient strip response. The elevation pulse parameters are $\mu = 50$ and $A = 0.4$. In Figure 3, a visual representation of the upper surface elevation

solution is shown against time. At the beginning of time, the free surface disturbance set to originate in Ω_2, splits into two propagating waves, travelling in opposite directions. The excitation is partially reflected when it reaches the inclined seabed, and later when it impacts the strip edge (at the interface between Ω_2 and Ω_0). As the pulse reaches the strip, the hydroelastic wave begins to propagate, showing clear signs of dispersion. The waves propagating over shallower bathymetry travel at lower speeds, as expected.

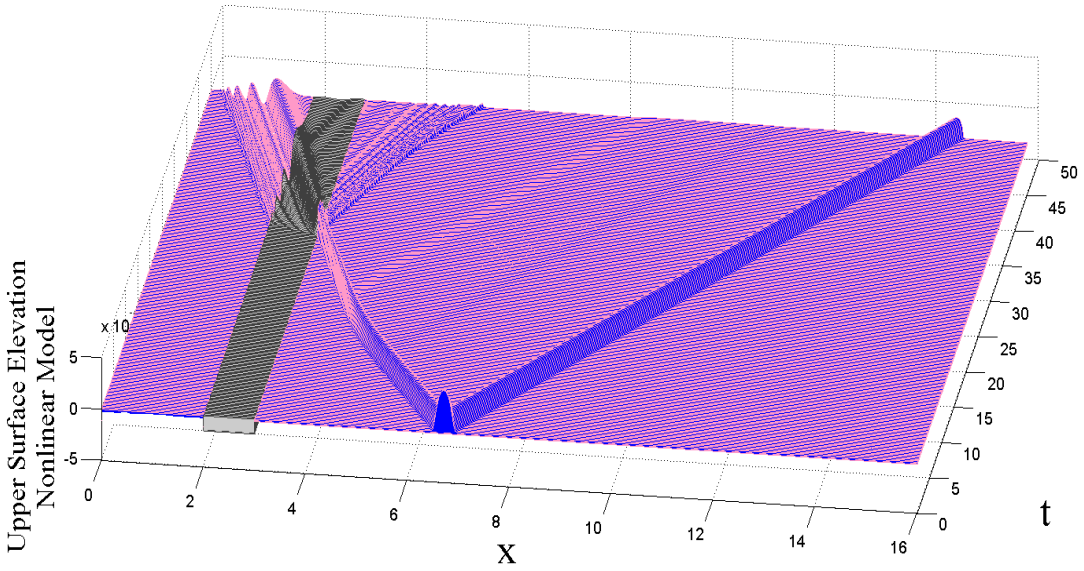


Figure 4: Space-time plot of the elevation pulse propagation. Mild reflection due to variable bathymetry and reflected pulses from the interaction with the floating strip are evident.

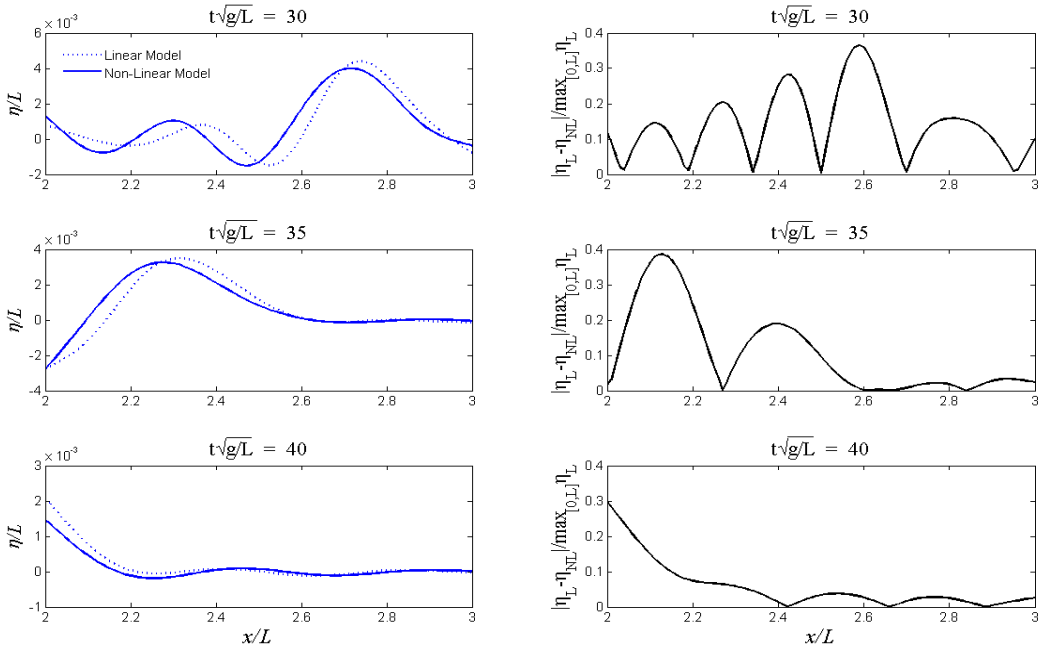


Figure 5: Deflection comparison between the linear and the non-linear models, $A = 0.2$ and $L = 265m$.

In Figures 5 and 6, solutions for the strip deflection using the previously described non-linear model (continuous line) and the linear Euler Bernoulli beam coupled with the linearised shallow water equations [17] (dashed line) are compared, for different values of initial disturbance amplitude, namely $A = 0.2$ and 0.4, at various moments in time. The same spatial and temporal discretizations were used for the derivation of the upper surface/deflection solutions. Notably, the wavelength was kept constant at $w = 265m$, hence the steepness of the initial pulse increases with increasing amplitude. The deviation between the calculated solutions is also presented. In Figure 3, the deviation between the two models, reaches 40 %, marking the importance of non-linear effects when studying the response of thin, elastic strips. As expected, when the elevation pulse steepness increases, in Figure 6, ($A = 0.4$) the deviation between the two models becomes clearer, reaching 65%. The above fact renders non-linear effects critical for the study of the transient hydroelastic response of large floating bodies.

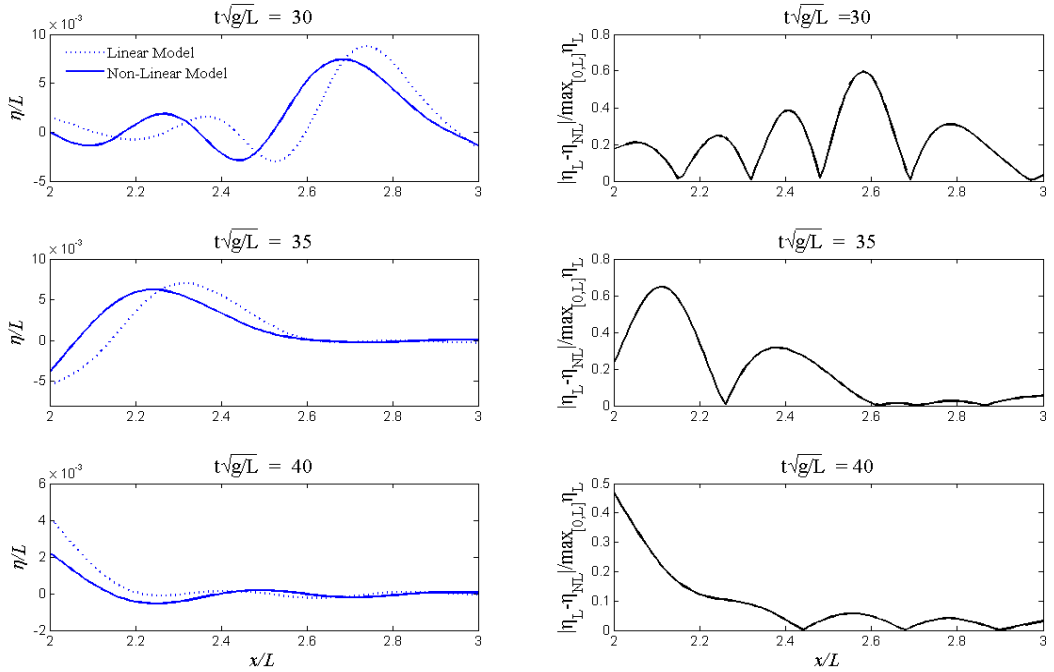


Figure 6: Deflection comparison between the linear and the non-linear models, $A = 0.4$ and $L = 265m$.

4.2 N-wave

Finally, the case of an incoming N-wave excitation was considered. The amplitude was kept at $A = 0.4$. Figure 7, in a similar manner, shows a comparison between the strip deflections calculated with the linear and the non-linear models. The deviations in the calculated deflection solutions, between the two models exceed 100% in some instances. This is attributed to the complex form of the excitation. As it can be seen in Fig. 3, the steepness of the N-wave form is greater than that of the elevated pulse (regardless of the value for the smoothness parameter μ) for the same amplitude and wavelength values. This is straightforward, since the same amplitude is reached within the halfwavelength by the N-

wave. Hence, once again the observed deviations are linked with increasing pulse steepness. Non-linear effects are bound to become dominant for N-wave excitations, making the large deflection assumption made in the present analysis a valid first approximation of such phenomena.

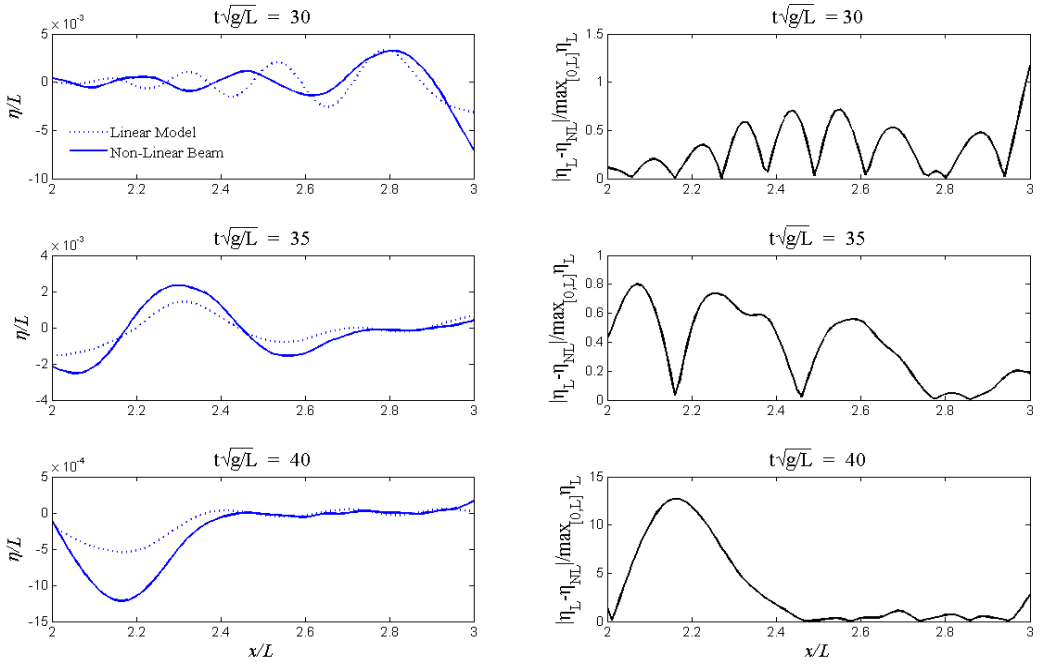


Figure 7: Deflection comparison between the linear and the non-linear models, N-wave case $A = 0.4$ and $L = 265m$.

5 CONCLUSIONS

A higher order finite element method is presented for the solution of the non-linear, hydroelastic problem of a floating strip over shallow water bathymetry. In the present work, the shallow water equations coupled with the Gao [21] non-linear thin beam model, are considered. The proposed higher-order, finite element method incorporates Hermite polynomials of fifth degree for the approximation of the beam deflection/upper surface elevation in the hydroelastic coupling region and 5-node Lagrange finite elements for the simulation of the velocity potential in the water region. The resulting second order ordinary differential equation system is converted into a first order one and integrated with respect to time with the Crank-Nicolson method. This finite element scheme has already been applied to the hydroelastic analysis of linear Euler-Bernoulli beams [17] and is now extended to a nonlinear strip simulation.

Subsequently, the effect of non-linearity, imposed by the large deflections assumption of the beam model is explored. Numerical results, using an elevation pulse and an N-wave as initial excitations, are presented. When compared with the linear Euler-Bernoulli model coupled with the linearised shallow water equations, it becomes evident that non-linear effects increase with increasing excitation steepness. The deviation between the two models peaks at 65% for the case of a ‘steep’ elevation pulse, while for the case of an N-wave, exhibiting the

same amplitude and wavelength characteristics, deviation exceeds 100%. The present work is a stepping stone towards the implementation of higher-order finite element schemes in the solution of the hydroelastic problem of floating bodies, including non-linear effects. Possible extensions include the incorporation of shear deformable floating bodies and inhomogeneous environments of general bathymetry.

REFERENCES

[1] Wang, C.M., Watanabe, E. and Utsunomiya, T. *Very large floating structures*. London, UK: Taylor and Francis (2008).

[2] Chen, X.J., Wu, Y.S., Cui, W.C. and Juncher Jensen, J. Review of hydroelasticity theories for global response of marine structures. *Ocean Eng.* (2006) **33**:439–457. (doi:10.1016/j.oceaneng.2004.04.010)

[3] Watanabe, E., Utsunomiya, T. and Wang, C.M. Hydroelastic analysis of pontoon-type VLFS: a literature survey. *Eng. Struct* (2004). **26**:245–256. (doi:10.1016/j.engstruct.2003.10.001)

[4] Squire, V. A. Synergies between VLFS Hydroelasticity and Sea Ice Reasearch, *Int. J. offshore polar* (2008), **18** (3):1-13.

[5] Squire, V.A. Of ocean waves and sea-ice revisited. *Cold Reg. Sci.Technol* (2007)*,* **49**:110–133.

[6] Evans, D.V. and Davies, T.V. Wave–ice interaction. Report No. 1313 (1968), Davidson Lab, Stevens Institute of Technology, Hoboken, NJ, USA.

[7] Meylan, M.H. and Squire, V.A. The response of ice floes to ocean waves. *J. Geophys. Res.* (1994), **99**:899–900.

[8] Meylan, M.H. Wave response of an ice floe of arbitrary geometry. *J. Geophys. Res.* (2002) **107**:5-1–5-11.

[9] Chen, XJ, Jensen, J.J., Cui, W.C. and Fu, S.X. Hydroelasticity of a floating plate in multidirectional waves. *Ocean Eng.* (2003) **30**:1997–2017.

[10] Endo, H., Yoshida, K. Timoshenko equation of vibration for plate-like floating structures. In *Proc. of 2nd Int. Conf. on Hydroelasticity in Marine Technol*, (1998), 255–264, Fukuoka, Japan.

[11] Papathanasiou, T.K. and Belibassakis K.A. Hydroelastic analysis of VLFS based on a consistent coupled mode system and FEM. *IES J. Part A Civil Struct. Eng* (2014), **7**:195–206.

[12] Kim, J.W. and Ertekin, R.C. An eigenfunction-expansion method for predicting hydroelastic behaviour of a shallow-draft VLFS. In *Proc. 2nd Int. Conf. Hydroelastic Marine Tech* (1998) (eds M Kashiwagi, WKoterayama, M Ohkusu). Fukuoka, Japan: RIAM.

[13] Kashiwagi, M.A. B-spline Galerkin scheme for calculating the hydroelastic response of a very large floating structure waves. *J. Mar. Sci. Tech.* (1998) **3**: 37–49.

(doi:10.1007/BF01239805)

[14] Watanabe, E., Utsunomiya, T. Transient response analysis of a VLFS at airplane landing. In *Proc. Int.Workshop on Very Large Floating Structures* (1996) (eds YWatanabe, EWatanabe), 243–247. Hayama, Japan.

[15] Watanabe, E., Utsunomiya, T. and Tanigaki, S. A transient response analysis of a very large floating structure by finite element method. *Struct. Eng. Earthquake Eng.* (1998), **15**:155–163.

[16] Sturova, I.V. Time- dependent response of a heterogeneous elastic plate floating on shallow water of variable depth. *J. Fluid Mech.* (2009), **637**:305-325.

[17] Papathanasiou, T. K., Karperaki, A., Theotokoglou, E. E. and Belibassakis, K. A. A higher order FEM for time-domain hydroelastic analysis of large floating bodies in an inhomogeneous shallow water environment, Proceedings of the Royal Society A (2014), **471**:2175.

[18] Belibassakis KA, Athanassoulis GA. A coupled mode model for the hydroelastic analysis of large floating bodies over variable bathymetry regions. *J. Fluid Mech.* (2005) **531**: 221–249.

[19] Belibassakis KA, Athanassoulis GA. A couple-mode technique for weakly non-linear wave interaction with large floating structures lying over variable bathymetry. *App. Ocean Res.* (2006) **28**:59–76.

[20] Athanassoulis, G.A. and Belibassakis K.A. A novel-coupled mode theory with application to hydroelastic analysis of thick, non-uniform floating bodies over general bathymetry. *J. Eng.Marit. Environ* (2009), **223**:419–437.

[21] Gao, D.Y. Nonlinear elastic beam theory with applications to contact problems and variational approaches, *Mech. Res. Commun.* (1996), **23**:11–17.

[22] Tadepalli, S. and Synolakis, K Model for the leading waves of tsunamis, *Phys. Rev. Lett.* (1996), **77**:2141-2144.

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis (Eds)

HYBRID ANALYTICAL-NUMERICAL METHODOLOGY FOR COMPUTATIONALLY EFFICIENT PRE-DESIGN ANALYSIS OF FLUID-STRUCTURE INTERACTION

JORDI RUBIO[†], SERGUEI IAKOVLEV[*] AND PEDRO DIEZ[†]

[†] Laboratori de Calcul Numeric, Departament de Matematica Aplicada III
Universitat Politecnica de Catalunya, Barcelona, Spain

[*] Department of Engineering Mathematics and Internetworking, Dalhousie University
Halifax, Canada
email: serguei.iakovlev@dal.ca

Key words: fluid-structure interaction, hybrid techniques, numerical-analytical techniques, pre-design analysis.

Abstract. Analytical approach to modeling the interaction between submerged elastic structures and non-stationary loads has long been recognized as an attractive tool of engineering analysis, especially at the pre-design stage where it has been particularly valued for its high computational efficiency. At the same time, the approach has a number of limitations, the most regrettable one being its inability to handle geometries that are more complex than the basic ones such as a spherical or cylindrical geometry. We present an attempt to overcome this limitation while still preserving the much favored computational efficiency by introducing a hybrid methodology that combines the analytical and finite-element approaches. We then validate the methodology using available experimental data and show that a good agreement with the experiments is observed.

1 INTRODUCTION

In recent years, an efficient semi-analytical methodology has been developed for modelling the interaction between shell structures and non-stationary loads, e.g. [1-4]. The methodology was based on the use of the so-called "response functions", the functions that only depend on the geometry of the system and not on the properties of the fluid(s) and structure(s). The methodology was extensively validated and has proven to be an attractive choice for the use by the practitioner due to its high computational efficiency; it is particularly attractive when an extensive parametric analysis of the system is intended, as often is the case at the pre-design stage.

Along with its significant advantages, the methodology also has some rather serious limitations. In particular, in its present form it is only applicable to very specific geometries of the structure such as cylindrical and spherical. This feature significantly limits the applicability of the methodology due to the fact that most industrial systems possess higher geometrical and/or material complexity. In an attempt to overcome this limitation, we propose a hybrid model where the structural part is handled using FEM, while the fluid domain is

modelled using the response-functions-based methodology.

We show that such an approach results in a model that is capable of accurately simulating the shock response of many commonly encountered in engineering practice structures that have a degree of geometrical and/or material complexity, while being more computationally efficient than "pure-FEM" approaches. Thus, the proposed methodology is demonstrated to combine the versatility of FEM with the computational efficiency of the response-functions-based methodology.

2 MATHEMATICAL MODEL

We consider a structure of the cylindrical outer shape, of radius r_0, and of arbitrary inner structural complexity submerged into inviscid, irrotational, and lineraly compressible fluid with density ρ_f and sound speed c_f. The transverse and normal displacements of the outer surface of the structure are v^* and w^*, respectively. The geometry of the problem is shown in Figure 1.

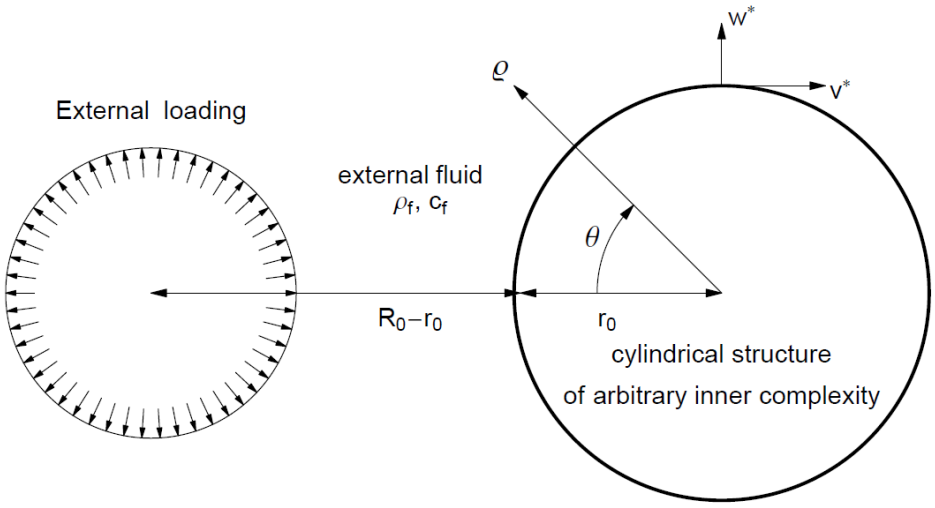


Figure 1: Geometry of the problem

We assume that the structure is subjected to a high-frequency, non-stationary external loading.

The fluid domain is modeled using the wave equation, an assumption which, combined with the simplicity of the geometry of the interface, enables the use of the classical apparatus of the mathematical physics. For the structural domain, however, we do not make any simplifying assumptions regarding its inner geometry, and model it assuming that the structure's material is linearly elastic and isotropic.

For the fluid domain, therefore, we obtain solution using the approach that we have developed in our earlier work [1-4], that is, the Laplace transform with respect to time is first applied to the wave equation, and then the separation of the spatial variables is used to yield the general solution in the form of a linear combination of two modified Bessel functions. Upon applying the boundary conditions and inverting the resulting expressions, the pressure

components are obtained in the form of Fourier series with time-dependent coefficients. For the radiated pressure (the component that plays the key role in the proposed methodology), these coefficients are (the dimensionless form of the normal displacement is used, thus the asterisk is not present):

$$p_n^r = -\int_0^t \frac{\mathrm{d}^2 w_n(\eta)}{\mathrm{d}\eta^2}\, \xi_n^e(r, t-\eta)\mathrm{d}\eta \tag{1}$$

where $\xi_n^e(r,t)$ are the response functions of the problem with the Laplace transforms given by

$$\Xi_n^e(r,s) = -\frac{\mathrm{K}_n(rs)}{s\mathrm{K}_n'(s)} \tag{2}$$

where $\mathrm{K_n}$ is the modified Bessel function of the second kind of order n, and $\mathrm{w_n}$ is the time-dependent component of the n-th harmonic of w. All the details of the methodology can be found in our earlier work [1-4].

For the structural part, the standard finite-element approach has been adopted, and we do not include any of the details here. The fluid and the structural parts are coupled at each time step, with the latter receiving the hydrodynamic loading from the former and passing back the normal displacements of the structural surface; the process is accompanied by the constant switching between the functional and modal forms of the quantities concerned.

3 RESULTS AND DISCUSSION

The ultimate objective of this study is to be able to simulate the interaction with structures that have a certain degree of geometrical complexity. But prior to attempting such simulations, we needed to validate the developed methodology. To that end, we considered a typical steel shell with the thickness-to-radius ratio of 0.03 submerged in water and subjected to a point-source acoustic pulse, and carried out two comparisons - one with the results of an experimental study [5], Figure 2, and the other with the results produced by the simulations based on a semi-analytical model [4], Figure 3, the former aimed at establishing that the methodology adequately reproduces the structure of the hydrodynamic field induced during the interaction, and the latter ensuing that the structural dynamics is also accurately represented. In both cases, a very good agreement has been observed, thus providing a rather solid evidence in favor of the developed approach. We note that the fact that the good match was seen for the two rather different aspects of the interaction is quite comforting as far as the use of the approach for modeling more complex systems is concerned.

We then considered a more structurally complex system, namely, a steel shell of radius 1 m and thickness 0.03 m, with two attached masses positioned at $\pm 45^{\mathrm{o}}$, of the radial extension of 0.03 m and angular one of 1.8^{o}, also made of steel; the aim of choosing such relatively "light" masses was to demonstrate that even a minor structural modification produces rather significant differences as far as the radiated hydrodynamic pattern is concerned. Figure 5 shows the snapshots of the radiated hydrodynamic field during the early interaction for the shell with and without the attached masses. The effect of the attached masses is apparent.

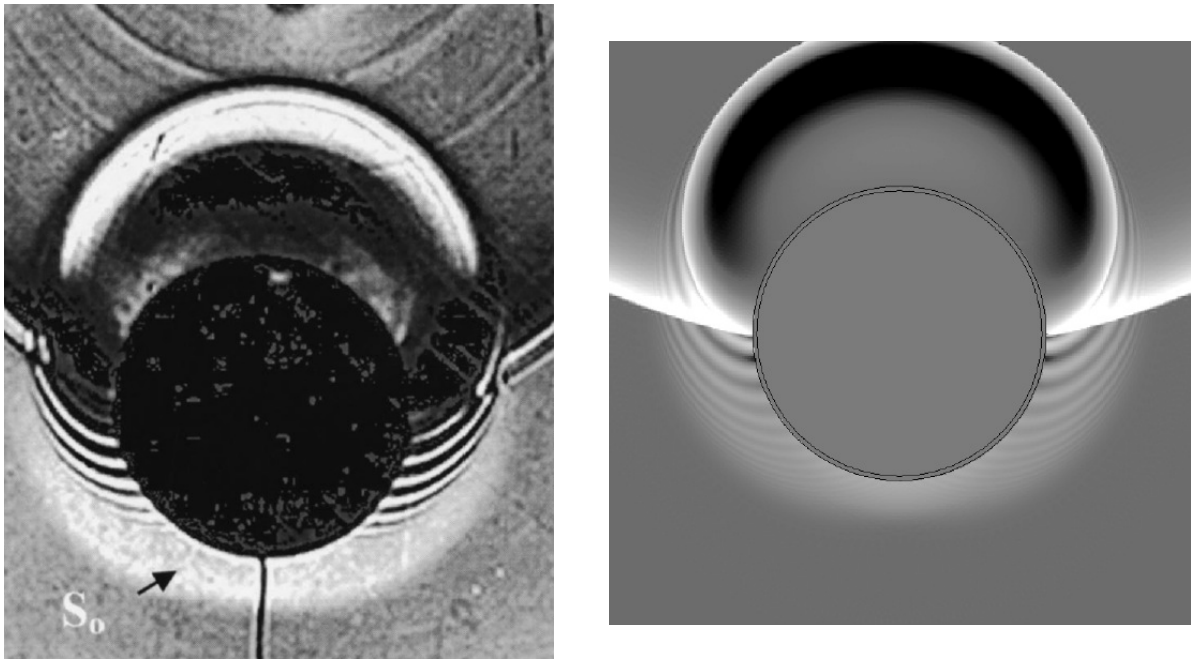


Figure 2: Numerically simulated hydrodynamic field around a steel shell with h_0/r_0=0.03, right, compared with an experimental image, left [reprinted with permission from Derbesse, L., Pernod, P., Latard, V., Merlen, A., Decultot, D., Touraine, N., and Maze, G. 2000 Acoutsic scattering from complex elastic shells: visualization of S0, A0 and A waves. Ultrasonics 38, 860-863, Figure 1 (b), © 2000 Elsevier].

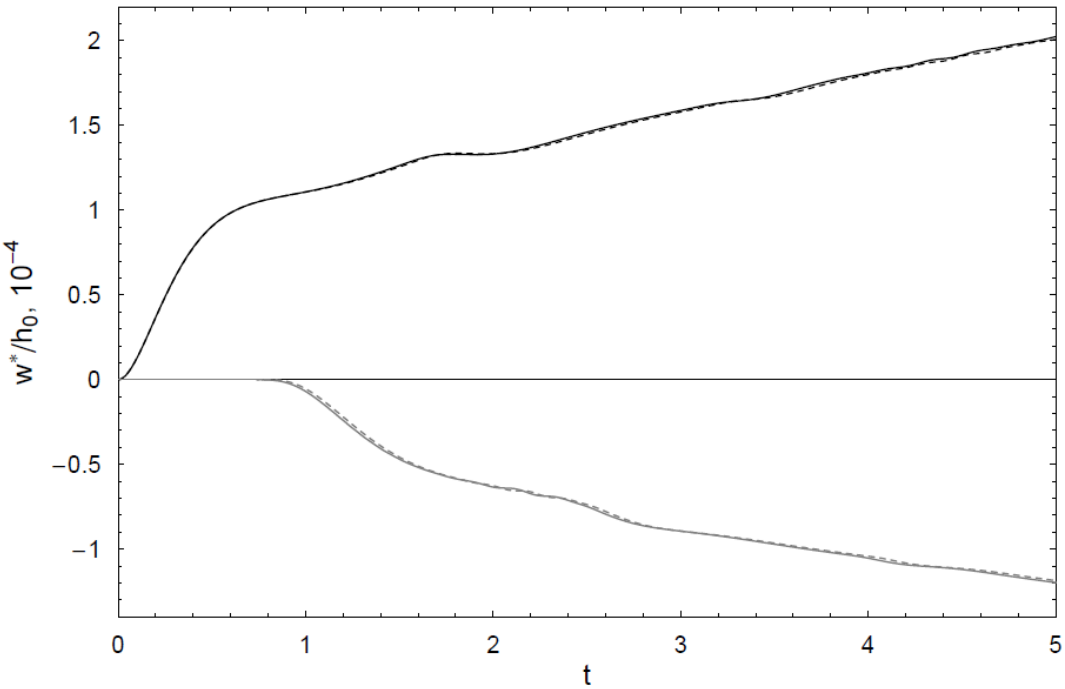


Figure 3: Normal displacements of a steel shell with h_0/r_0=0.03 computed using a semi-analytical approach based on the Reissner-Mindlin model (dashed lines) and using the present approach (solid lines) evaluated at the head point (black lines) and the tail point (grey lines).

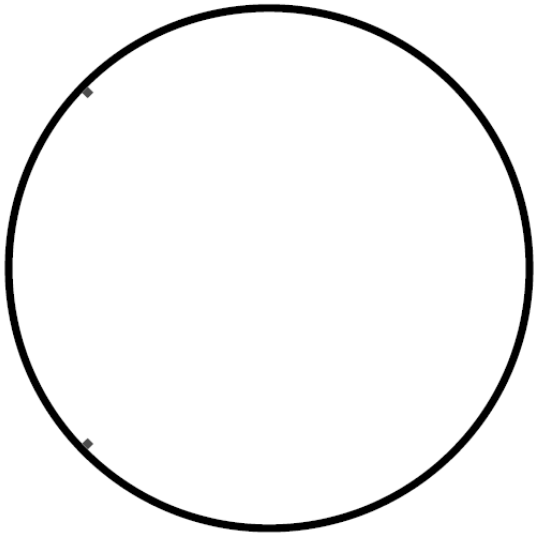


Figure 4: Shell with attached masses.

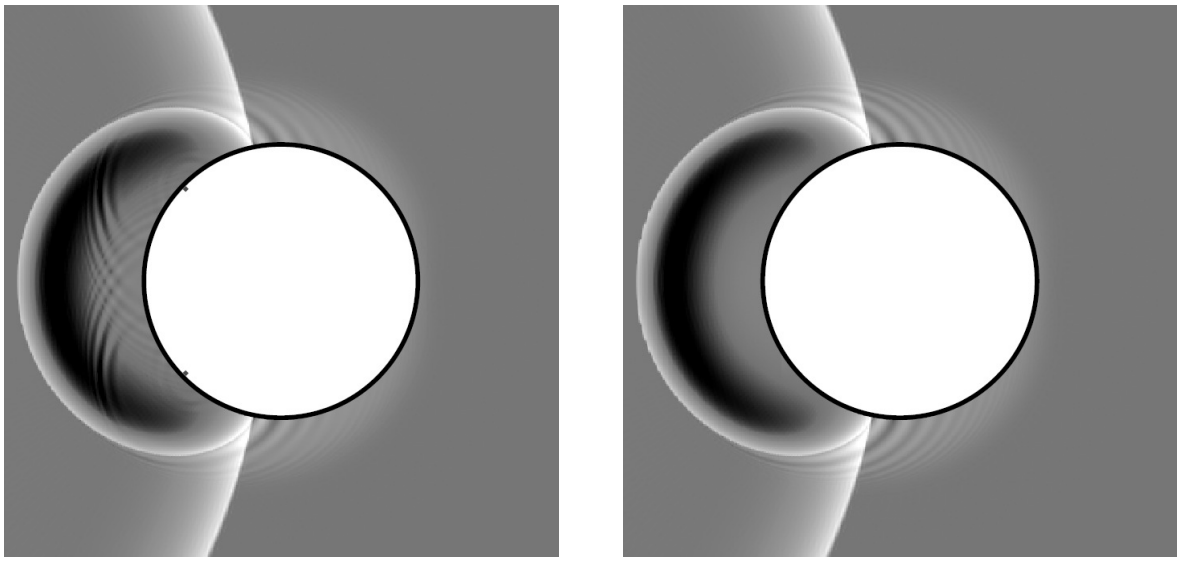


Figure 5: Numerically simulated hydrodynamic field around a steel shell with h_0/r_0=0.03 and two small attached masses during the early interaction (left), and around the same shell at the same instant but without attached masses (right).

We point out that although the hydrodynamic patterns are noticeably different, the structural dynamics is hardly affected at all, as is evident from the time-histories of the normal displacements, Figure 6. This observation highlights the importance of the focused analysis of the hydrodynamic fields carried out along with the structural analysis when a complete understanding of all aspects of the interaction is necessary.

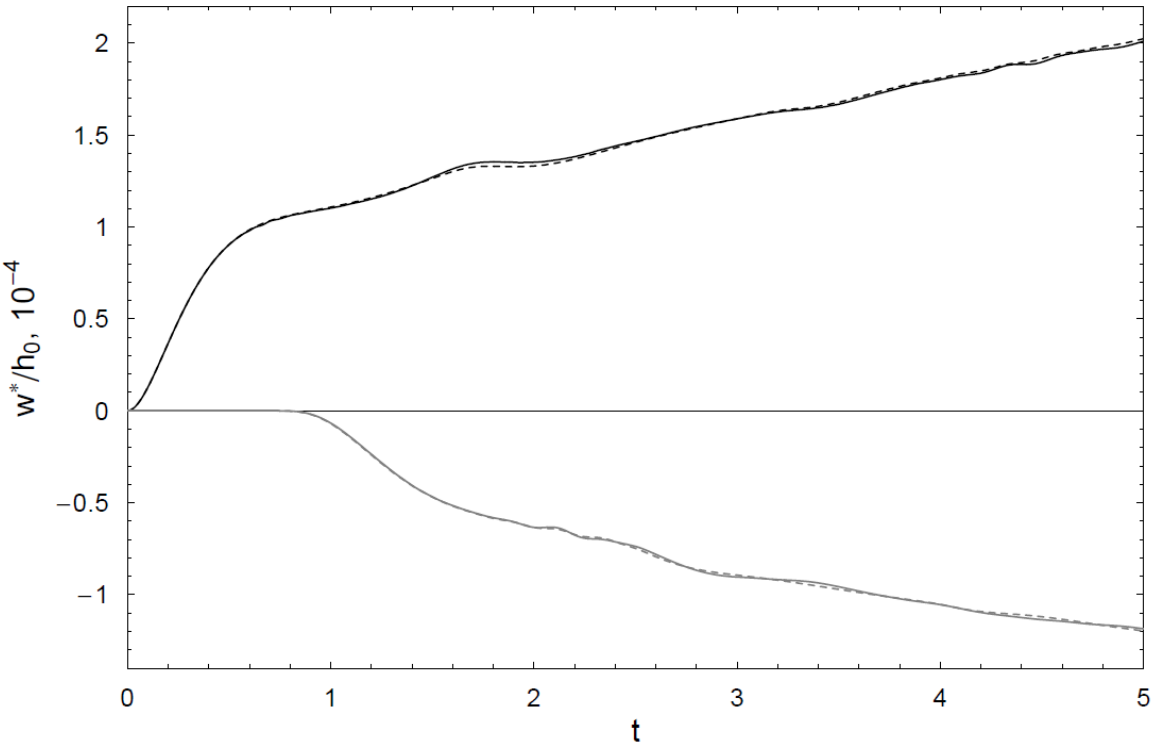


Figure 6: Normal displacements of a steel shell with h_0/r_0=0.03 and no structural enhancements (dashed lines) compared to the normal displacements of the same shell with attached masses (solid lines) evaluated at the head point (black lines) and the tail point (grey lines).

4 CONCLUSIONS

We have proposed a hybrid analytical-numerical methodology for simulating the interaction between a structure which has a cylindrical outer shape but arbitrary inner complexity. The methodology combines the analytical approach based on the classical apparatus of mathematical physics for the fluid domain with the finite-element approach for the structural domain, with coupling occurring on the interface at each time step.

The methodology has been validated using both available semi-analytical solution and experimental data, and a good agreement was observed in both cases.

We then considered a more complex system, namely, a cylindrical shell with two attached masses, and demonstrated the capabilities of the model by simulating both the fluid and structural dynamics of such a system. In particular, we showed that although the structural effect of the attached masses, as manifested in the time-histories of the normal displacements, is very insignificant, their effect on the overall hydrodynamic pattern is quite pronounced, thus highlighting the necessity of a careful consideration of all aspects of the interaction.

ACKNOWLEDGEMENTS

S.I. gratefully acknowledges the support of the Natural Sciences and Engineering Research

Council (NSERC) of Canada. The mobility grant for scholars provided by the Erasmus Mundus program in the framework of the Master of Science in Computational Mechanics (FPA 2013-0220 2014 Erasmus+) is also acknowledged.

REFERENCES

[1] Iakovlev, S., External shock loading on a submerged fluid-filled cylindrical shell, *Journal of Fluids and Structures* (2006) 22: 997-1028.

[2] Iakovlev, S., Interaction between a submerged evacuated cylindrical shell and a shock wave. Part I: Diffraction-radiation analysis, *Journal of Fluids and Structures* (2008) 24: 1077-1097.

[3] Iakovlev, S., Interaction between a submerged evacuated cylindrical shell and a shock wave. Part II: Numerical aspects of the solution, *Journal of Fluids and Structures* (2008) 24: 1098-1119.

[4] Iakovlev, S., Santos, H. A. F. A., Williston, K., Murray, R., Mitchell, M., Non-stationary radiation by a cylindrical shell: numerical modeling using the Reissner-Mindlin theory, *Journal of Fluids and Structures* (2013) 36: 50-69.

[5] Derbesse, L., Pernod, P., Latard, V., Merlen, A., Decultot, D., Touraine, N., Maze, G., Acoustic scattering from complex elastic shells: visualization of S0, A0 and A waves, *Ultrasonics* (2000) 38: 860–863.

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis(Eds)

IMMERSED BOUNDARY METHOD FOR THE SIMULATION OF LID-DRIVEN CAVITY FLOW WITH AN EMBEDDED CYLINDER

Shang-Gui CAI*, Abdellatif OUAHSINE*, Julien FAVIER, Yannick HOARAU† AND Hassan SMAOUI*,††**

* Sorbonne Universités, Université de Technologie de Compiègne, CNRS, Laboratoire Roberval UMR 7337, Centre de Recherches de Royallieu, CS 60319, 60203 Compiègne Cedex, France
e-mail: shanggui.cai@utc.fr, ouahsine@utc.fr

** Aix Marseille Université, CNRS, M2P2 UMR 7340, 13451 Marseille, Frace

† Université de Strasbourg, CNRS, ICUBE UMR 7357, 67000 Strasbourg, France

††Direction Technique Eau, Mer et Fleuves (CEREMA)
134, Rue de Beauvais, CS 60039, 60280 Margny-lès-Compiègne, France

Key words: Immersed Boundary Method, Projection Method, Lid-driven cavity flow, Embedded Cylinder

Abstract. We present the numerical simulation of lid-driven cavity flow with an embedded cylinder by using a modified implicit immersed boundary method. The immersed boundary method features the non body-conforming mesh merit, which facilitates the simulation with complex geometry in a flow. The paper compares the numerical solutions between the modified implicit immersed boundary method with the traditional body-conforming mesh method as well as the results in the literature, which shows an excellent agreement.

1 INTRODUCTION

Immersed boundary (IB) method has emerged in recent years as an alternative to the conventional body-conforming mesh technique [1, 2], demonstrating its advantage to handle complex rigid or flexible boundaries embedded in a flow [3]. Although meshless methods have a successful applications [4], they have difficulties of their own. The establishment of a well body fitted Cartesian mesh generally possess most of the CFD work. IB method significantly simplifies this procedure and speedup the solution by eliminating the re-meshing or mesh deforming process. Through adopting an appropriate boundary force for the presence of immersed structure or solid, the simulation is simply performed

on a regular Cartesian mesh which can be accomplished by highly efficient solvers. A recent review of the immersed boundary method can be found in [5, 6].

Ever since introduced by Peskin [7] for the simulation of blood flow in a heart with flexible valves, IB method has been extended to a great variety of applications, including civil engineering, aeronautics, hydrodynamics, etc. The original IB method is categorized as "continuous forcing"in [5] as it is implemented into the Navier-Stokes equations before discretizated. However, when applied to rigid boundary problems, the original IB method appeals to a spring with an extremely large stiffness to pull the boundary back to its equilibrium position [8], which in general gives arise to numerical difficulties. To overcome this shortness, Mohd-Yusof [9] proposed the direct forcing IB method using a discrete-time derivation of the forcing value to eliminate the spring, which falls in the "discrete forcing"group. Later Uhlamnn [10] suggested evaluating the force on the Lagrangian coordinates instead of the Eulerian location, which is much simpler to perform. Nonetheless, those IB methods are all based on an explicit scheme, the no-slip boundary condition on the interface is never exactly satisfied even for the steady flow [11]. However full implicit scheme is too cumbersome to perform, we need to iterate the whole system to ensure the incompressibility and no-slip boundary condition on the interface in the same time. In the present study we adopt the modified implicit IB method in [12] for accuracy and efficiency.

Almost all IB methods are able to deal with elastic and rigid boundaries. However, in the present work we only focus on the direct forcing IB method with rigid boundaries. Actually the simulation of lid-driven cavity flow with an embedded rigid cylinder is not a difficult task for the traditional body-conforming mesh method. One can mesh the domain with one O-type structured mesh near the cylinder. However the purpose of this paper is mainly to compare the numerical accuracy between the IB method and the traditional one.

In the next section, we discuss the governing equations and its numerical solution by the rotational incremental pressure correction projection method. In the following the implementation of the modified implicit IB method into the projection method is recalled. Then numerical results of lid-driven cavity flow with an embedded cylinder are compared between IB method and traditional method as well as the result in the literature. Conclusions are drawn in the final section.

2 METHODOLOGY

2.1 Governing equations and projection method

The unsteady viscous incompressible Navier-Stokes equations read

$$\frac{\partial u_i}{\partial x_i} = 0 \tag{1}$$

$$\frac{\partial u_i}{\partial t} + \frac{\partial (u_i u_j)}{\partial x_j} = -\frac{1}{\rho}\frac{\partial p}{\partial x_i} + \nu \frac{\partial}{\partial x_j}(\frac{\partial u_i}{\partial x_j}) \tag{2}$$

where u_i are the velocity components, p is the pressure, ρ is the fluid density and ν is the fluid kinematic viscosity. The equations are integrated in time using the rotational incremental pressure correction projection method of Guermond et al. [13] which is proved to be second order in time for the velocity fields and has a consistent boundary condition for the pressure. The projection method consists of three sub-steps: prediction, projection and correction.

1. Predict an intermediate velocity u_i^* with the pressure value at past time level.

$$\frac{u_i^* - u_i^n}{\Delta t} + \frac{1}{2}[3C_i^n - C_i^{n-1}] = -\frac{1}{\rho}\frac{\delta p^n}{\delta x_i} + \frac{\nu}{2}(D_i^* + D_i^n) \tag{3}$$

where $C_i = \frac{\delta(u_i u_j)}{\delta x_j}$ and $D_i = \nu\frac{\delta}{\delta x_j}(\frac{\delta u_i}{\delta x_j})$ represent the convective and diffusive terms, respectively. The convective term is treated explicitly in time for simplicity by using a second order Adams-Bashforth scheme, and the diffusive term is discretized with an implicit Crank-Nicolson scheme to eliminate the viscous stability constraint. Therefore, the total scheme is only subjected to the standard CFL condition. The diffusion term is spatially discretized by a second-order central difference operator while the convective term is treated with the TVD scheme described in [14].

2. Project the tentative velocity u_i^* to its divergence-free part u_i^{n+1}, by solving the pressure Poisson equation in conjunction with the incompressibility constraint.

$$\frac{u_i^{n+1} - u_i^*}{\Delta t} = -\frac{1}{\rho}\frac{\delta \phi^{n+1}}{\delta x_i} \tag{4}$$

Taking divergence of Eq. (4) and applying Eq. (1), we obtain

$$\frac{1}{\rho}\frac{\delta}{\delta x_i}(\frac{\delta \phi^{n+1}}{\delta x_i}) = \frac{1}{\Delta t}\frac{\delta u_i^*}{\delta x_i} \tag{5}$$

3. Correct or update the velocity and pressure field. Once the pressure correction ϕ^{n+1} is known, the final pressure and velocity can be updated as

$$u_i^{n+1} = u_i^* - \Delta t\frac{1}{\rho}\frac{\delta \phi^{n+1}}{\delta x_i} \tag{6}$$

$$p^{n+1} = p^n + \phi^{n+1} - \frac{\mu}{2}\frac{\delta u_i^*}{\delta x_i} \tag{7}$$

where μ is the fluid dynamic viscosity. In Eq. (7) the last term is actually the missing term due to time splitting when implicit scheme is applied to the viscous term in Eq. (3), which is now absorbed into the pressure.

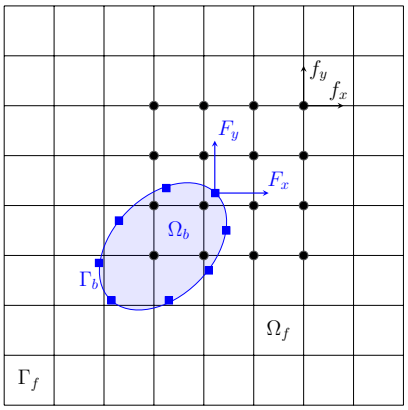


Figure 1: Immersed boundary configuration.

2.2 Immersed boundary method implementation

The principal idea of immersed boundary method is to represent the structure or solid boundary immersed in a fluid by an artificial force. The force is then added to the momentum equation as a source term to take effect. In fact, how the force is formulated determines each kind of IB method. In the present work, we choose the modified implicit IB method in [12], which can be briefed as

1. After the prediction step Eq. (3), interpolate the tentative velocity from the relying Cartesian mesh (see Fig. 1) to the structure or solid boundary Γ_b: $U_i^* = T(u_i^*)$, where upper case letter represents the quantities on the Lagrangian locations and T is the field transfer operator.

2. Evaluate the force on the immersed boundary: $F_i = (U_i^d - U_i^*)/\Delta t$, where U_i^d is the desired velocity boundary condition we want to impose on Γ_b. Here i is the vector component index ($i = 1, 2$ in 2D, $i = 1, 2, 3$ in 3D) and not a specific Lagrangian point.

3. Spread the force to the relying mesh: $f_i^* = S(F_i^*)$, where S is the spreading operator. In Fig. 1 the "•"donates the neighbouring points associated with the Lagrangian variable. The number of points depend on the support width of discrete delta function. In the present work, Peskin's 4-point version is employed [8].

$$\delta_h(r) = \begin{cases} \frac{1}{8}(3 - 2|r| + \sqrt{1 + 4|r| - 4r^2}), & |r| \leqslant 1, \\ \frac{1}{8}(5 - 2|r| - \sqrt{-7 + 12|r| - 4r^2}), & 1 \leqslant |r| \leqslant 2, \\ 0, & \text{otherwise.} \end{cases} \tag{8}$$

4. Add the force to Eq. (4) to mimic the presence of structure or solid.

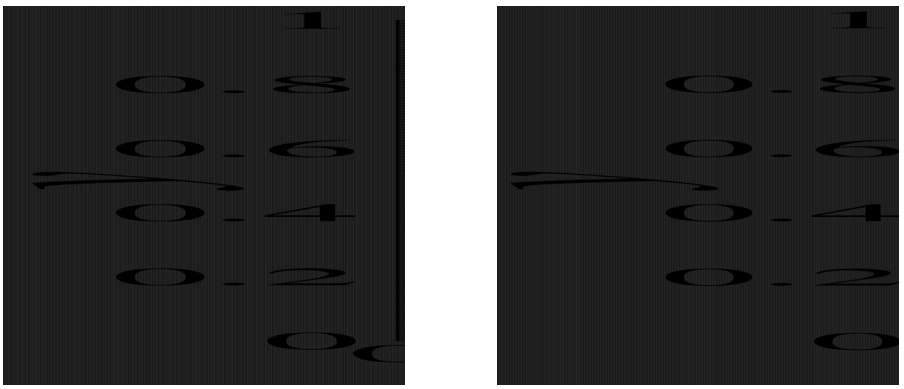
5. Solve the pressure correction Eq. (5) with the artificial force and update the velocity Eq. (6).

Repeat 1-5 until the desired velocity boundary condition is satisfied on these Lagrangian

points. Noting that the pressure correction equation is solved only once at each iteration for efficiency, we propose to perform Eq. (5) after the forcing loop to obtain better convergence and then update the velocity and pressure with Eq. (6), (7). In the present paper, the Krylov iterate solver of conjugate gradient along with the multigrid preconditioner is applied to solving the pressure Poisson equation.

3 NUMERICAL RESULTS

The flow of lid-driven cavity with an embedded cylinder is investigated in the present work to assess the accuracy of the improved implicit immersed boundary method. The length and height of cavity are unity ($L = 1$). The top wall is moving with a constant velocity $u_\infty = 1$ while the other three are stationary no slip walls. The cylinder is fixed at the center of the cavity with a diameter $D = 0.4L$, as shown in Fig. 2a for the IB method with uniform mesh arrangement. For comparison, a body-conforming mesh is applied in Fig. 2b. A fine mesh is used here, where the ratio of cylinder diameter to the mesh size is $D/h = 200$ in both cases. The flow pattern is characterized by the Reynolds number $Re = u_\infty L/\nu$, which is set to 1000 in the this study.



(a) (b)

Figure 2: (a) Uniform mesh arrangement for the IB method (the embedded cylinder is represented by the thick dashed line), (b) body-conforming mesh arrangement.

The convergence criteria is set to $||u_i^{n+1} - u_i^n||_2/||u_i^{n+1}||_2 < 1.0E-8$ so that the flow reaches the steady state. Fig. 3 shows the vorticity contour and streamline of the flow at $Re = 1000$ for both IB method and traditional body-conforming mesh method, where the vorticity contour value is varied from -3 to 3 with an increment of 0.4. Fig. 3a, 3c correspond to the IB method for the vorticity contour and streamline, while Fig. 3b, 3d are the results from the traditional method for vorticity and streamline respectively. The numerical simulation presented in this study is very similar to the results of Vanella and Balaras [15]. Apparently, three vortices emerge: one at the upper right near the cylinder and two lower at the left and right corners. It is noteworthy that the upper vortex is

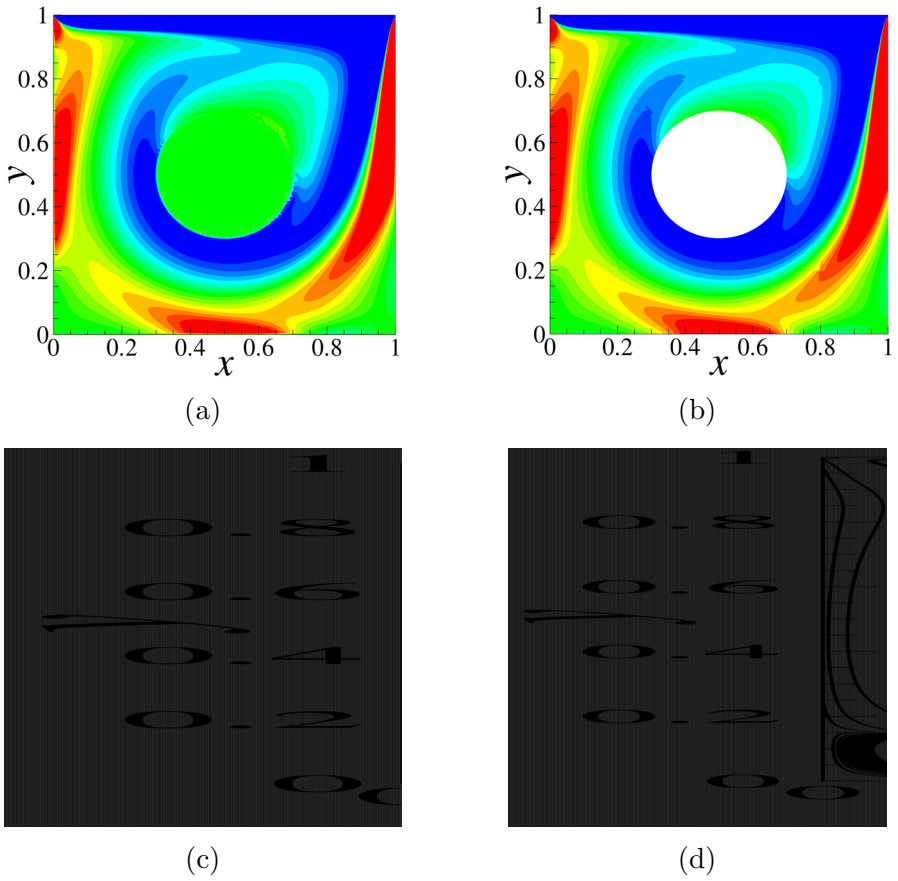


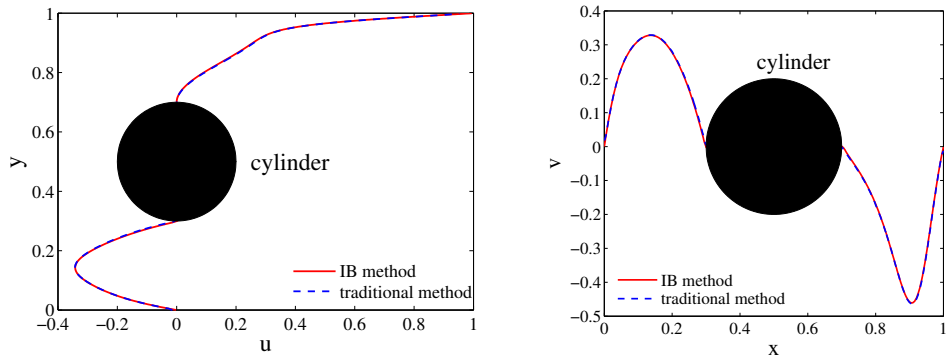
Figure 3: Vorticity contour and streamline of lid-driven cavity flow with an embedded cylinder.

generated by the presence of the fixed cylinder. The locations of three vortices center of both methods are listed in Table 1. The two methods yield nearly the same results.

Table 1: Comparison of vortices center positions

	upper right vortex	lower left vortex	lower right vortex
IB method	(0.6946, 0.6872)	(0.0790, 0.0721)	(0.8856, 0.1061)
Traditional method	(0.6906, 0.6872)	(0.0791, 0.0721)	(0.8849, 0.1063)

The velocity component u at the vertical midline and the velocity component v at the horizontal midline are plotted in Fig. 4. Since the modified implicit IB method imposes the no-slip boundary condition on the interface exactly, the velocity profiles of both methods match very well.



(a) (b)

Figure 4: Comparison of velocity profiles (a) distribution of velocity component u along vertical midline, (b) distribution of velocity component v along horizontal midline.

4 SUMMARY

We have presented the numerical simulation of the lid-driven cavity flow with an embedded cylinder by the improved implicit immersed boundary method. Comparison has been made between the IB method, traditional body-conforming mesh method and the results in the literature. This validation case has demonstrated good agreements for both methods. Even though the fixed cylinder in the present work is not complex, the IB method can be easily extended to the flows with more complicated rigid or deforming, and/or moving geometries. In particular for the fluid-structure interaction problems with multi-structure in artificial shapes, the traditional body-conforming mesh method becomes extremely difficult and time-consuming to ensure the mesh quality especially after re-meshing. This task with IB method will be included in the future work.

ACKNOWLEDGMENT

The first author acknowledges the financial support of the China Scholarship Council. The calculation is performed on the HPC platform PILCAM2 at the Université de Technologie de Compiègne.

REFERENCES

[1] Ji, S.C., Ouahsine, A., Smaoui and H., Sergent, P. 3-D numerical simulation of convoy-generated waves in a restricted waterway. *J. Hydrodyn. Ser. B* (2012) **24**:420–429.

[2] Kaidi, S., Rouainia, M. and Ouahsine, A. Stability of breakwaters under hydrodynamic loading using a coupled DDA/FEM approach. *Ocean Eng.* (2012) **55**:62–70.

[3] Favier, J., Revell, A. and Pinelli, A. A lattice-Boltzmann-immersed boundary method to simulate the fluid interaction with moving and slender flexible objects. *J. Comput. Phys.* (2014) **261**:145–161.

[4] Darbani, M., Ouahsine, A., Villon, P., Naceur, H. and Smaoui, H. Meshless method for shallow water equations with free surface flow. *Appl.Math.Comput.* (2011) **217**:5113–5124.

[5] Mittal, R. and Iaccarino, G. Immersed boundary method. *Annu. Rev. Fluid Mech.* (2005) **37**:239–261.

[6] Sotiropoulos, F. and Yang, X. Immersed boundary methods for simulating fluid-structure interaction. *Prog. Aerosp. Sci.* (2014) **65**:1–21.

[7] Peskin, C.S. Flow patterns around heart valves: A numerical method. *J. Comput. Phys.* (1972) **10**:252–271.

[8] Lai, M.-C. and Peskin, C.S. An immersed boundary method with formal second-order accuracy and reduced numerical viscosity. *J. Comput. Phys.* (2000) **160**:705–719.

[9] Mohd-Yosuf, J. Combined immersed boundary/b-spline methods for simulation of flow in complex geometries. *Annu. Res. Briefs, Cent. Turbul. Res.* (1997)

[10] Uhlmann, M. An immersed boundary method with direct forcing for the simulation of particulate flows. *J. Comput. Phys.* (2005) **209**:448–476.

[11] Kempe, T. and Fröhlich, J. An improved immersed boundary method with direct forcing for the simulation of particle laden flows. *J. Comput. Phys.* (2012) **231**:3663–3684.

[12] Cai, S.-G., Ouahsine, A., Smaoui, H., Favier, J. and Hoarau, Y. An efficient implicit direct forcing immersed boundary method for incompressible flows. *Proc. of the 3rd International Conference on Mathematical Modelling in Physical Sciences (IC-MSQUARE).* Madrid, Spain. (2014) **574**:012165.

[13] Guermond, J.L., Minev, P. and Jie, S. An overview of projection methods for incompressible flows. *Comput. Methods Appl. Mech. Engrg.* (2006) **195**:6011–6045.

[14] Smaoui, H., Zouhri, L. and Ouahsine, A. Flux-limiting techniques for simulation of pollutant transport in porous media: Application to groundwater management. *Math. Comput. Model.* (2008) **47**:47–59.

[15] Vanella, M. and Balaras, E. A moving-least-squares reconstruction for embedded-boundary formulations. *J. Comput. Phys.* (2009) **228**:6617–6628.

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis(Eds)

NUMERICAL MODELLING OF FLUID-STRUCTURE INTERACTIONS

A. FORTIN*, A. JENDOUBI* AND J. DETEIX*

*GIREF
Département de mathématiques et de statistique
1045 Avenue de la médecine, Université Laval, Québec, Canada, G1V 0A6.
e-mail: afortin@giref.ulaval.ca,
web page: http://http://www.giref.ulaval.ca/accueil.html

Key words: ALE formulation, Moving mesh, Fluid-Structure interactions.

Abstract. Fluid-structure interaction (FSI) problems occur when a potentially deformable solid interact with a surrounding fluid. The flow of the fluid deforms the solid and/or changes its position thus modifying the geometry of the fluid domain. In this paper, we present an Arbitrary Lagrangian Eulerian (ALE) formulation of FSI problems and we present a few numerical examples.

1 INTRODUCTION

Fluid-structure interaction problems occur in various engineering applications and their numerical simulations is still challenging nowadays. Many difficulties arise and we will briefly present some of them in this paper. We consider the general case where the flow of the fluid deforms the solid and changes its position thus modifying the geometry of the fluid domain. Arbitrary Lagrangian Eulerian (ALE) formulations (see [1]) are frequently used to solve such problems where the solid and the fluid are respectively in Lagrangian and Eulerian coordinates. One of the major difficulties in ALE description of fluid-structure interactions is, knowing the displacement of the solid domain, to determine how to update the fluid mesh. If care is not taken, elements adjacent to the solid boundary can degenerate leading to computational problems and eventually to the divergence of the solution process. In this paper, we compare moving mesh methods such as kriging, Laplacian smoothing and displacement methods based on the solution of an elasticity problem. We consider a number of test problems and comment their respective performances.

2 THE COUPLED FSI PROBLEM

2.1 Governing equations

In order to introduce the general non-linear fluid-structure problem, let us consider a time-dependent domain $\Omega(t) \subset \mathbb{R}^d$, $d = 2$ or 3. We assume, for all time t, that $\Omega(t) = \Omega_f(t) \cup \Omega_s(t)$ and $\Omega_f(t) \cap \Omega_s(t) = \emptyset$, where $\Omega_f(t)$ is occupied by an incompressible viscous fluid and $\Omega_f(t)$ by an elastic solid. The reference (initial) configuration of the system is defined $\Omega_0 = \Omega_{f0} \cup \Omega_{s0}$. The fluid-structure interface at time t is denoted by $\Sigma(t) = \overline{\Omega_f(t)} \cap \overline{\Omega_s(t)}$ and $\Sigma_0 = \overline{\Omega_{f0}} \cap \overline{\Omega_{s0}}$ on the initial configuration.

For the fluid, since we are dealing with a moving interface, we consider the incompressible Navier-Stokes equations in ALE formulation (see [2]) :

$$\begin{cases} \rho_f \left(\dfrac{\partial \boldsymbol{v}_f}{\partial t} + (\boldsymbol{v}_f - \boldsymbol{v}_{s\mapsto f}) \cdot \nabla \boldsymbol{v}_f \right) + \nabla p - \nabla \cdot \big(2\eta \dot{\gamma}(\boldsymbol{v}_f)\big) &= \boldsymbol{r}_f \ \text{ in } \ \Omega_f(t) \\ \nabla \cdot \boldsymbol{v}_f &= 0 \ \text{ in } \ \Omega_f(t) \\ \boldsymbol{v}_f(\boldsymbol{x}, t) &= \boldsymbol{v}_s(\boldsymbol{x}, t) \ \text{ on } \ \Sigma(t), \end{cases} \tag{1}$$

where $\boldsymbol{v}_f(x, t)$ is the Eulerian velocity field of the fluid. The vector $\boldsymbol{v}_{s\mapsto f}(x, t)$ is the Eulerian domain (mesh) velocity that will be described later on. As we will see, the domain velocity is naturally defined in the solid domain and must somehow be extended to the fluid domain thus the subscript $s \to f$. Note that in general $\boldsymbol{v}_{s\mapsto f}(\boldsymbol{x}, t) \neq \boldsymbol{v}_f(\boldsymbol{x}, t)$. The continuity of the fluid and solid velocities is imposed on the interface $\Sigma(t)$. The system is completed by initial and boundary conditions on the other parts of the boundary $\Gamma_f^D(t)$ (Dirichlet) and $\Gamma_f^N(t)$ (Neumann). The complete problem is thus defined (and solved) on the current geometry $\Omega_f(t)$.

The motion of the structure is described by the following equations:

$$\begin{cases} \rho_{s0} \dot{\boldsymbol{V}}_s - \nabla \cdot \big(\boldsymbol{\Pi}_s(\boldsymbol{U}_s)\big) &= \rho_{s0} \boldsymbol{r}_s \ \text{ in } \ \Omega_{s0} \\ \dot{\boldsymbol{U}}_s &= \boldsymbol{V}_s \quad \text{ in } \ \Omega_{s0} \end{cases} \tag{2}$$

where $\boldsymbol{U}_s(\boldsymbol{X}, t)$ and $\boldsymbol{V}_s(\boldsymbol{X}, t)$ are respectively the Lagrangian displacement and velocity fields of the solid, ρ_{s0} its density in the undeformed geometry and $\boldsymbol{r}_s$ a given body force (usually vanishing). The problem is thus written in a total Lagrangian formulation on Ω_{s0}.

The tensor $\boldsymbol{\Pi}_s = \boldsymbol{\Pi}_s(\boldsymbol{U}_s)$ is the first Piola-Kirchhoff tensor (see Bonet [3]) and is related to the Cauchy stress tensor $\boldsymbol{\sigma}_s$ by the relation:

$$\boldsymbol{\Pi}_s = J_s \boldsymbol{\sigma}_s \cdot \boldsymbol{F}_s^{-T} \tag{3}$$

where $\boldsymbol{F}_s$ is the gradient of deformation tensor and J_s is the Jacobian of the transformation from the initial geometry Ω_{s0} to the current configuration $\Omega_s(t)$. These quantities come

out naturally from the solution of problem (2). This problem is also completed with a proper set of initial and boundary conditions (Dirichlet on Γ_s^D and Neumann on Γ_s^N).

From a mechanical point of view, the coupling between the systems (1) and (2) is realized by imposing the equilibrium of the stresses at the interface:

$$\boldsymbol{\sigma}_s \cdot \boldsymbol{n}_s \;=\; -\boldsymbol{\sigma}_f \cdot \boldsymbol{n}_f \quad \text{on } \Sigma(t),$$

The Cauchy tensor $\boldsymbol{\sigma}_f$ comes out of the solution of system (1) and is defined on the current configuration. System (2) is however solved on the reference configuration. $\boldsymbol{\sigma}_f$ must therefore be transferred on the reference geometry. It is easily seen from (3) and from the transformation of the normal vectors that the above equilibrium condition on the deformes geometry becomes:

$$\boldsymbol{\Pi}_s \cdot \boldsymbol{N}_s = -J_{s\mapsto f}\boldsymbol{\sigma}_f \cdot \boldsymbol{F}_{s\mapsto f}^{-T} \cdot \boldsymbol{N}_f \quad \text{on } \Sigma_0. \tag{4}$$

on the reference geometry. In the above equation, $\boldsymbol{N}$ and $\boldsymbol{n}$ stand for the normal vectors to the initial and deformed geometries respectively. The equilibrium condition (4) is added to system (2) as a Neumann boundary condition. Note that $J_{s\mapsto f}$ and $\boldsymbol{F}_{s\mapsto f}$ are extensions to Ω_{f0} of similar quantities already defined in Ω_{s0}. Their computation will require the definition of a Lagrangian displacement in Ω_{f0}. In practice, any reasonable extension (denoted $\boldsymbol{U}_{s\mapsto f}$) of $\boldsymbol{U}_s|_{\Sigma_0}$ over Ω_{f0} can be used and we will see in Section 3 a number of ways to achieve that goal. It is however quite natural to impose the continuity of the displacements at the interface:

$$\boldsymbol{U}_{s\mapsto f} = \boldsymbol{U}_s \quad \text{on } \Sigma_0 \tag{5}$$

The Lagrangian mesh velocity $\boldsymbol{V}_{s\mapsto f}$ in the fluid domain can then be easily obtained from a finite difference in time. For instance, if an Euler scheme is used, then:

$$\boldsymbol{V}_{s\mapsto f}(\boldsymbol{X},t) = \frac{\boldsymbol{U}_{s\mapsto f}(\boldsymbol{X},t+\Delta t) - \boldsymbol{U}_{s\mapsto f}(\boldsymbol{X},t)}{\Delta t} \tag{6}$$

which, together with (5), imposes the continuity of the velocities on the interface.

We therefore suppose that we have defined an extension $\boldsymbol{U}_{s\mapsto f}$ of the mesh displacement in Ω_{f0} from which we can evaluate $J_{s\mapsto f}$ and $\boldsymbol{F}_{s\mapsto f}$. As a consequence, the current configuration of the fluid domain, $\Omega_f(t)$, is parametrized by:

$$\Omega_f(t) = \Omega_{f0} + \boldsymbol{U}_{s\mapsto f}(\Omega_{f0}) \tag{7}$$

In other terms, each node $\boldsymbol{X}$ of the initial fluid domain has a corresponding node $\boldsymbol{x}$ in the deformed configuration satisfying:

$$\boldsymbol{x} = \boldsymbol{X} + \boldsymbol{U}_{s\mapsto f}(\boldsymbol{X},t)$$

In this way, we can follow each node of the fluid mesh. Consequently, the Eulerian domain velocity $\boldsymbol{v}_{s\mapsto f}$ at node $\boldsymbol{x}$ satisfies:

$$\boldsymbol{v}_{s\mapsto f}(\boldsymbol{x},t) = \boldsymbol{V}_{s\mapsto f}(\boldsymbol{X},t) \tag{8}$$

and will be used in System (1).

3 MESH-UPDATE PROCEDURES

We propose in this section different constructions of the extension $\boldsymbol{U}_{s\mapsto f}$. The main difficulty is to avoid the occurrence of overly distorted elements. Once $\boldsymbol{U}_{s\mapsto f}$ has been constructed, quantities such as $J_{s\mapsto f}$ and $\boldsymbol{F}_{s\mapsto f}$ are easily computed. The mesh velocity is then obtained from (6) and (8). The update of the mesh is thus a crucial step.

3.1 Laplacian smoothing

Laplacian smoothing is a classical method for updating meshes (see [4]). The following problem is solved separately for each component of the displacement.

$$\left\{\begin{array}{rcll} -\Delta \boldsymbol{U}_{s\mapsto f} & = & 0 & \text{in } \Omega_{f0}, \\ \boldsymbol{U}_{s\mapsto f} & = & \boldsymbol{U}_s & \text{on } \Sigma_0, \\ \boldsymbol{U}_{s\mapsto f} & = & 0 & \text{on } \partial\Omega_{f0}\backslash\Sigma_0, \end{array}\right. \tag{9}$$

3.2 Elasticity model

Another possibility (see [5]) is to solve the small deformation elasticity system:

$$\left\{\begin{array}{rcll} -\nabla\cdot\left[2\mu\boldsymbol{\gamma}(\boldsymbol{U}_{s\mapsto f}) - \dfrac{2\mu}{3}(\nabla\cdot\boldsymbol{U}_{s\mapsto f})\boldsymbol{I}\right] + \nabla p & = & 0 & \\ -\dfrac{1}{k}p - \nabla\cdot\boldsymbol{U}_{s\mapsto f} & = & 0 & \\ \boldsymbol{U}_{s\mapsto f} & = & \boldsymbol{U}_s & \text{on } \Sigma_0, \\ \boldsymbol{U}_{s\mapsto f}\cdot\boldsymbol{n} & = & 0 & \text{on } \partial\Omega_{f0}\backslash\Sigma_0, \end{array}\right. \tag{10}$$

where λ and k (bulk modulus) are obtained from the Poisson coefficient ν and Young's modulus E as

$$k = \frac{E}{3(1-2\nu)} \text{ and } \mu = \frac{E}{2(1+\nu)}$$

This problem cannot be solved component per component. It is thus more expensive that the Laplacian smoothing. The value of the Poisson coefficient ν can be chosen close to 0.5 in order to preserve the area (volume) of the elements.

Note also that the last boundary condition allows slip on the boundary nodes which, in some instances, prevents the distortion of elements close to the boundary. It can also be imposed for the Laplacian smoothing if one is ready to solve all the components at the same time.

3.3 Kriging method

Kriging is an interpolation method first introduced by Krige [6] in geostatistics. The displacement is defined as

$$\boldsymbol{U}_{s\mapsto f}(\boldsymbol{X},t)=\sum_{j=1}^{n}\alpha_j g(|\boldsymbol{X}-\boldsymbol{X}_j|)+a_0+a_1X_1+a_2X_2+a_3X_3$$

where the coefficients α_j and a_j have to be determined so that

$$\boldsymbol{U}_{s\mapsto f}(\boldsymbol{X}_j,t)=\boldsymbol{U}_s(\boldsymbol{X}_j,t)\quad j=1,2,\cdots,n$$

and to satisfy non-bias conditions (see also Matheron [7]). The function g must be carefully chosen and in this work,

$$g(h)=\begin{cases} h^2\ln h & \text{for 2D problems}\\ h & \text{for 3D problems}\end{cases}$$

corresponding to a thin shell approximation. It is easily seen that kriging requires the solution of a dense matrix which is still reasonable for two-dimensional applications but quickly becomes prohibitive for three-dimensional ones. More details can be found in Olivier [8]. This approach is also known as the radial basis function interpolation method (see [9]).

3.4 Parabolic smoothing

We propose a slightly modified version of the Laplacian smoothing.

$$\begin{cases} k_0\boldsymbol{U}_{s\mapsto f}-\Delta\boldsymbol{U}_{s\mapsto f} &=& 0 & \text{in } \Omega_{f0},\\ \boldsymbol{U}_{s\mapsto f} &=& \boldsymbol{U}_s & \text{on } \Sigma_0,\\ \boldsymbol{U}_{s\mapsto f} &=& 0 & \text{on } \partial\Omega_{f0}\backslash\Sigma_0,\end{cases}\tag{11}$$

where k_0 is a constant coefficient chosen large ($k_0\simeq 500$). In the various two and three-dimensional applications we have considered so far, this is the most robust and the most effective method. Here again, it can be computed for each component separately and very efficient iterative methods exist for this kind of problems so that its solution is not a major issue in terms of computational cost.

4 SOLUTION STRATEGY

The following algorithm is used to solve the global FSI problem. For each time step:

1. Solve the Navier-Stokes equations (1) to obtain the velocity $\boldsymbol{v}_f$, the pressure and the Cauchy stress tensor $\boldsymbol{\sigma}_f$. A mixed formulation (velocity-pressure) is used with a quadratic ($O(h^2)$) Taylor-Hood P_2-P_1 element [10]. The system is solved using a very efficient projection method described in [11].

2. Computation of the right-hand side of (4) and its reinterpolation on the solid side of the interface Σ_0. Note that the meshes in the fluid and solid domains may be non conforming on Σ_0.

3. Solve system (2) to get the solid displacement $\boldsymbol{U}_s$ and velocity $\boldsymbol{V}_s$. A quadratic P_2 element is used for that purpose. Note that a mixed formulation (displacement-pressure) can also be used for incompressible material.

4. Computation of the mesh displacement $\boldsymbol{U}_{s\mapsto f}$ of the fluid domain from the solid displacement field $\boldsymbol{U}_s$ using one of the methods described in Section 3;

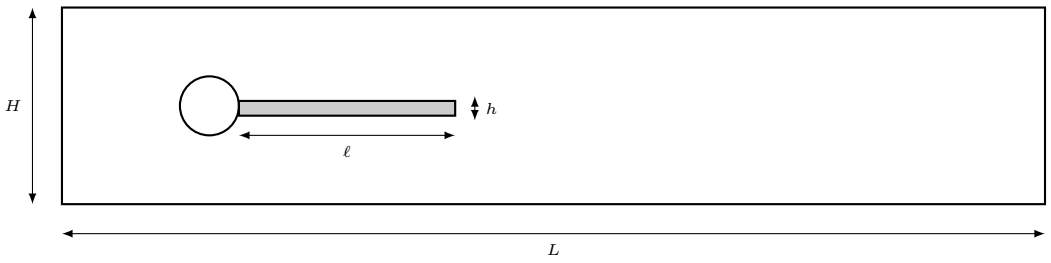
5. Update the fluid domain using (7).

6. Compute the mesh velocity $\boldsymbol{V}_{s\mapsto f}$ from equation (6) from which we deduce $\boldsymbol{v}_{s\mapsto f}$.

7. Go back to step 1. if the solution is not converged. Convergence is explicitly verified on each variable of the problem. The norm of the corrections must be smaller than 10^{-8} for all variables before we go to the next time step.

5 NUMERICAL EXAMPLES

We present a numerical example proposed by Turek et Hron [12]. An elastic bar is attached behind a circular cylinder. The vortex shedding behind the cylinder will provoque oscillations of the flexible structure. The geometry is described in the following figure.



The same dimensions as in [12] were used: $H = 0.41$m, $L = 2.5$m, $h = 0.02$m, and $l = 0.35$m. The cylinder (with radius 0.05m) is centered at $(0.2, 0.2)$. A parabolic velocity profile is imposed at the inlet of the fluid domain as in [12]. The fluid is Newtonian and the solid is modelled using a Saint-Venant-Kirchhoff hyperelastic model

$$\Pi^s = \boldsymbol{F}_s \left(\lambda_s(\text{tr } \boldsymbol{E})\boldsymbol{I} + 2\mu_s \boldsymbol{E}\right)$$

where $\boldsymbol{E} = \frac{\boldsymbol{F}^t\boldsymbol{F}-\boldsymbol{I}}{2}$ is the Green-Lagrange tensor. The parameters for the problem are:

Parameters for the solid	Values	Parameters for the fluid	Valeurs
Density ρ_s $[\mathrm{kg} \cdot \mathrm{m}^{-3}]$	10^4	Density ρ_f $[\mathrm{kg} \cdot \mathrm{m}^{-3}]$	10^3
Lamé coefficient λ_s [Pa]	2.0×10^6	Kinematic viscosity ν_f $[\mathrm{m}^2 \cdot \mathrm{s}^{-1}]$	10^{-3}
Shear modulus μ_s [Pa]	0.5×10^6	Reynold's number Re	100

As an illustration of the displacement methods described in section 3, we present the fluid meshes obtained after a rigid rotation of $\pi/8$ of the structure. Fig. 1 presents the different meshes near the tip of the structure where the displacement is the largest. As can be easily observed, all meshes contain reversed and unusable elements except for the parabolic smoothing method.

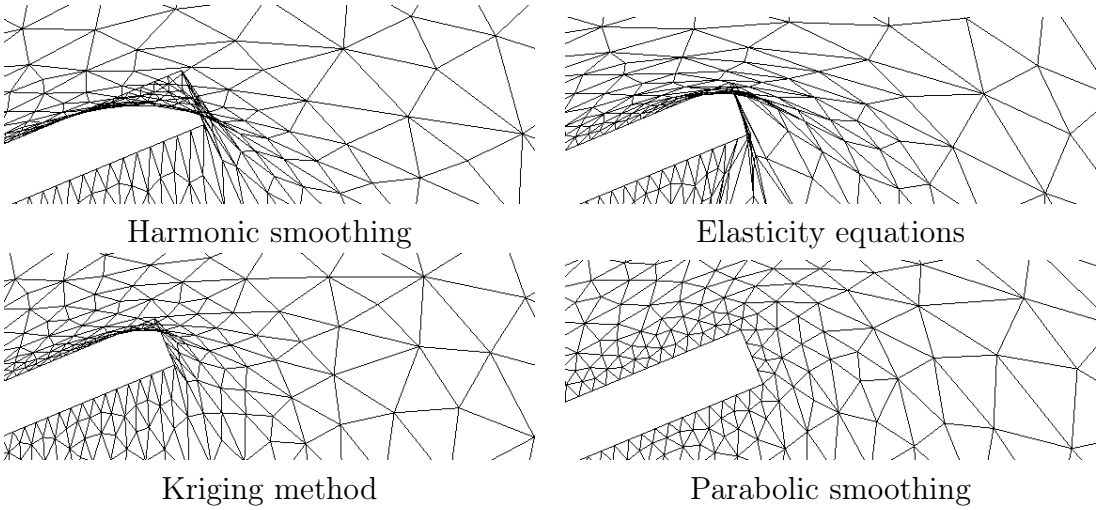


Figure 1: Mesh displacement

Using thus the parabolic smoothing method, we have successfully solved Turek's benchmark problem. We have been able to reproduce accurately the amplitude end frequency of the displacement of the tip of the structure. Figure 2 shows the velocity field and the position of the structure at three time steps. Note that the kriging method can also be used for the update od the fluid domain. The Laplacian smoothing and the elasticity equations led to reversed elements. Other examples will be presented at the conference including 3D problems.

6 CONCLUSIONS

We have presented in this paper an ALE formulation for fluid-structure interaction problems. In the long term, we are interested in the numerical modeling on aquaplaning which is a major issue in the tyre industry.

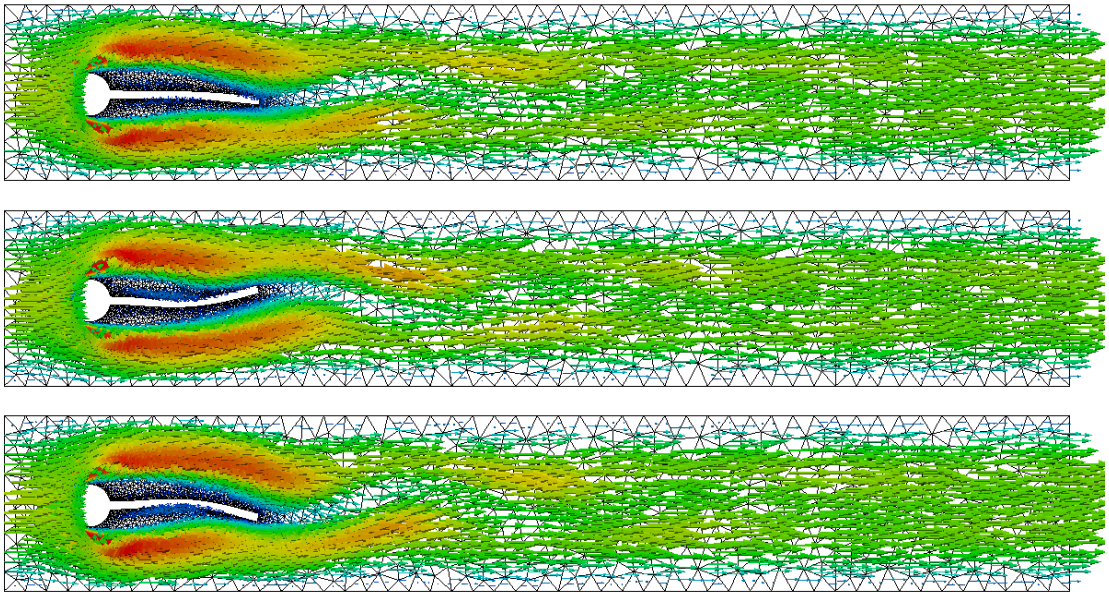


Figure 2: Velocity field and position of the structure

REFERENCES

[1] J. Donea, A. Huerta, J.-Ph. Ponthot, and A. Rodríguez-Ferran. *Arbitrary Lagrangian–Eulerian Methods*. John Wiley & Sons, Ltd, 2004.

[2] M.A. Fernández. *Modèles Simplifiés d'Interaction Fluide-Structure*. PhD thesis, University of Paris IX, 2001.

[3] J. Bonet and R. D. Wood. *Nonlinear Continuum Mechanics for Finite Element Analysis*. Cambridge University Press, Cambridge (UK), New York, 2008.

[4] J.U. Brackbill and J.S. Saltzman. Adaptive zoning for singular problems in two dimensions. *J. Comput. Phys.*, 46(3):342–368, 1982.

[5] J.S. Chen, W.K. Liu, and T. Belytschko. Arbitrary Lagrangian–Eulerian methods for materials with memory and friction. In Hughes TJR Tezduyar TE, editor, *Recent developments in computational fluid dynamics*, volume 95, San Francisco, 1988.

[6] D.G. Krige. A statistical method for mine variation problems in the Witwatersrand. *Journal of Chemistry and Metallurgy of the Mining Society of South Africa*, 52:119–139, 1951.

[7] G. Matheron. The intrinsec ramdom functions and their applications. *Advances in Applied Probability*, 5:439–468, 1973.

[8] M. Olivier. *A Fluid-Structure Interaction Partitioned Algorithm Applied to Flexible Flapping Wing Propulsion.* PhD thesis, Département de génie mécanique, Université Laval, Québec, Canada, 2014.

[9] A. De Boer, M.S. Van der Schoot, and H. Bijl. Mesh deformation based on radial basis function interpolation. *Comp. & Struct.*, 85:784–795, 2007.

[10] D. Boffi, F. Brezzi, and M. Fortin. *Mixed Finite Element Methods and Applications*, volume 44 of *Springer Series in Computational Mathematics*. Springer, Berlin, Heidelberg, 2013.

[11] J. Deteix, A. Jendoubi, and D. Yakoubi. A coupled prediction scheme for solving the Navier–Stokes and heat equations. *SIAM Journal of Numerical Analysis*, 52(5):2415–2439, 2014.

[12] S. Turek and J. Hron. Proposal for numerical benchmarking of fluid-structure interaction between an elastic object and laminar incompressible flow. In H.-J. Bungartz and M. Schäfer, editors, *Fluid-Structure Interaction*, volume 53 of *Lecture Notes in Computational Science and Engineering*, pages 371–385. Springer Berlin Heidelberg, 2006.

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis(Eds)

NUMERICAL PREDICTION OF SHIP RESISTANCE AND SQUAT IN CONFINED WATERS

F. LINDE*†, A. OUAHSINE*, N. HUYBRECHTS*† and P. SERGENT†

*Sorbonne Universités, Université de Technologie de Compiègne
Laboratoire Roberval, UMR-CNRS 7337
Centre de Recherches de Royallieu, CS 60319, 60203 Compiègne Cedex, FRANCE
e-mail: florian.linde@developpement-durable.gouv.fr

†DTecEMF, CEREMA
134 rue de Beauvais, CS 60039, 60280 Margny Lès Compiègne, France

Key words: Ship Resistance, Ship Squat, Restricted Waterways, Computation Fluid Dynamics, Free Surface Flow, Fluid-Structure Interaction.

Abstract. Accurate prediction of hydrodynamic forces opposing a ship displacement in restricted waterways is necessary in order to improve energy efficiency of inland transport. When a ship moves in restricted waterways, a significant increase in ship squat (combination of sinkage and trim) and resistance occurs compared to a movement in open waters. In this paper, a 3D numerical model based on fluid-structure coupling is presented and used to investigate the effect of limited water depth and channel width on ship resistance and squat.

1 INTRODUCTION

When a ship moves in restricted waterways the water in front of its bow is pushed away creating an increase in pressure and a void appears behind the stern creating a decrease in pressure. Due to the reduction of the section where the water can flow, the flow around the ship is also accelerated. The increase in the water speed under the ship causes a decrease in pressure and as a result, a vertical force is applied on the ship making it drop vertically down in the water. The uneven pressure distribution along the ship hull creates a moment along the transverse axis which leads the ship to trim by the bow or the stern. These phenomena have a direct impact on ship resistance: the acceleration of the flow around the hull of the ship causes an increase of friction and ship sinkage increases the quantity of water that the ship needs to push in order to move forward.

In the past, ship resistance in open waters has been studied in towing tank with ship models. The results of those studies have been used to develop empirical models for open waters [1]. However, over the past few years, with the increase in computing power and

development of new numerical methods, there has been a growing interest in Computational Fluid Dynamics (CFD) techniques and its application to ship hydrodynamic and open water ship resistance. Stern et al. [2] reviewed the progress made over the last ten year in CFD applied to ship hydrodynamics and The Gothenburg 2010 Workshop on Numerical Hydrodynamics [3] gathered the results of 33 groups on 18 cases and concluded that the mean error for all computed resistance cases was practically zero.

In shallow and restricted water, near Schijf-limiting speed [4], squat effects become significant and ship resistance increases dramatically [5]. Briggs et al. [6] reviewed PIANC empirical formulas for predicting squat in canals and in restricted and open channels. Ship squat has also been studied with numerical methods often based on Boundary Elemement Method [7], Tucks theory [8], CFD [9] and mathematical models [10]. However, most of these methods do not model ship sinkage and as result the flow modification induced by the squat phenomenon is not accounted for. Some empirical model have been developed in order to predict ship resistance in shallow water [11] and restricted waterways [12]. These methods use the open water formulas with a correcting factor for ship's speed. Few numerical studies of ship resistance in shallow water and even less in restricted waters have been done. Most of them are based on the resolution of the Reynolds Averaged Navier-Stokes (RANS) equations associated with a phase transport model (Volume Of Fluid method, Level-Set) [13]. To the knowledge of the author, none of those studies include ship squat effect on the resistance. In general, ship resistance and ship squat are studied separately. On one hand, numerical models predicting ship squat solve non-viscous fluid equations (Euler) and thus cannot evaluate ship resistance; on the other hand, numerical ship resistance models do not take ship squat into account in their calculations. Therefore, this study aims to develop a ship resistance prediction model allowing to take ship squat and viscous effects into account.

2 THEORETICAL APPROACH

2.1 Fluid equations

The fluid is considered incompressible. The governing equations for the viscous free surface flow problem are the RANS equations:

$$\frac{\partial u_i}{\partial x_i} = 0 \tag{1}$$

$$\frac{\partial \rho \overline{u}_i}{\partial t} + \frac{\partial}{\partial x_j}(\rho \overline{u}_i \overline{u}_j) = \frac{\partial \overline{p}}{\partial x_i} + \frac{\partial}{\partial x_j}(\mu \frac{\partial \overline{u}_i}{\partial x_j} - \rho \overline{u_i' u_j'}) \tag{2}$$

where x_i is the i^{th} (i=1,2,3) component of the fixed coordinate system, ρ is the fluid density, $\overline{u}_i$ and $\overline{p}$ are the mean velocity and pressure, μ is the fluid dynamic viscosity and $-\rho\overline{u_i' u_j'}$ is the Reynolds stress. To solve the closure problem, the SST $\kappa - \omega$ turbulence model has been used.

2.2 Solid equations

The ship's hull is modelled as a rigid body. The movement of the ship is governed by the rigid body equations of motion written at the body's centre of mass:

$$m\overrightarrow{a_G} = \sum \overrightarrow{F} \tag{3}$$

$$\frac{d}{dt}(I_G \cdot \overrightarrow{\omega_G}) = \sum \overrightarrow{M_G} \tag{4}$$

where m is the mass of the body, $\overrightarrow{a_G}$ is the body acceleration, $\sum \overrightarrow{F}$ is the sum of the external forces, I_G is the moment of inertia tensor, $\overrightarrow{\omega_G}$ is the angular velocity and $\sum \overrightarrow{M_G}$ is the sum of the applied torques.

3 NUMERICAL SOLUTION

3.1 Fluid equations

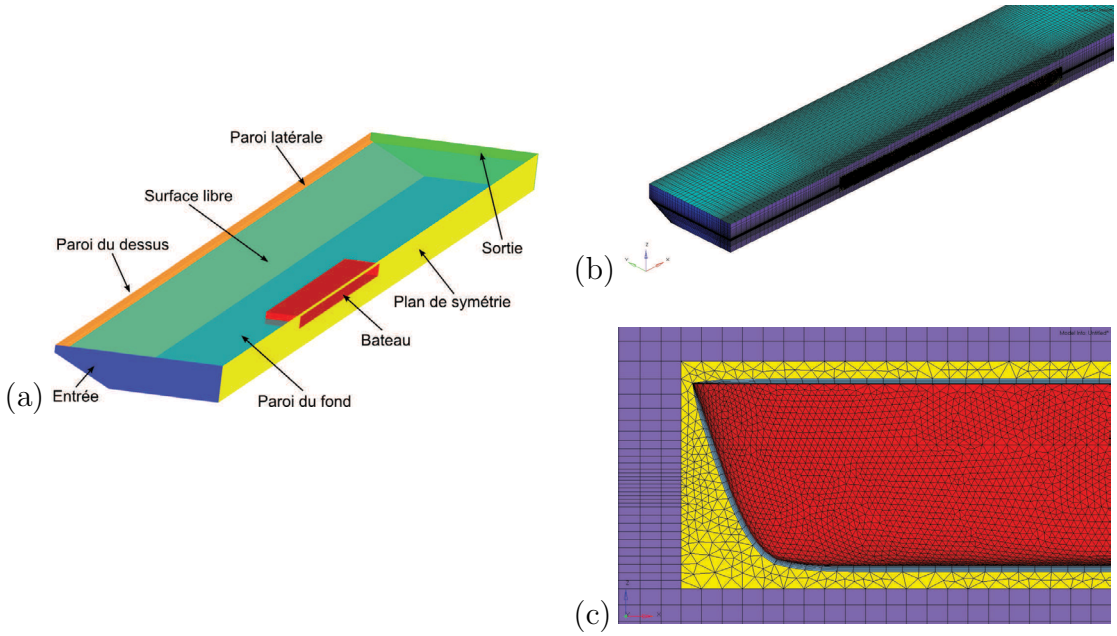


Figure 1: Geometric domain and associated mesh: (a) computational domain with boundaries, (b) meshed domain with higher density on free surface and around ship hull, (c) box with unstructured and boundary layer mesh

Fluid equations (eqs. (1) and (2)) are solved by using Ansys Fluent CFD code. This solver uses finite volume method to discretise the domain into a finite set of control volumes on which the general conservation (transport) equations are solved. The pressure-velocity coupling is solved by using the SIMPLE algorithm. The VOF method is used

in order to simulate the free surface and ship waves [14, 15]. The 2^{nd} order Upwind scheme is used for the dicretization of the convective terms. Pressure conditions were specified at the open boundaries (inlet and outlet); wall conditions with no slip were used for the bottom, the side and the ship hull; and symmetry conditions were set for the symmetry plane and the top. A structured-unstructured hybrid mesh consisting of about 1.3 millions cell was generated for the whole domain. Structured elements are used on the whole domain except around the hull where an unstructured mesh was generated in order to remesh the domain during the actualisation of the ship position. A prismatic boundary layer mesh has also been used around the ship hull. Figure 1 below shows the geometric domain and associated mesh.

3.2 Solid equations

In order to find the ship equilibrium position, the rigid body motion equations (eqs. (3) and (4)) are usually solved in time until the body reach equilibrium. As a result, this method requires to model the transient state and ship oscillations around the hull. However, this study only takes interest in the equilibrium position and steady state. Moreover, here, only ship sinkage is modelled. Therefore, the equilibrium position is defined as follows:

$$F_z = (\overrightarrow{P} + \overrightarrow{F_f}) \cdot \overrightarrow{e_z} = (\iint_S \boldsymbol{\sigma} \cdot \overrightarrow{n} dS - m\overrightarrow{g}) \cdot \overrightarrow{e_z} = 0 \tag{5}$$

where F_z is the projection of forces on $\overrightarrow{e_z}$ axis, m is the mass of the ship, $\overrightarrow{g}$ is the gravitational acceleration, $\overrightarrow{F_f}$ is the force from the fluid acting on the hull of the shp which can be decomposed as viscous and pressure forces, $\boldsymbol{\sigma} = \boldsymbol{\tau} - p\mathbf{I}$ is the stress tensor, $\boldsymbol{\tau}$ is the shear stress, p is the pressure, $\mathbf{I}$ is the identity matrix, $\overrightarrow{n}$ is the normal vector to the ship's hull and S is the wetted surface of the ship's hull. Newton's method was used in order to solve this nonlinear equation. Let z^N be the vertical position of the gravity center and $F_z(z^N)$ the sum of the vertical forces at the N^{th} iteration of the Newton algorithm. The iteration of the algorithm is given by:

$$z^{N+1} = z^N - \frac{F_z(z^N)}{F'_z(z^N)} \tag{6}$$

As there is no analytical expression for F_z, its derivative is calculated using finite difference.

3.3 Fluid-structure interaction (FSI)

The coupling between fluid and solid models is as follows: solving the fluid equations allows to evaluate the forces acting on the hull, which are then used as an input in order to iterate Newton's method and calculate the displacement of the ship; and finally the fluid equations are solved again. This iterative process is presented in fig. 2.

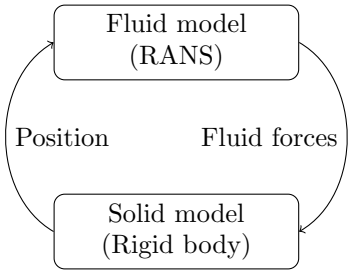


Figure 2: Fluid-structure interaction

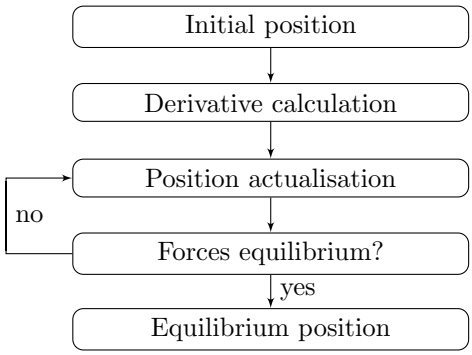


Figure 3: Numerical resolution process

As the derivative of F_z is calculated by finite difference, it is necessary to evaluate the forces at two positions: z^N and $z^N + dz_0$, where dz_0 is a small increment of the vertical position. Moreover, after each actualisation of the position, the solution of the fluid problem is disturbed and n iterative resolutions of the fluid equations are required to reach a new stable solution and change the ship's position again. As a result, calculating the derivative of the forces can be time consuming and in order to make the calculations faster, a quasi-Newton method has been used, where the derivative is calculated only once and kept constant during all the iterative process. A study comparing the speed convergence of the two methods has been carried out and it has been shown that the quasi-Newton method converges faster but it needs more iterations N compared to the Newton algorithm. Figure 3 describes the numerical resolution process.

4 EXPERIMENTAL DATA

In 2013, in collaboration with the University of Liege (Belgium), an experimental work was carried out at the ANAST towing tank in order to validate the numerical model. In these experiments a 135 m (full size) inland vessel was used at 1/25 scale. A wide range of parameters were tested such as channel width (W), channel depth (H) and ship draft (T). The boat was free to sink and trim and forces and moments acting on the ship as well as sinkage and trim were recorded.

In this paper, the effect of limited water depth and channel width on ship resistance and squat is investigated. According to ITTC87 [16], some of the parameters to consider in order to estimate the nature and level of restriction for waterways are:

- water depth to draught ratio $\frac{H}{T}$, if $\frac{H}{T} < 4$, there is an influence of the bottom on the flow around the hull;
- water width to ship breadth (B) ratio $\frac{W}{B}$, if $\frac{W}{B} < 4$, the flow around the hull changes;
- canal section (A_C) to midship section (A_B) $\frac{A_C}{A_B}$ ratio, a general restriction of the waterway starts when $\frac{A_C}{A_B} < 15$.

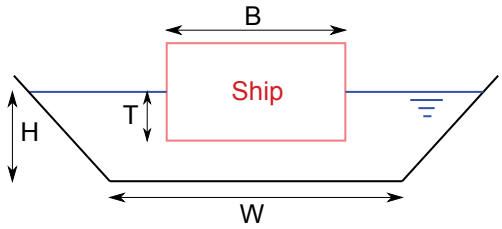


Figure 4: Schematic representation of the model and waterway geometric parameters

Table 1 sums up the different configurations for which the experimental data were compared to the numerical results as well as the values for the restriction parameters described above.

Table 1: Modelled configurations and corresponding parameters

Conf.	**H** [m]	**T** [m]	**W** [m]	$\frac{\mathbf{H}}{\mathbf{T}}$ [-]	$\frac{\mathbf{W}}{\mathbf{B}}$ [-]	$\frac{\mathbf{A_C}}{\mathbf{A_B}}$ [-]	**V** [m/s]	$\mathbf{V_L}$ [m/s]
C1	0.18	0.10	0.72	1.80	1.58	4.26	0.11-0.51	0.58
C2	0.18	0.10	1.44	1.80	3.16	7.11	0.22-0.56	0.74
C3	0.18	0.10	2.88	1.80	6.32	12.79	0.33-0.67	0.88
C4	0.24	0.10	2.88	2.40	6.32	17.68	0.44-0.89	1.10
C5	0.18	0.04	2.88	4.50	6.32	31.97	0.44-0.89	1.04

V_L=Schijf limiting speed [4]

5 RESULTS AND DISCUSSION

5.1 Mesh convergence study

A grid convergence study was carried out for the configuration 3 using a fine grid and coarsening it two times by a factor $\sqrt{2}$ in all directions. Table 2 shows the number of elements as well as the change in drag (C_D) and lift (C_L) coefficients.

Table 2: Dependence of drag (C_D) and lift (C_L) coefficients on grid density (configuration 3, V=0.67 m/s)

Grid	**No. Elements**	$\mathbf{C_D}$ [-]	$\mathbf{C_L}$ [-]	$\frac{\mathbf{C_D}}{\mathbf{C_{Dfine}}}$ [%]	$\frac{\mathbf{C_L}}{\mathbf{C_{Lfine}}}$ [%]
Coarse	0.34 M	0.310	2.847	106.07	99.53
Medium	0.85 M	0.296	2.856	101.31	99.85
Fine	2.23 M	0.293	2.860	100	100

Table 2 shows that the grid sensitivity between the medium and fine grid is very limited : the change for the lift coefficient is less than 1% and around 1.3% for the drag coefficient. As a result, an intermediate grid between the medium and fine grid was chosen to maintain an affordable computational cost. This grid was generated by keeping similar mesh parameters outside the ship box and increasing mesh density inside ship box which resulted in 1.3 M elements.

5.2 Quasi-Newton algorithm convergence

Figure 5 below illustrates the convergence of the quasi-Newton algorithm for configuration 3 at speed $V = 0.57$ m/s.

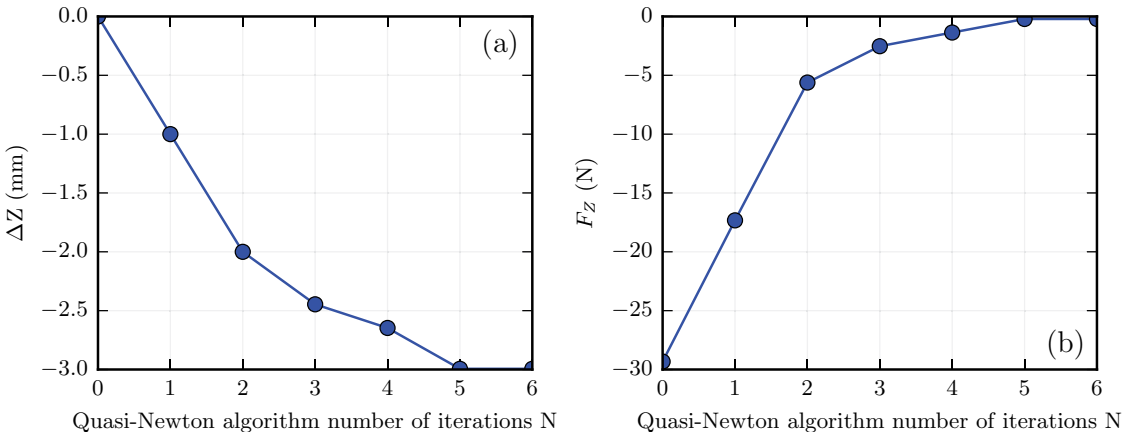


Figure 5: Quasi-Newton algorithm convergence for configuration 1 at speed $V = 0.57$ m/s : change in (a) ship sinkage ΔZ and (b) vertical forces F_Z

In this case, the equilibrium position is reached after 5 iterations of the algorithm. The stopping criteria is set for $F_Z < 0.5\ N$ and in order to prevent any remeshing error (negative volume) the maximum displacement is set to 1mm. It is worth noticing that this algorithm can easily be extended to two degrees of freedom (in order to take trim into account for instance) by introducing the Jacobian matrix. Linde et al. (2014) [17] studied the influence of taking sinkage into account on the numerical results but in the experiments carried out at the University of Liege, the trim angle was so low that it had no impact on ship resistance and ship sinkage was the predominant factor. As a result, only ship sinkage is modelled in this study.

5.3 Restricted-water effect and comparison with experimental data

Figure 6 compares numerical results with experimental data for the 5 configurations listed in table 2.

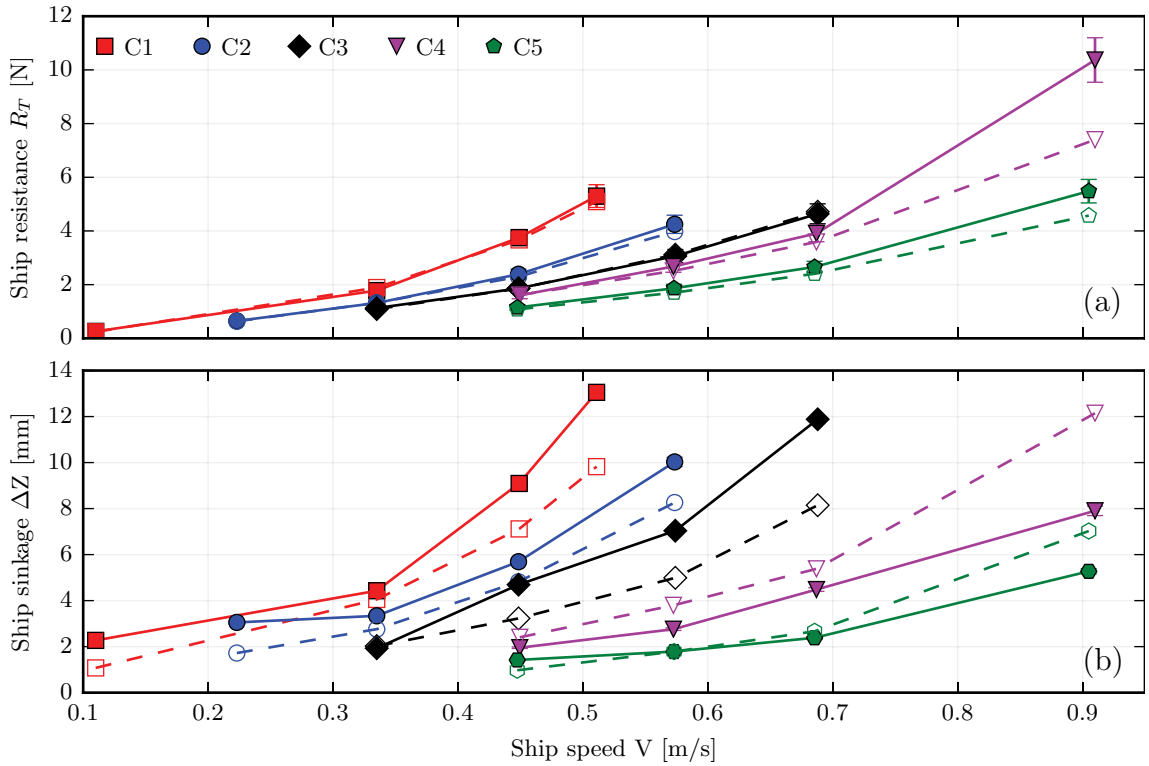


Figure 6: Restricted-water effect: comparison between numerical (full line with filled markers) and experimental (dashed line with empty markers) results for conf. 1-5. Change in (a) ship resistance R_T and (b) ship sinkage ΔZ against ship speed V.

From fig. 6, it can be seen that as restriction increases (i.e $\frac{A_C}{A_B}$ ratio decreases - C5 to C1) ship resistance and sinkage increase. For instance, for $V = 0.44\ m/s$, ship resistance and sinkage in configuration 1 are respectively 1.9 and 4.7 times higher than in configuration 4. The comparison between numerical and experimental results shows that for ship speed up to 0.7 m/s the predicted resistance is in a very good agreement with the experimental data. For ship sinkage, the numerical results show the same trend as the experiment however, there seems to be an offset. This could be explained by the potential error made when measuring the initial draft during the experiment as the uncertainty on the measure is $\pm$ 1 mm. For the ship speed of 0.9 m/s the numerical model overestimates the sinkage and underestimate the resistance. It is possible that for this speed the model is not able to reproduce correctly the pressure field around the ship hull therefore underestimating vertical and horizontal pressure forces. This prediction error is currently under investigation. So far, the influence of parameters such as mesh density, boundary conditions or turbulence model have been tested, but none of them had an impact on the error.

5.4 Influence of ship sinkage on the resistance prediction error

Figure 7 shows the predicted resistance with sinkage without sinkage as well as the experimental data for configuration 1,2 and 3.

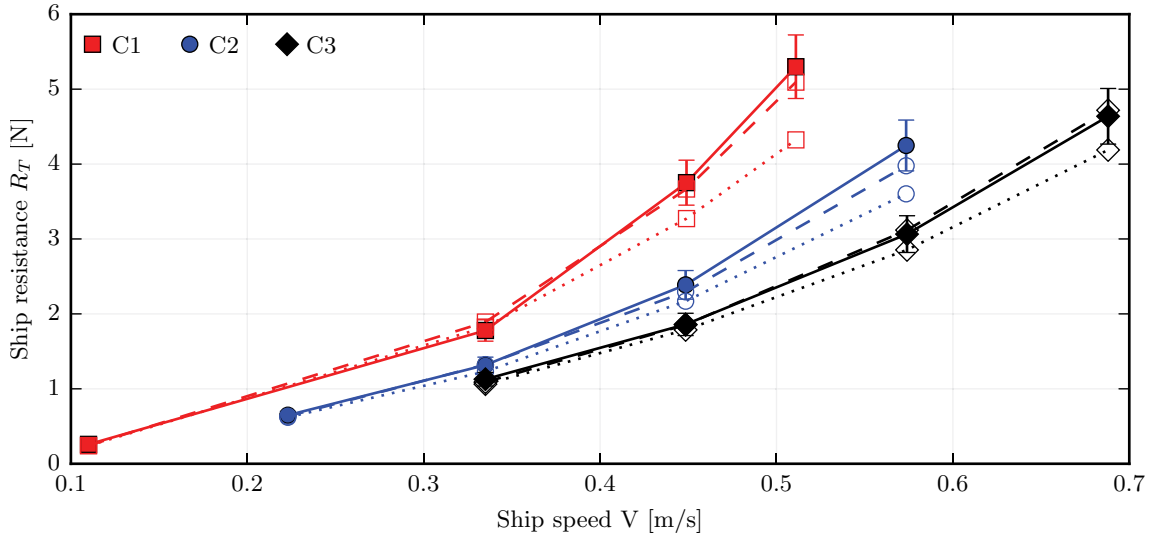


Figure 7: Comparison between the predicted resistance with sinkage (dashed line, empty markers) and without sinkage (dotted line, empty markers) as well as the experimental data (full line with filled markers) for configuration 1,2 and 3.

Figure 7 shows that up to a speed of 0.33 m/s, the predicted resistance with or without sinkage does not differ much. However, above that speed, there is a gap between the measured data and predicted resistance without sinkage; meanwhile the plot of the resistance with sinkage remains very close to the experimental data. This observation is sustained by the calculated prediction error: the error with sinkage is almost always smaller than that without sinkage. Moreover, for those three configurations, the maximum error is 6.3 % with sinkage whereas it reaches 18.4 % without sinkage. Moreover, the calculated increase in ship resistance due to ship sinkage can be very significant: it reaches a maximum value of 18% for configuration 1 at 0.51 m/s. It can also be seen that the prediction error without sinkage significantly increases with the speed and the restriction of the water, which is not the case for the predicted resistance with sinkage. Those facts highlight the significance of taking ship sinkage into account in order to accurately predict ship resistance in restricted waterways.

5.5 Influence of restriction parameters on ship resistance increase

In order to study if the depth restriction ($\frac{H}{T}$) or the width restriction ($\frac{W}{B}$) has more effect on resistance, two sets of simulations were carried out. Starting from a reference configuration with a rectangular channel, in one case the water depth was decreased while

the channel width remained constant, and the opposite was done in the other case. The increase in ship resistance is then compared with the reference resistance R_{Tref} calculated for the initial configuration. The use of a rectangular channel ensures the similarity of the definition of the two ratio and the reference configuration was chosen so that both have the same value ($\frac{H}{T} = \frac{W}{B} = 5$, equal restriction). For this study, the ship described in section 4 was used, the draught remained constant and two speeds were tested (0.4 m/s and 0.6 m/s). The two set-ups are described in table 3 and illustrated in fig. 8.

Table 3: Modelled set-ups used to study the influence of restriction parameters

Setup	**H** [m]	**W** [m]	$\frac{\mathbf{H}}{\mathbf{T}}$ [-]	$\frac{\mathbf{W}}{\mathbf{B}}$ [-]
Ref	0.5	2.28	5	5
S1	0.15-0.4	2.28	1.5-4	5
S2	0.5	0.68-1.82	5	1.5-4

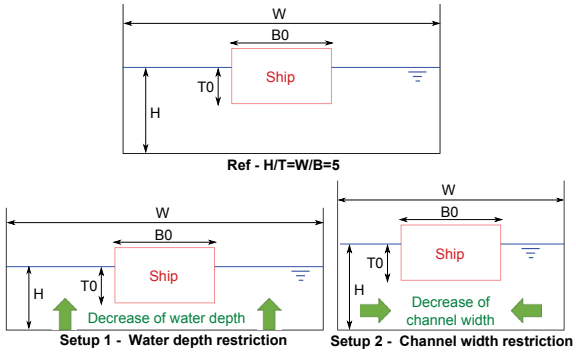


Figure 8: Illustration of the two tested set-ups

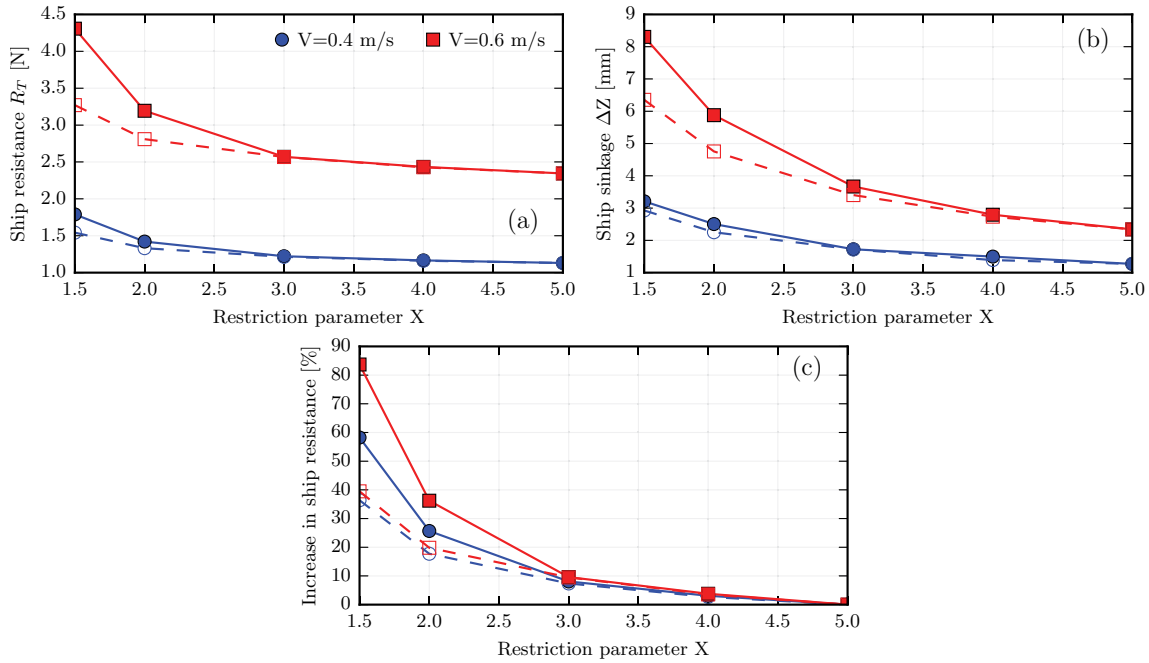


Figure 9: Evolution of (a) ship resistance, (b) ship sinkage and (c) increase of ship resistance against restriction parameter X ($X = \frac{H}{T}$ for set-up 1 (full line with filled markers) and $X = \frac{W}{B}$ for set-up 2 (dashed line with empty markers)

Figure 9 shows the evolution of ship resistance (fig. 9 (a)) and ship sinkage (fig. 9 (b)) as well as the increase in ship resistance $\frac{R_T(X)-R_{Tref}}{R_{Tref}}$ (fig. 9 (c)) due to depth restriction in set-up 1 and channel width restriction in set-up 2. Figure 9 (a) and fig. 9 (b) show that significant differences in ship resistance and sinkage between the water depth and channel width restriction begin to appear when the restriction parameter is inferior to 3. This difference also increases with the increase in ship velocity. It can also be seen that for a given speed, the resistance and sinkage start to rise for a value of the restriction parameter below 4, which is in agreement with the findings of the ITTC87 mentioned in section 4. It is clear from fig. 9 (c) that water depth restriction has more influence on ship resistance increase than channel width restriction: for the 0.6 m/s velocity, the increase amounts to 84 % for $\frac{H}{T} = 1.5$ against 39 % for $\frac{W}{B} = 1.5$. This can be explained by the fact that ship sinkage is less sensitive to vertical restriction than horizontal restriction. Indeed, under the hull, the increase of flow velocity and the pressure drop will be less impacted by a decrease of channel width than a decrease of water depth.

6 CONCLUSIONS

The numerical model presented in this article allows to calculate ship resistance in restricted waterways by taking ship sinkage into account. The use of a Newtonian method in order to find the equilibrium position allows to skip the transient state. Moreover, this method can easily be extended to two degrees of freedom by using a Jacobian matrix. The comparison with experimental data has shown that predicted resistance and sinkage were in good agreement for velocities below 0.7 m/s but discrepancies appeared for the highest tested velocity (0.9 m/s). Taking ship sinkage into account allowed to significantly reduce the prediction error. Finally, the model was used to study the influence of channel width and water depth restriction. The results showed that in the case of water depth restriction, ship sinkage was more significant which resulted in higher ship resistance increase.

ACKNOWLEDGEMENTS

The authors thank André Hage, Mathieu Lapy, Mathieu Vandescuren and Adrian Constantinescu (ANAST, University of Liege, Belgium) for their contribution with the experimental campaign and their continuous support.

REFERENCES

[1] Holtrop, J. and Mennen, G.G. An approximate power prediction method. *International Shipbuilding Progress* (1982) **19**:166–170.

[2] Stern, F. *et al.* Computational ship hydrodynamics: nowadays and way forward. *International Shipbuilding Progress* (2013) **60**:3–105.

[3] Larsson, S., Stern, F. and Visonneau, CFD in ship hydrodynamics—results of the Gothenburg 2010 workshop. *Marine 2011* (2013) 237–259.

[4] Schijf, J.B. Influence on the Form and Dimensions of the Cross-Section of the Canal, of the Form, of the Speed and the Propulsion System of Vessels. *XVIIth PIANC* (1949).

[5] Constantine, T. The behaviour of ship moving in restricted waterways. *ICE Proceedings* (1961) **19**:549–562.

[6] Briggs, M.J., Vantorre, M., Uliczka, K., Debaillon, P. Prediction of squat for underkeel clearance. *Handbook of Coastal and Ocean Engineering* (2013).

[7] Härting, A., Laupichler, A., Reinking, J. Considerations on the squat of unevenly trimmed ships. *Ocean Engineering* (2009) **36**:193–201.

[8] Gourlay, T. Slender-body methods for predicting ship squat. *Ocean Engineering* (2008) **35**:191–200.

[9] Debaillon, P. Numerical investigation to predict ship squat. *Journal of Ship Research* (2010) **54**:133–140.

[10] Lataire, E., Vantorre, M., Delefortrie, G. A prediction method for squat in restricted and unrestricted rectangular fairways. *Ocean Engineering* (2012) **55**:71–80.

[11] Schlichting, O. Ship Resistance in Water of Limited Depth-Resistance of Sea–Going Vessels in Shallow Water. *Jahrbuch der STG* (1934) **35**:127–148.

[12] Geerts, S., Verwerft, B., Vantorre, M., Van Rompuy, F., *et al.* Improving the efficiency of small inland vessels. *European Inland Waterway Navigation Conference* (2010).

[13] Toxopeus, S. L. Viscous-flow calculations for KVLCC2 in deep and shallow water. *Marine 2011* (2013) 151–169.

[14] JI, S.C., Ouahsine, A., Smaoui, H., Sergent, P. 3-D numerical simulation of convoy-generated waves in a restricted waterway. *Journal of Hydrodynamics, Ser. B* (2012) **24**:420–429.

[15] JI, S.C., Ouahsine, A., Smaoui, H., Sergent, P. 3D Numerical modeling of Sediment resuspension induced by the compounding effects of ship-generated waves and the ship propeller. *Journal of Engineering Mechanics* (2013) **140 (6)**:04014034.

[16] Resistance and flow committee. Report of the resistance and flow committee. *18th International Towing Tank Conference* (1987).

[17] Linde, F., Ouahsine, A., Huybrechts, N., Sergent, P. Effet de la pronfondeur limitée sur la résistance à l'avancement et l'enfoncement dynamique d'une barge fluviale. *14èmes Journées de l'Hydrodynamique* (2014).

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis (Eds)

NUMERICAL SIMULATION OF FLOW INTERACTION BETWEEN STATIONARY AND DOWNSTREAM ELASTICALLY MOUNTED CYLINDERS AT LOW REYNOLDS NUMBERS

PAULO R. F. TEIXEIRA[*] AND ERIC DIDIER[†]

[*] Universidade Federal do Rio Grande (FURG)
Avenida Itália km8, Campus Carreiros, 96203-900, Rio Grande, RS, Brazil
E-mail address: pauloteixeira@furg.br; web page: http://www.furg.br

[†] Departamento de Hidráulica e Ambiente, Laboratório Nacional de Engenharia Civil (LNEC)
Avenida do Brasil, 101, 1700-066, Lisbon, Portugal
E-mail address: edidier@lnec.pt; web page: http://www.lnec.pt

Key words: Finite Element Method, Flow-induced Vibration, Circular Cylinders, Tandem Arrangement, Wake Interference.

Abstract. The vortex-induced vibration phenomenon can occur as a result of the action of wind on bridges, slender buildings, chimneys and energy transmission cables besides the action of water flow on pipelines and risers, among others. Despite the simplicity of the geometry of the circular cylinders, the uniform flow around them is very complex and important, since it may induce unsteady forces on structures associated with vortex shedding. This paper describes the study of two circular cylinders in tandem arrangement subject to bi-dimensional uniform laminar flows at low Reynolds numbers. The numerical model Ifeinco, which is based on the finite element method and uses a partitioned scheme that considers two-way interaction of fluid flow and structure, has been employed in the analysis. The fluid flow model uses a semi-implicit two-step Taylor-Galerkin method to discretize the Navier-Stokes equations whereas the arbitrary Lagrangean-Eulerian formulation to follow the cylinder motion. This movement has been described by the one DOF dynamic equation for the transverse direction discretized in time by the implicit Newmark method. Both cylinders are immersed in water and the downstream one is elastically mounted in transversal direction. Firstly, stationary cylinders in tandem arrangement for Re = 100 are analysed for L/D from 1.5 to 6.0. Results of lift and drag coefficients and Strouhal number are compared with other numerical results and good agreement is found. These analyses show that the vortex shedding occurs for both cylinders for gaps $L/D > 4.0$ and the wake behind the downstream cylinder is formed by the combination of vortex shed of both cylinders. Secondly, numerical simulations considering downstream elastically mounted cylinder for $L/D = 5.25$ are analysed for Reynolds numbers ranging from 100 to 140. It shows that the resonance occurred for Reynolds numbers between 115 and 120, unlike the range obtained for a single cylinder, from 102 to 113, submitted to the same conditions. Furthermore, the maximum dimensionless amplitude of oscillation is 0.721 for Re = 118, which is much higher than the one of the single cylinder (0.422 for Re = 103). The interaction between cylinders changes the Strouhal number in relation to the one of the single cylinder; because of this, there are differences between the lock-in regions.

1 INTRODUCTION

The vortex-induced vibration phenomenon can occur as a result of the action of wind on bridges, slender buildings, chimneys and energy transmission cables besides the action of water flow on pipelines and risers. The wake around circular cylinders due to a uniform flow leads to several complex phenomena. Despite the simplicity of geometry, the flow around cylinders requires deep studies, since it may induce unsteady forces on structures associated with vortex shedding.

This phenomenon, which appears in many practical situations of different bluff body arrangements, has its complexity increased due to the wake interference among circular cylinders. Several interference regimes have been classified, according to the Reynolds number and distance between cylinders, based on experimental studies, such as Igarashi's [1], Zdravkovich's [2,3] and Sumner et al.'s [4]. Considering tandem arrangement of circular cylinders [1,5], a critical distance, Lc, between both cylinders is identified ($L/D \approx 4$, where L is the center to center distance between cylinders and D is the diameter) in which discontinuity of the flow behavior and vortex shedding occur. For shorter distances ($L < Lc$), the mean drag of the downstream cylinder is small and negative. For longer distances ($L > Lc$), fluctuating lift and drag forces of the downstream cylinder become higher than the upstream ones.

Researchers, such as Zdravkovich [6], Brika and Laneville [7], investigated experimentally the flow induced oscillations of two interfering circular cylinders considering different arrangements and distances between them. The authors observed that the vortex induced oscillations are strongly dependent on the arrangement of cylinders and the gap between cylinders. Assi et al. [8], Okajima et al. [9] and Huera-Huarte and Gharib [10] analyzed experimentally the flow induced oscillations specifically for tandem arrangements.

Other researchers, such as Li et al. [11], Slaouti and Stanby [12], Mittal et al. [13], Meneghini et al. [14], Sharman et al. [15], Carmo [16], Carmo and Meneghini [17], studied several parameters of the flow for different arrangements between cylinders based on numerical simulations. Some numerical analyses (Carmo et al. [18], Mittal and Kumar [19], Papaioannou et al. [20]) have been developed for the study of fluid-structure interaction in tandem arrangements of circular cylinders considering two-dimensional cases at different Reynolds numbers, by using vortex discrete method, finite volume method and spectral element method.

This paper describes the study of two circular cylinders in tandem arrangement subject to bi-dimensional uniform laminar flows at low Reynolds numbers (from 90 to 140). Firstly, stationary cylinders in tandem arrangement for Re = 100 are analysed for L/D from 1.5 to 6.0. Afterwards, cases with center to center distance between cylinders equal to $5.25D$ for both fixed cylinders and for fixed upstream cylinder and a downstream one elastically mounted in transversal direction are analysed. The numerical model Ifeinco [21] which is based on the finite element method and uses a partitioned scheme that considers two-way interaction of fluid flow and structure is employed for the analysis.

2 IFEINCO MODEL

The numerical model Ifeinco is based on a partitioned scheme, in which the fluid flow and

the structure are solved in two-way interaction. Basically, updating the variables of the flow consists of following steps [22]:

a) Calculate non-corrected momentum per volume $\tilde{U}_i$ at $t+\Delta t/2$, where the pressure term is at t instant, according to Eq. (1).

$$\tilde{U}_i^{n+1/2} = U_i^n - \frac{\Delta t}{2}\left(\frac{\partial f_{ij}^n}{\partial x_j} - \frac{\partial \tau_{ij}^n}{\partial x_j} + \frac{\partial p^n}{\partial x_i} - w_j^n \frac{\partial U_i^n}{\partial x_i}\right) \tag{1}$$

where p is the pressure, w_i are the velocity components of the reference system, τ_{ij} is the viscous stress tensor, $U_i = \rho \mathrm{v}_i$, $f_{ij} = \mathrm{v}_j(\rho \mathrm{v}_i) = \mathrm{v}_j U_i$ $(i, j = 1, 2)$, ρ is the specific mass and v_i are the velocity components.

b) Update the pressure p at $t+\Delta t$, given by the Poisson equation:

$$\frac{1}{c^2}\Delta p = -\Delta t\left[\frac{\partial \tilde{U}_i^{n+1/2}}{\partial x_i} - \frac{\Delta t}{4}\frac{\partial}{\partial x_i}\frac{\partial \Delta p}{\partial x_i}\right] \tag{2}$$

where $\Delta p = p^{n+1} - p^n$ and $i = 1, 2$.

c) Correct the velocity at $t+\Delta t/2$, adding the pressure variation term from t to $t+\Delta t/2$, according to the equation:

$$U_i^{n+1/2} = \tilde{U}_i^{n+1/2} - \frac{\Delta t}{4}\frac{\partial \Delta p}{\partial x_i} \tag{3}$$

d) Calculate the velocity at $t+\Delta t$ using variables updated in the previous steps as follows:

$$U_i^{n+1} = U_i^n - \Delta t\left(\frac{\partial f_{ij}^{n+1/2}}{\partial x_j} - \frac{\partial \tau_{ij}^{n+1/2}}{\partial x_j} + \frac{\partial p^{n+1/2}}{\partial x_i} - w_j^{n+1/2}\frac{\partial U_i^{n+1/2}}{\partial x_i}\right) \tag{4}$$

The classical Galerkin weighted residual method is applied to the space discretization of Eq. (1), (2), (3) and (4), and a triangular element is employed. In the variables at $t+\Delta t/2$ instant, a constant shape function is used, and in the variables at t and $t+\Delta t$, a linear shape function is employed [21].

The Poisson equation, which is a result of spatial discretization of Eq. (2), is solved by employing the conjugate gradient method with a diagonal preconditioning [23]. Spatially discretized Equation (4) is explicitly solved by the iterative process using the lumped matrix [24].

The mesh velocity transversal component w_2 is computed to diminish element distortions, keeping prescribed velocities on moving and stationary boundary surfaces. The mesh movement algorithm adopted in this paper uses a smoothing procedure for the velocities based on these boundary lines [21].

In order to update the rigid body motion of the structure, it is necessary to calculate displacements and rotations of a hypothetical concentrated mass at its gravity center. In this case study, there is only movement in transverse direction (one degree of freedom – DOF) and, consequently, displacement, velocity and acceleration in this direction are the variables to be determined at each time step. To update the variables of the structure, the rigid motion

of the cylinder is calculated at each instant, after the variables of the flow (pressure and viscous stress) are known. For this case study, one DOF dynamic equation is considered for the transverse direction, as follows:

$$m\ddot{y} + c\dot{y} + ky = F \tag{8}$$

where $\ddot{y}$, $\dot{y}$ and y are the transverse acceleration, velocity and displacement, respectively; m is the mass; c is the damping coefficient; k is the stiffness; and F is the dynamic force. In Ifeinco code, Eq. (8) is discretized in time by using the implicit Newmark method [25].

3 NUMERICAL SIMULATION

Both cylinders, which are immersed in water, have diameters D = 0.0016 m and their center to center distance between cylinder is equal to L (Fig. 1). In cases when the downstream cylinder (mass m = 0.2979 kg) is considered elastically mounted in transversal direction, the spring stiffness, k, is equal to 579 N/m, resulting in a natural frequency of f_n = 7.016 Hz while the damping coefficient, c, is equal to 0.0325 kg/s, its correspondent damper ratio is $\zeta = c / 2m\omega_n$ = 0.0012, where ω_n is the angular frequency, and the mass ratio ($M = m/\rho D^2$, where ρ is the specific mass of fluid) is equal to 166 [26]. Constant velocity (U_∞) is imposed on the inlet boundary whereas, on the lateral boundaries (far from cylinders), a slide condition is imposed; the outlet boundary is free exit, but null pressure is imposed on its center.

The computational domain consists of a rectangle whose sides have the minimum distance from the cylinders of 100D. The cylinder boundary is discretized in 200 segments and the size of the first element around the cylinders is 0.016D, totalizing from 228000 to 252000 nodes and from 455000 to 505000 triangular elements, depending on the distance between both cylinders. The time step is between $6\text{x}10^{-5}$ s and $8.5\text{x}10^{-5}$ s, according to the Reynolds number [27].

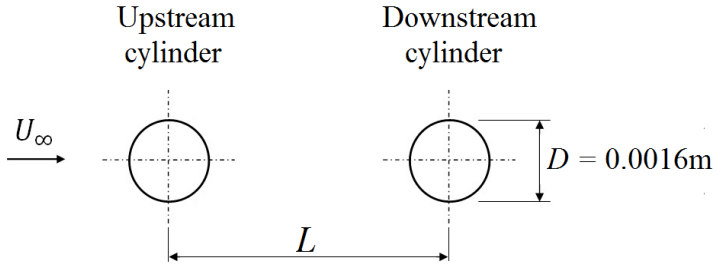


Figure 1: Sketch of the case study

3.1 Stationary upstream and downstream cylinders

In this section, stationary cylinders in tandem arrangement for Reynolds number, Re, equal to 100 (Re = $\rho U_\infty D/\mu$, where μ is the viscosity) are analysed for L/D from 1.5 to 6.0. Results are compared with other numerical ones presented by Didier [5] and Sharman et al. [15].

Figure 2 shows Strouhal numbers ($St = f D/U_\infty$, where f is the vortex shedding frequency) for L/D from 1.5 to 6.0. Ifeinco results had good agreement with those obtained by other authors. Figure 2 indicates the strong influence of the cylinder distances on the Strouhal

number, which experiments an abrupt variation at *L/D* around 4.0 and increases almost 30% from *L/D* = 3.5 to 4.0. The Strouhal numbers are lower than 0.1641, which is the value obtained by Ifeinco for a single cylinder, but the more the distance between both cylinders increases (diminishing the interference between them), the closer the Strouhal number gets to that value.

The distance between cylinders of 4.0 *D*, in which an abrupt variation in the Strouhal number occurs, is named the critical spacing (*Lc*). Different flow behaviors are observed when cases with distances shorter and longer than *Lc* are compared. The analysis of the streamlines for *L/D* = 2.0 (*L* < *Lc*) and 4.0 (*L* > *Lc*), shown in Fig. 5, explains these differences. For *L/D* = 2.0, the wake off the upstream cylinder reattaches to the downstream one and a symmetrical flow pattern occurs between both cylinders. Moreover, long vortices are shed off the downstream cylinder, a fact that results in the lower Strouhal number. For *L/D* = 4.0, vortices are also shed off the upstream cylinder; it moves the reattachment point and energizes the fluid that flows completely around the downstream cylinder.

Figures 3 and 4 show the root mean square lift (C_{Lrms}) and the mean drag (C_{Dmean}) coefficients for different cylinder distances, respectively. The critical spacing (*Lc*) is clear in both figures and the lower values of C_{Lrms} and C_{Dmean} before *Lc* have been explained by the flow behavior shown in Fig. 5. Cases in which *L* > *Lc*, the interference between cylinders provides values of C_{Lrms} higher than those obtained by Ifeinco for the single cylinder (0.2268), considering the same Reynolds number. It occurs with more intensity for the downstream cylinder that is subject to the shed of the upstream one. Figure 4 shows that the mean drag of downstream cylinder is negative for smaller spacing, i.e., this cylinder is pushed forwards by the fluids. From that moment on, the more the distance between both cylinders increases, the more the mean drag slowly increases to small positive values and then varies abruptly at critical spacing. The upstream cylinder presents higher values of C_{Dmean} by comparison with downstream cylinder one, but lower than that obtained by Ifeinco for single cylinder (1.3208).

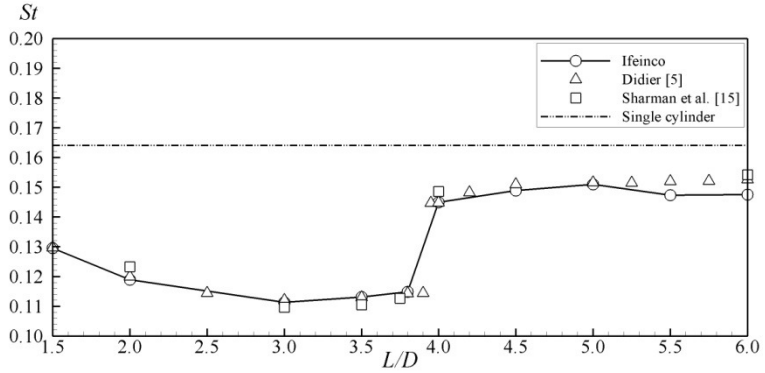


Figure 2: Strouhal number versus *L/D* for stationary cylinder cases.

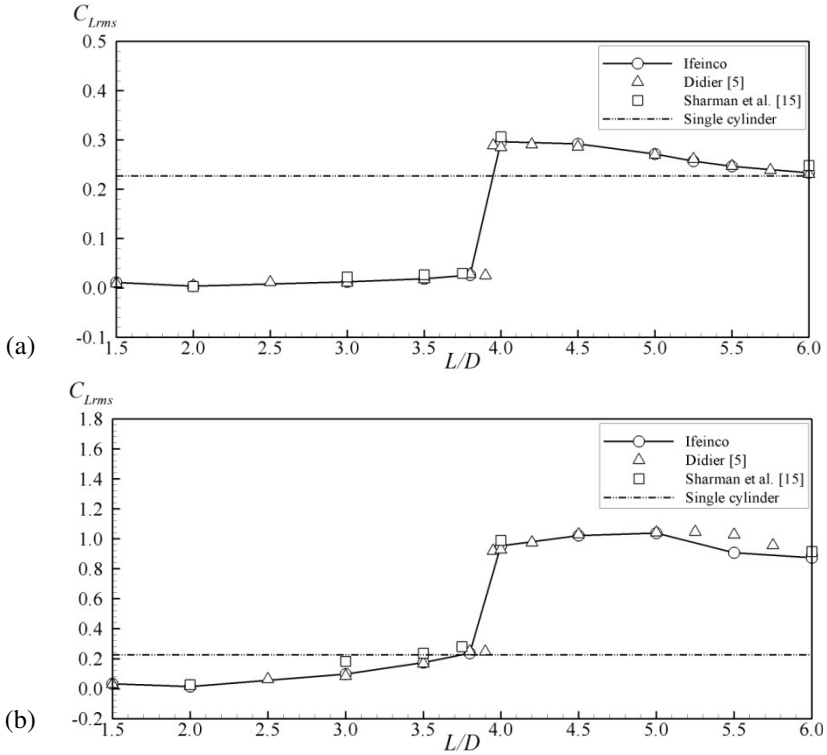


Figure 3: Root mean square of lift coefficient versus *L/D* for upstream cylinder (a) and downstream one (b) in stationary cylinder cases.

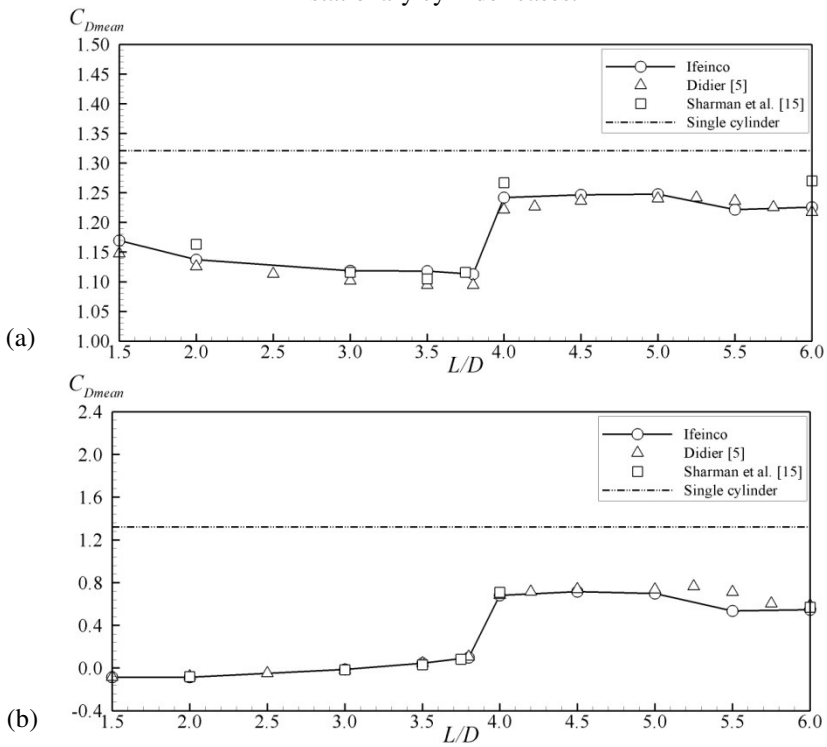


Figure 4: Mean drag coefficient versus *L/D* for upstream cylinder (a) and downstream one (b) in stationary cylinder cases.

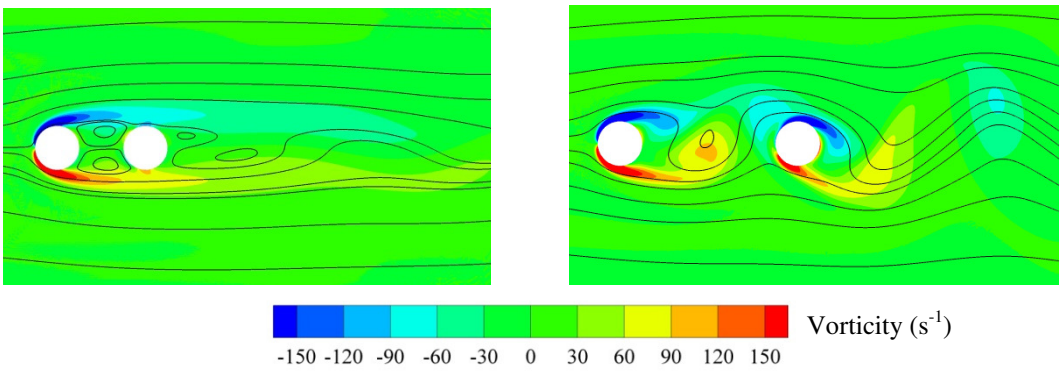


Figure 5: Streamlines around the stationary cylinders for L/D = 2.0 (a) and 4.0 (b).

3.2 Stationary upstream and oscillating downstream cylinders

In this section, the center to center distance between cylinders is 5.25D, the downstream cylinder is elastically mounted and the analysis is performed for Reynolds numbers ranging between 90 and 140, thus, characterizing a regime in which the vortex street is fully laminar and bi-dimensional. Therefore, vortices are expected to shed off the upstream cylinder and roll up before striking the downstream cylinder, thus, strongly interacting with it. For L/D = 5.25, an interesting maximum occurs in C_{Lrms} due to the synchronization of vortex shedding: vortices shed from the upstream cylinder add to the vortices shed from the downstream one, in a way that the fluctuating lift force reaches its maximum. The main dimensionless parameter that influences the VIV behavior is the reduced velocity ($V_R = U_\infty/D\,f_n$), ranging from 5 to 7.8 in this case.

In this case, a significant increase in C_{Lrms} from 1.001 to 1.701 at Reynolds number from 110 (V_R = 6.12) to 119 (V_R = 6.63) and an abrupt decrease in 0.687 at Re = 120 (V_R = 6.68) are observed. This phenomenon is characteristic of the resonance (lock-in) and also occurs in the single cylinder, but with some differences in coefficient magnitude and Reynolds number range, as shown in Fig. 6. The Reynolds number range in which this phenomenon occurs for single cylinder is wider than the range in the tandem arrangement case. Besides, in these regions, the mean drag coefficient for oscillating cylinder gets higher values abruptly, reaching 0.85 for Reynolds number of 117 (V_R = 6.51), as shown in Fig. 7. Similar behavior occurs in the single cylinder case, but with a wider Reynolds number range.

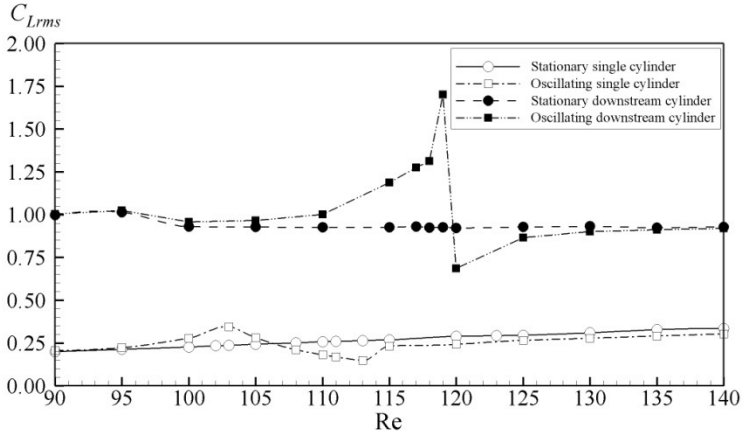


Figure 6: Root mean square of the lift coefficient versus Reynolds number for stationary and oscillating downstream cylinders in tandem arrangement and for fixed and oscillating single cylinders

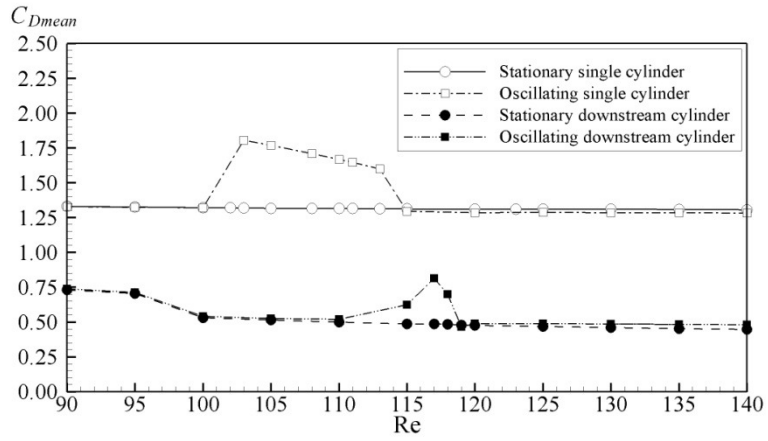


Figure 7: Mean drag coefficient versus Reynolds number for stationary and oscillating downstream cylinders in tandem arrangement and for fixed and oscillating single cylinders

Figure 8 shows the root mean square of lift coefficient and the mean drag coefficient versus Reynolds number for stationary upstream and oscillating downstream cylinders. It may be noticed that while the downstream cylinder experiments high variations of coefficients in the lock-in region, these coefficients decrease smoothly for the upstream cylinder. In this case, flow interference between both cylinders is different from the stationary downstream cylinder due to the motion of the downstream cylinder and changes in wake pattern and pressure field.

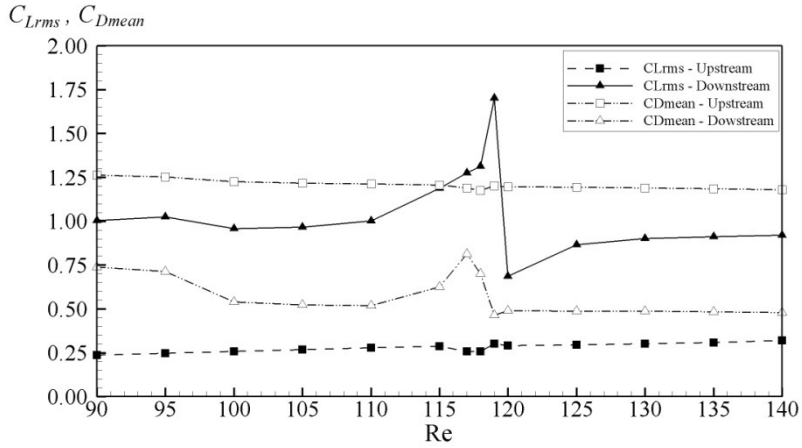


Figure 8: Root mean square of lift and mean drag coefficients versus Reynolds number for stationary upstream and oscillating downstream cylinders in tandem arrangement

Figure 9 shows the dimensionless amplitude (Y/D) and the vortex induced vibration frequency (f/f_n) for downstream cylinder and a comparison with the single one. It can be noticed that the resonance, for the single cylinder, occurred for Reynolds numbers between 102 (V_R = 5.68) and 113 (V_R = 6.29) case in which the vibration amplitude is higher and the vortex frequency is approximately equal to the natural frequency (f_n) of the dynamic system. In tandem arrangement with L/D = 5.25, the downstream cylinder experiments the resonance

in a shorter Reynolds number range, between 115 (V_R = 6.40) and 120 (V_R = 6.68). Furthermore, the maximum dimensionless amplitude is 0.721 for Re = 118 and this is much higher (more than 70%) than the one for the single cylinder (0.422 for Re = 103). The interference effects on tandem arrangement change the Strouhal number in relation to the single cylinder one and to the mode of interference between both cylinders, since the downstream cylinder is immersed in the periodic wake of the upstream one. This is why there are differences between both lock-in regions.

Figure 10 shows the temporal series of lift coefficient (C_L) and dimensionless vibration amplitude (Y/D) for Re = 118. Vorticity distributions and streamlines are shown in Fig. 11 where time positions of each picture are indicated in Fig. 10. It may be noticed that the lift coefficient presents higher non linearity by comparison with vibration amplitude which shows harmonic behavior. The phase angle between them is around 42° in this case. Figure 10 shows that the vortex shedding occurs for both cylinders, as expected for $L/D > 4$, and the wake behind the downstream cylinder is formed by the combination of vortex shed from both cylinders. The maximum vibration amplitude (Y/D = 0.721) occurs at instants shown in Fig. 11a and 11b.

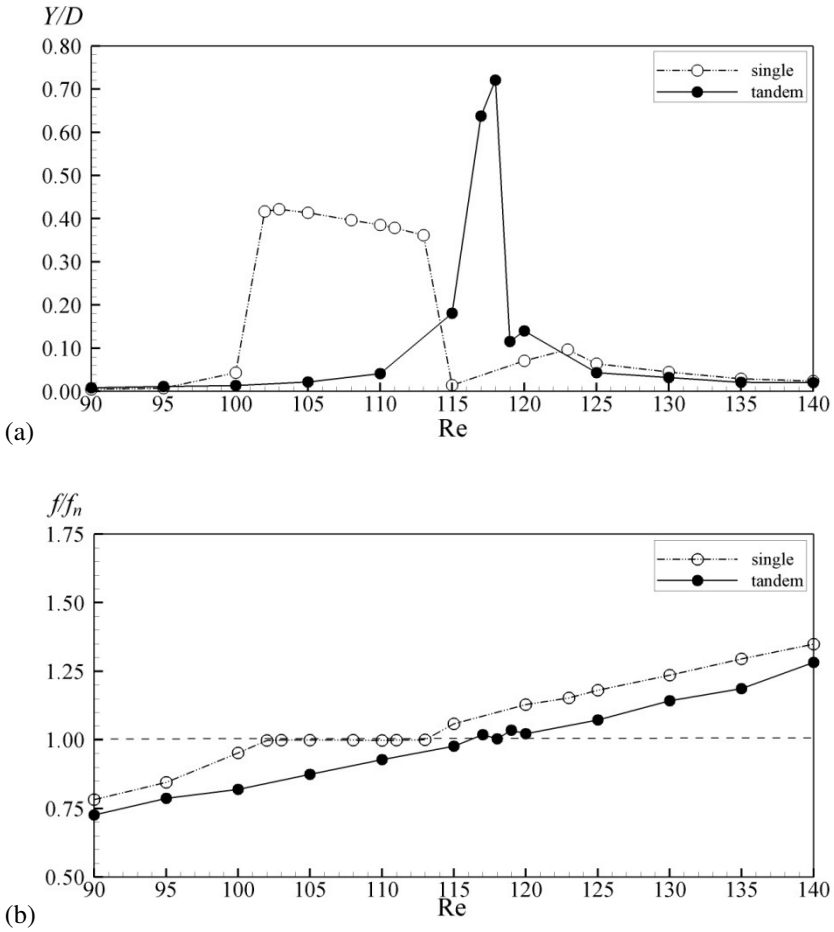


Figure 9: Dimensionless vibration amplitude (Y/D) (a) and vortex shedding frequency (f/f_n) (b) versus Reynolds numbers for single cylinder and cylinders in tandem arrangement

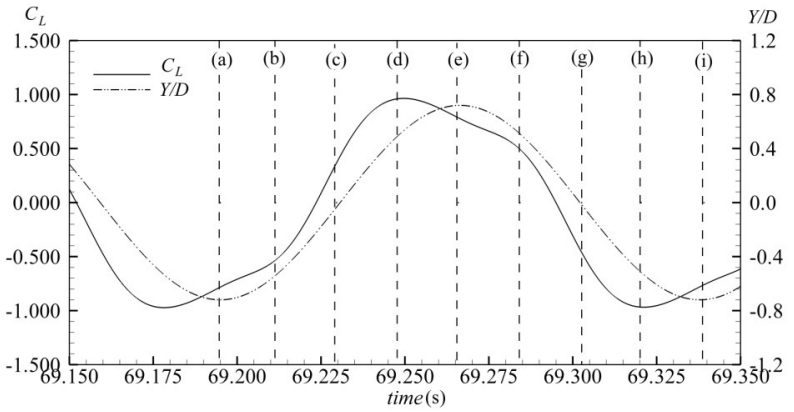


Figure 10: Temporal series of lift coefficient (C_L) and dimensionless vibration amplitude (Y/D) for Re = 118 (letters represent time positions of the vorticity distributions, shown in Fig. 10)

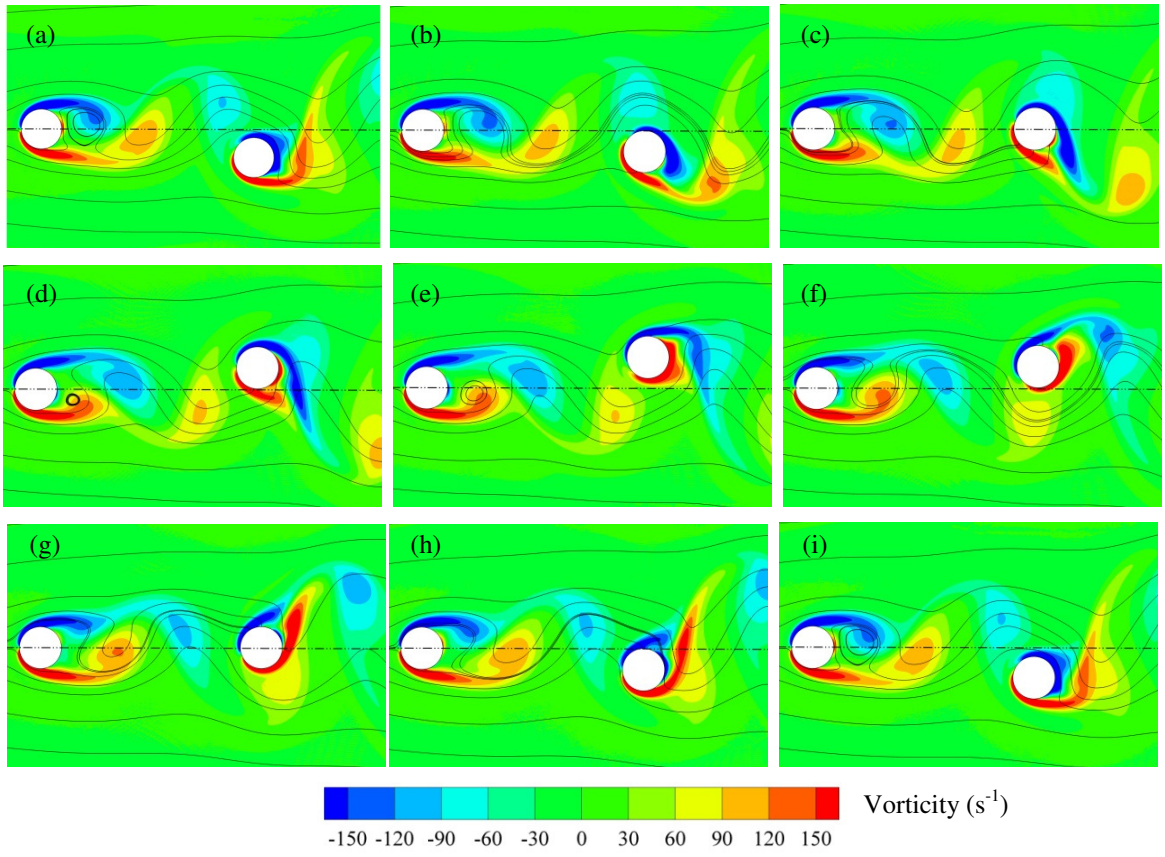


Figure 11: Vorticity and streamlines around cylinders for Re = 118 at nine instants (shown in the graph in Fig. 9) during one cycle

4 CONCLUSION

This paper presented a numerical analysis of a uniform flow over circular cylinders in tandem arrangement at low Reynolds numbers. Simulations were carried out by using Ifeinco

model which is based on the finite element method and employs the semi-implicit two-step Taylor-Galerkin method to discretize the Navier-Stokes equations and the Newmark method for the dynamic equation of the structure.

Ifeinco results (Strouhal number, root mean square lift and mean drag coefficients) had good agreement with those obtained by other authors for stationary cylinders, *L*/*D* from 1.5 to 6.0 and Re = 100. The critical spacing (*Lc*) equal to 4.0 *D* was obtained and abrupt variations in the flow parameters were observed in this case. Different flow behaviors were noticed when distances were shorter and longer than *Lc*. For $L < Lc$, vortices wake off the upstream cylinder and reattach to the downstream one while a symmetrical flow pattern occurs between both cylinders. For $L > Lc$, vortices shed off both cylinders.

Considering the center to center distance equal to 5.25*D* and the downstream cylinder elastically mounted in transversal direction, the lock-in region was captured by Reynolds number from 115 to 120; this range was different from and shorter than the one for the single cylinder (102 to 113). The maximum vibration amplitude, which was 0.721*D*, was higher than the one for the single cylinder (0.422*D*). At Re = 120 (in the lock-in region), the drag coefficient and the lift coefficient showed abrupt variations; the latter was much more intense. Complex interference occurs between both cylinders since the downstream cylinder is immersed in the periodic wake of the upstream one.

REFERENCES

[1] Igarashi, T. Characteristics of the flow around two circular cylinders arranged in tandem. *Bulletin of JSME* (1977) **24**(188):323-331.

[2] Zdravkovich, M.M. Review of flow interference between two circular cylinders in various arrangements. *ASME Journal of Fluids Engineering* (1977) **99**:618-633.

[3] Zdravkovich, M.M. The effects of interference between circular cylinders in cross flow. *Journal of Fluids and Structures* (1987) **1**:618-633.

[4] Sumner, D., Price, S.J. and Païdoussis, M.P. Flow-pattern identification for two-staggered circular cylinders in cross-flow. *Journal of Fluid Mechanics* (2000) **411**:263-303.

[5] Didier, E. Numerical simulation of low Reynolds number flows over two circular cylinders in tandem. *Conference on Modelling Fluid Flow*, Budapest, September 9-12, (2009) 347-354.

[6] Zdravkovich, M.M. Flow induced oscillations of two interfering circular cylinders. *Journal of Sound and Vibration* (1985) **101**(4):511-521.

[7] Brika, D. and Laneville, A. The flow interaction between a stationary cylinder and a downstream flexible cylinder. *Journal of Fluids and Structures* (1999) **13**:579-606.

[8] Assi, G.R.S., Meneghini, J.R., Aranha. J.A.P., Bearman, P.W. and Casaprima, E. Experimental investigation of flow-induced vibration interference between two circular cylinders. *Journal of Fluids and Structures* (2006) **22**: 819:827.

[9] Okajima, A., Yasui, A., Kiwata, T. and Kimura, S. Flow-induced treamwise oscillation of two circular cylinders in tandem arrangement. *International Journal of Heat and Fluid Flow* (2007) **28**:552-560.

[10] Huera-Huarte, F.J. and Gharib, M. Vortex- and wake-induced vibrations of a tandem arrangement of two flexible circular cylinders with far wake interference. *Journal of*

Fluids and Structures (2011) **27**:824-828.
[11] Li, J., Chambarel, A., Donneaud, M. and Martin, R. Numerical study of laminar flow past one and two circular cylinders. *Computers & Fluids* (1991) **19**:155-170.
[12] Slaouti, A. and Stanby, P.K. Flow around two circular cylinders by the random-vortex method. *Journal of Fluids and Structures* (1992) **6**:641-670.
[13] Mittal, S., Kumar, V. and Raghuvanshi A. Unsteady incompressible flows past two cylinders in tandem and staggered arrangements. *International Journal for Numerical Methods in Fluids* (1997) **25**:1315-1344.
[14] Meneghini, J.R., Saltara, F., Siqueira, C.L. and Ferrari, J.A. Numerical simulation of flow interference between two circular cylinders in tandem and side by side arrangements. *Journal of Fluids and Structures* (2001) **15**:327-350.
[15] Sharman, B.. Lien, F.S., Davidson, L. And Norberg, C. Numerical predicitions of low Reynolds number flows over two tandem circular cylinders. *International Journal for Numerical Methods in Fluids* (2005) **47**:423-447.
[16] Carmo, B.S. Estudo numérico do escoamento ao redor de cilindros alinhados. Master Thesis, Escola Politécnica da Universidade de São Paulo (2005).
[17] Carmo, B.S. and Meneghini, J.R. Numerical investigation of the flow around two circular cylinders in tandem. *Journal of Fluids and Structures* (2006) **22**:979-988.
[18] Carmo, B.S., Sherwin, S.J., Bearman, P.W. and Willden, R.H.J. Flow-induced vibration of a circular cylinder subjected to wake intereference at low Reynolds number. *Journal of Fluids and Structures* (2011) **27**:503-522.
[19] Mittal, S. and Kumar, V. Flow-induced oscillations of two cylinders in tandem and staggered arrangements. *Journal of Fluids and Structures* (2001) **15:**717-736.
[20] Papaioannou, G.V., Yue, D.K.P., Triantafyllou, M.S. and Karniadakis, G.E. On the effect of spacing on the vortex-induced vibrations of two tandem cylinders. *Journal of Fluids and Structures* (2008) **24**:833-854.
[21] Teixeira, P.R.F. and Awruch, A.M. Numerical simulation of fluid-structure interaction using the finite element method. *Computers & Fluids* (2005) **34**:249–273.
[22] Teixeira, P.R.F. and Awruch, A.M. Three–dimensional simulation of high compressible flows using a multi-time-step integration technique with subcycles. *Applied Mathematical Modelling* (2001) **25**:613–627.
[23] Argyris, J., Doltsinis, J.St., Wuestenberg, H. and Pimenta, P.M. *Finite element solution of viscous flow problems*. Finite Elements in Fluids. Wiley, New York, Vol. 6, 89–114, (1985).
[24] Donea J., Giuliani, S., Laval and H., Quartapelle, L. Finite element solution of the unsteady Navier-Stokes equations by a fractional step method. *Computer Methods in Applied Mechanics and Engineering* (1982) **33**:53–73.
[25] Bathe, K.J. *Finite element procedures*. Prentice-Hall, (1996).
[26] Fredsϕe, J. and Sumer, B.M. *Hydrodynamics around cylindrical structures*. World Scientific Publishing Co. Pte. Ltd., Singapore, Vol. 12, (1997).
[27] Gonçalves, R.A., Teixeira P.R.F. and Didier, E. Numerical simulations of low Reynolds number flows past elastically mounted cylinder. *Thermal Engineering* (2012) **11**(1-2):61-67.

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis (Eds)

NUMERICAL SIMULATION OF TSUNAMI FORCE ACTING ON A FLOATING/SUBMERGED TSUNAMI SHELTER

HIDEMI MUTSUDA[*], SHUNSUKE FUJII[*], ANDI ARDIANTI[*], YASUAKI DOI[*]

[*] Dept. of Energy and Environmental Engineering
Hiroshima University
1-4-1 Kagamiyama Higashi-Hiroshima, 739-8527, Hiroshima, Japan
e-mail: mutsuda@hiroshima-u.ac.jp

Key words: Tsunami, Shelter, Mooring, Particle based method, SPH.

Abstract. After the Tohoku earthquake tsunami in Japan, many evacuating options from huge tsunami, e.g. tsunami tower, large building, moving onto hill, super-sized breakwater, underground shelter and small lifeboat, have been proposed in recent years. This disaster is one of the strongly nonlinear coupled interactions problem between wave and structures. We have proposed and developed one kind of a large-sized tsunami shelter with mooring that is capable of accommodating at least one hundred people or more to evacuate from run-up tsunami. Using the particle method, SPH, to evaluate the optimized configuration for reducing tsunami force and its motions due to tsunami attacking with breaking, the present work has proposed and developed one of tsunami shelter, "Ellipsoid type" including electric device and storage system. The size of "Ellipsoid type" is 75m length, 20m width, 9m height in practical use. The wall at the front and back face can be constructed by pre-stressed concrete to protect from several kinds of debris such as floating car, wood and ship. The "Ellipsoid type" of tsunami shelter can vary smoothly tsunami flow going through downstream and then it could be one of useful options evacuating from tsunami disaster. The result was validated with experimental results in tsunami attacking on the shelter near coastal region.

1 INTRODUCTION

In March 2011, the huge tsunami hitting on the Tohoku region, northern part of Japan, caused the great damage in not only many infrastructures but also many people, about 16,000 for died, 2,500 for missing and 6,000 for injured. Another huge earthquake and tsunami categorized in mega class will take place in NANKAI trough of the Pacific Ocean near Japan coast. Therefore many evacuating options for tsunami disaster have been proposed and developed in practical use since 2011. There are so many tsunami options such as tsunami tower, super-sized breakwater, moving to hill, underground shelter and coastal tsunami shelter on the ground, see in [1-3]. We have to consider some needs in hazard region and autonomous characteristics. It is important to select options based on them as rationalized measures against huge tsunami. For instance, it is necessary to protect a public safety with understanding information such as composition of population, locations of the main hospital and their capacity, support service especially for the aged and small children and industrial structure including fishery.

Many types of tsunami shelter have been proposed for few people and hundreds people in order to be available to evacuate them from tsunami in recent year. Some of them are floating or drifting type storing oxygen cylinder and some foods for surviving. However we should not ignore not only safety driving of those kinds of shelters but also influence on floating/drifting motions to other infrastructures because the shelters can be pushed away and freely drifting with no-control. Therefore the shelter with mooring have been also proposed to keep the position itself. It is difficult for the mooring system to design its length and strength which is dependent of tsunami water level and velocity on the ground.

Based on these background as the mentioned above, in the previous research [4,5], a new type of tsunami shelter whose capacity is hundreds people, has been proposed and developed to be able to evacuate from huge tsunami with easy access and safety. The proposed shelter can keep its position with mooring and it can also control floating/submerging to avoid more than 10m higher tsunami water level and strongly velocity due to run-up tsunami. Based on numerical simulation using SPH [6] and experiment, the purpose of this study is to compute fluid structure interaction between run-up tsunami and tsunami shelter and to optimize design of the shelter in order to reduce its motions and tsunami force.

2 FLOATING/SUBMERGED TSUNAMI SHELTER

A new type of tsunami shelter, floating/submerged type, has been proposed in our previous works [4, 5] as shown in Fig.1. The floating/submerged tsunami shelter has an important concept that is not only to protect tsunami attacking with highly water level and strongly velocity but also to avoid from them. This study focuses on a large sized shelter on the land near coastal area because the capacity is enough for large population and many equipment for surviving. Using mooring and ballast water, the shelter can be floated on tsunami water surface when tsunami water level is low less than 3-5m and it can be submerged under tsunami wave when the water level is more than 5 to 10m in huge tsunami condition. After tsunami hitting, the floating/submerged tsunami shelter can be used as temporary accommodation and emergency medical center. Under normal condition, this shelter can be usually utilized as communication space for living people and storage space for fish and marine products.

Figure 2 shows one example of some arrangements of equipment and seats in the shelter. The front and back face are protected by double hull structure which is normally applied for ship building. There are many prepared facilities consists of emergency medical system, oxygen cylinder, life support system, food storage area, power generation and its battery, communication network system and so on for surviving during a few weeks or more. The windows and doors are stronger for high pressure and watertight. The evacuated people should be supported by seatbelt during tsunami forcing. In normal condition, the seats are put back under the floor and wider open-space can be kept to be community zone.

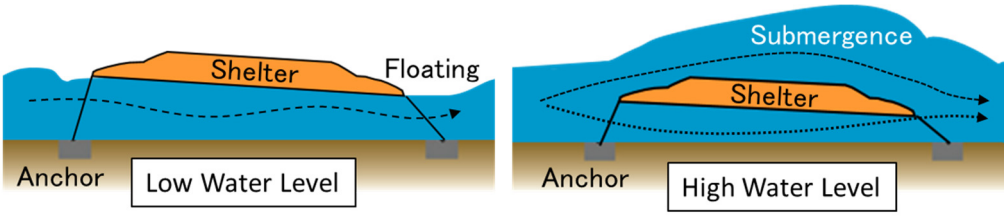
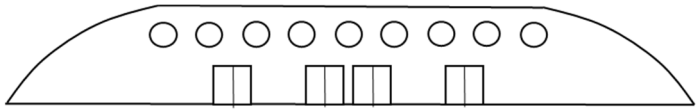
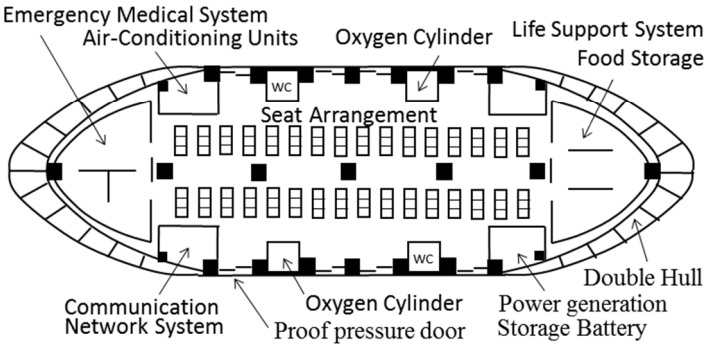


Figure 1: Concept of the floating/submerged tsunami shelter



(a) Side View



(b) Top View

Figure 2: Arrangement of equipment and seats in tsunami shelter

3 PARTICLE BASED METHOD

3.1 Governing Equations and computational method

To compute tsunami force and impact pressure, numerical work was conducted using particle based method, Smoothed Particle Hydrodunamics, SPH [6] which can track breaking wave and splashing after colliding on tsunami shelter. The gaverning equations for all phases can be discretized based on SPH method.

In this study, the governing equations for fluid phase consist of the mass conservation equation, the incompressible Navier-Stokes equation and the equation of continuity. The equations are expressed as follow:

$$\frac{\partial u_i}{\partial x_i} = 0 \tag{1}$$

$$\frac{\partial u_i}{\partial t} + u_j \frac{\partial u_i}{\partial x_j} = -\frac{1}{\rho}\frac{\partial P}{\partial x_i} + \frac{\mu}{\rho}\frac{\partial^2 u_i}{\partial x_j \partial x_j} + g_i + F_{fsi} \tag{2}$$

where ρ is the density, u_i is the velocity, P is the pressure, g_i is the gravity acceleration, F_{fsi} is the fluid structure interaction term.

The governing equations are solved using the splitting method as well-known conventional multiphase technique. The pressure including jump conditions caused by density between different phases can be solved by the Poisson equation given by

$$\nabla \cdot \left(\frac{\nabla P^{n+1}}{\rho^*} \right) = \frac{\nabla \cdot u^*}{\Delta t} \tag{3}$$

where * denotes the physical value after the advection step. The pressure of solid-phase can be obtained by this equation and it can be also applied for solving solid deformation if necesarry.

The governing equations for solid phase are the continuity equation and momentum equation as follows:

$$\frac{D\rho}{Dt} + \rho \frac{\partial u^i}{\partial x^i} = 0 \tag{4}$$

$$\rho \frac{Du^i}{Dt} = \frac{\partial \sigma^{ij}}{\partial x^j} + g^i - F^i_{fsi} \tag{5}$$

where ρ is the density, u^i is the velocity, σ^{ij} is the stress tensor for solid phase, and F_{fsi} is the fluid structure interaction term as the same in eq.(2) for fluid phase. The stress tensor in eq.(5) is given by

$$\sigma^{ij}_s = -P\delta^{ij} + S^{ij} \tag{6}$$

where S^{ij} is the deviatoric stress tensor and P is the pressure solved by eq.(3).

The model can consider large deformation of elastoplastic body if necessary. The solid body changes at every calculation step by using the following equation:

$$\{dS^{ij}\} = [D^{ep}]\{d\varepsilon^{ij}\} \tag{7}$$

where D^{ep} is the elastoplastic matrix, $d\varepsilon^{ij}$ is the strain increment for a time, and dS^{ij} is the deviatoric stress increment for the time.

To compute rotations of solid phase during motions, the Jaumann derivative is employed to ensure material frame indifference with respect to rotation as follow:

$$\frac{dS^{ij}}{dt} = 2\mu\left(\dot{\varepsilon}^{ij} - \frac{1}{3}\delta_{ij}\dot{\varepsilon}^{ij} \right) + S^{ik}\Omega^{jk} + \Omega^{ik}S^{kj} \tag{8}$$

where ε is the strain rate tensor and Ω is the spin tensor.

The fluid structure interaction F_{fsi} is computed by acceleration obtained from pressure on particles. In the model, the fluid structure interaction F_{fsi} in Eqs.(2) and (5) can be given by the following equation:

$$F_{fsi}(\mathbf{r}_a) = -\frac{1}{\rho(\mathbf{r}_a)}\sum_b m_b \frac{P(\mathbf{r}_b)}{\rho(\mathbf{r}_b)} \nabla_a \cdot W(\mathbf{r}_a - \mathbf{r}_b, h) \tag{9}$$

where W is the kernel function, m is the mass, P is the pressure, ρ is the density and h is the referenced area where interaction between particles can be considered.

In the model, the shelter model consists of SPH particles to compute motions in 3D. Therefore, the 3D motions of the shelter is represented by describing translation and rotation of the center of gravity of the shelter using the following equations:

$$\frac{\partial^2 x_{s,k}}{\partial t^2} = \frac{F_{s,k}}{m_i} - F_{fsi} \tag{10}$$

$$I\frac{\partial \omega_i}{\partial t} = T_i \tag{11}$$

$$\frac{\partial \theta_i}{\partial t} = \omega_i \tag{12}$$

where θ_i is the rotational angle, ω is the angular velocity, T_i is the torque, I is the inertia moment, and F is the fluid structure interaction. In addition, the center of gravity of the shelter can be obtained by solving the inertia moment of particles, and this is calculated by using Baraff theory[7]. Based on this theory in the model, equations for 3D motion are given by

$$\mathbf{r}_g = \frac{1}{N}\sum_{i=1}^{n} \mathbf{r}_i \tag{13}$$

$$I = \sum_{i=1}^{N} m\left|\mathbf{r}_i - \mathbf{r}_g\right|^2 \tag{14}$$

where N is the number of particles for shelter, $\mathbf{r}_g$ is the position of the gravity center, I is the inertia moment, $\mathbf{r}_i$ is the position of the i th particle and m is the mass of the particle. The inertia moment can be set at initial condition. Therefore, the coordinates of velocity of each particle in every time step can be tracked by using rotation matrix and the amount of angle rotation of the center of gravity to avoid the Gimbal lock phenomenon, and the quaternion is used instead of the rotation matrix. The numerical model can be enhanced for applying to shelter motions, overturning and sinking caused by nonlinear wave with breaking.

To keep computational efficiency and stability, the time increment for solid phase is approximately 1/10 to 1/50 of that for fluid phase. In this research, tsunami shelter can be supported by pole to simulate the experimental work. The tsunami shelter can be assumed as solid model in this work. In future work, motions of the tusnami shelter and its elastic deformation would be considered for practical design. Other details can be referred in Mutsuda et al. [8].

3.3 Shelter model

Tsunami force would be increasing with attacking angle to tsunami shelter when separation flow with vortex around tsunami shelter could occur in larger deadrise angle. To reduce tsunami force on tsunami shelter, based on the previous work [4, 5], the ellipsoidal tsunami shelter is focused in this study. Iida et al. [9] has verified that super-express in Japan, it called SHINKANSEN, has been optimized to investigate aerodynamics. Based on the results, especially the aspect ratio was set to be from 3 to 5, in front face and back one to reduce pressure resistance caused by tsunami attacking. Moreover the side face of the tsunami shelter is formed by the curved surface with streamline to control separation flow behind the tsunami shelter. Figure 3 shows tsunami shelters with different aspect ratio to compare tsunami force acting on them. The Ellipsoid3 (E3) and Ellipsoid3s (E3s) have aspect ratio 1/3 at the front face and back one. In Ellipsoid35 (E35) and Ellipsoild35s (E35s), the aspect ratios 1/3 and 1/5 are combined. The side face in E3s and E35s is also formed by streamline. The Box type, which is a typical building, is employed to compare pressure resistance with them.

3.4 Computational conditions

The computational domain is shown in Figure 4. To reduce computational time and cost, the tsunami conditions before breaking and run-up, such as surface profile, internal velocity and pressure, can be computed by Boundary Element Method [10] using exact solitary wave condition. These numerical results are set as the initial conditions in SPH.

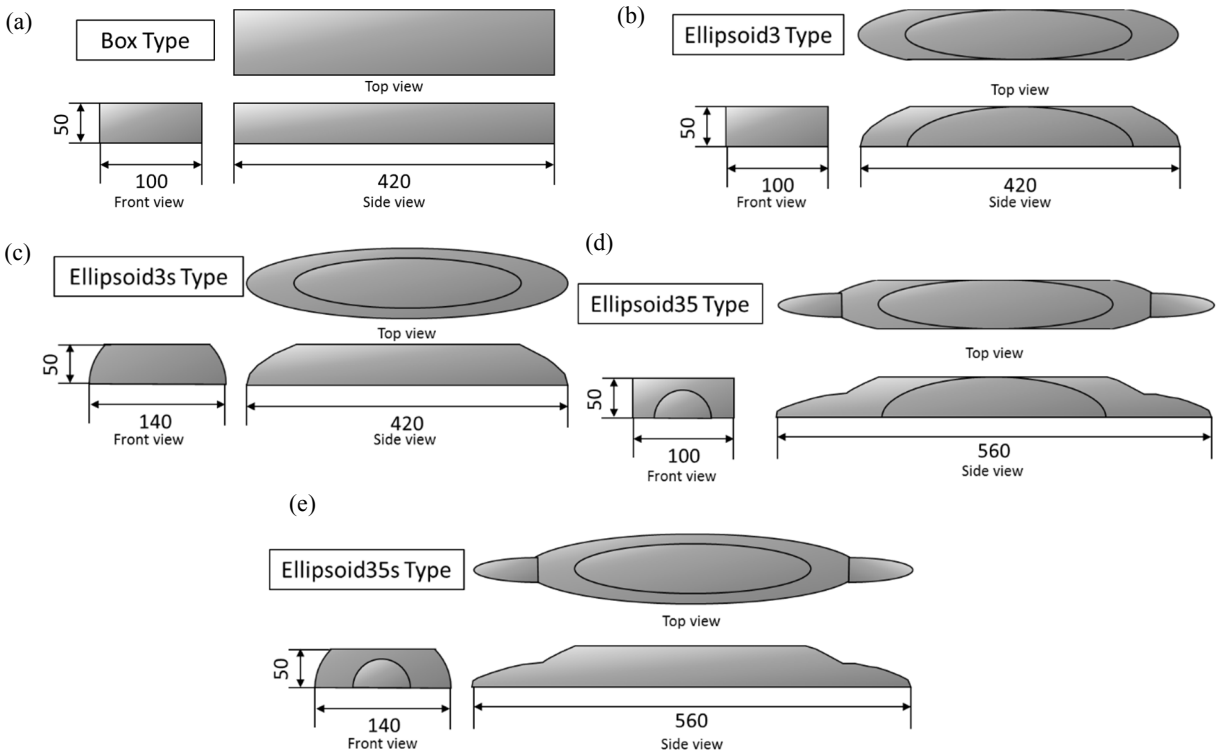


Figure 3: Configuration of tsunami shelter with different aspect ratio (unit:mm)

The tsunami shelters can be assumed as solid phase without deformation in this study. The time increment is 0.001s and the particle diameter is set to be 1cm. The attack angles are set to be 0, 15, 30, 45, 90 deg. to tsunami wave direction. The tsunami height condition corresponds to a fully developed bore propagating on the ground.

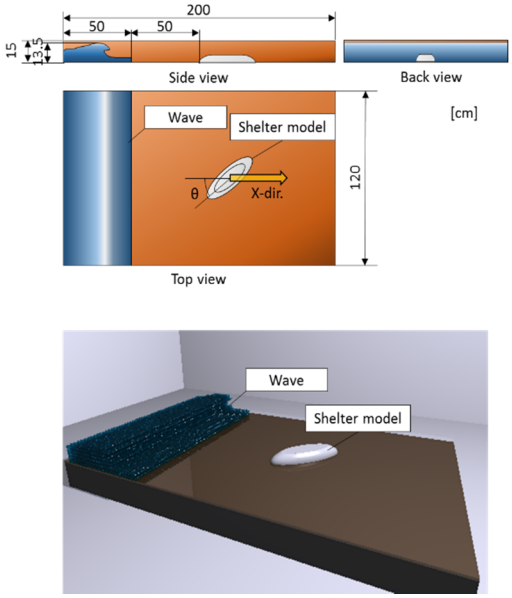


Figure 4: Computational domain

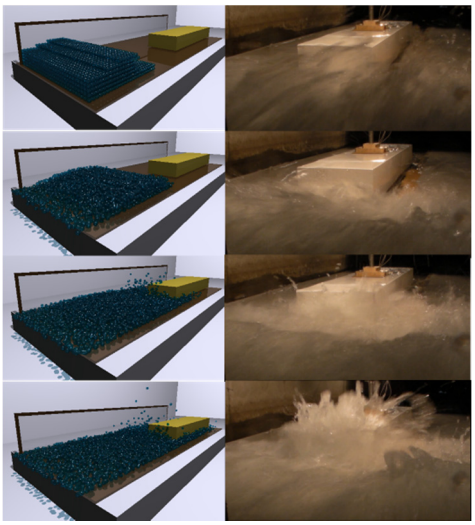


Figure 5: Numerical result (left) of tsunami attacking on box shelter with experimental result (right)

4 NUMERICAL RESULTS

4.1 Tsunami wave motion and validation

Figure 5 shows one example of snapshots tsunami phenomena acting on the box shelter with experimental result. It can be found that the strongly splashing occurs on the front face of the box shelter and it causes highly tsunami force and impact pressure.

Figure 6 and 7 show tsunami wave propagation with the attack angles 0 and 45 deg. in Box and E35s. Figure 8 shows comparison of time history of tsunami force acting on Box type for the attack angle 0 deg. with experimental result. The numerical result is in good agreement with the experimental one and it is reasonable for evaluating fluid force on tsunami shelter.

4.2 Reduction of tsunami force and influence on tsunami attacking angle

Maximum pressure and averaged tsunami force are examined and their characteristics are clarified in this section. The force can be calculated by summation of tsunami pressure acting on shelter in wave direction.

Figure 9 shows time histories of tsunami force in the attack angle 0, 45 and 90 deg. to compare with the normal type, Box. Figure 10 and 11 shows relationships between maximum tsunami force, averaged tsunami force and attack angle, H is the tsunami water level on the

ground, ρ is the water density and g is the gravity acceleration.

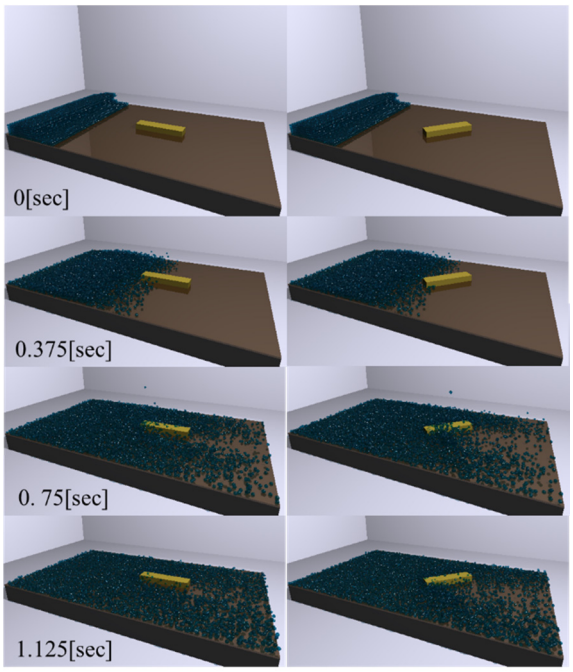


Figure 6: Box type (Left: 0 deg. Right: 45deg.)

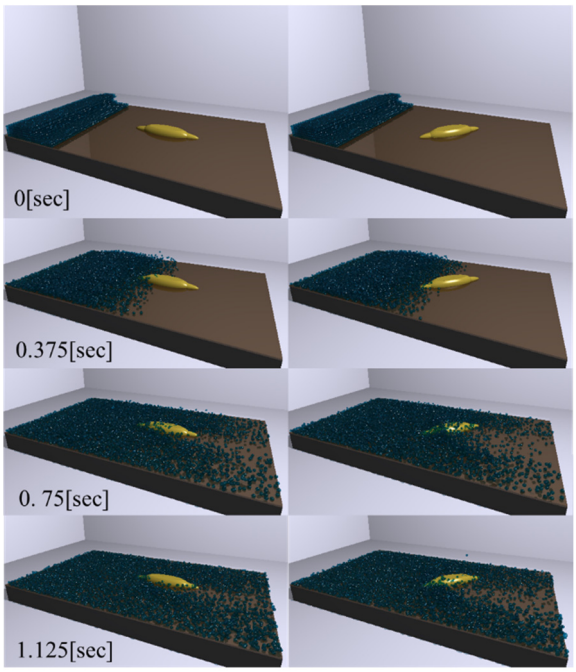


Figure 7: Ellipsoid35s type (Left: 0 deg. Right: 45deg.)

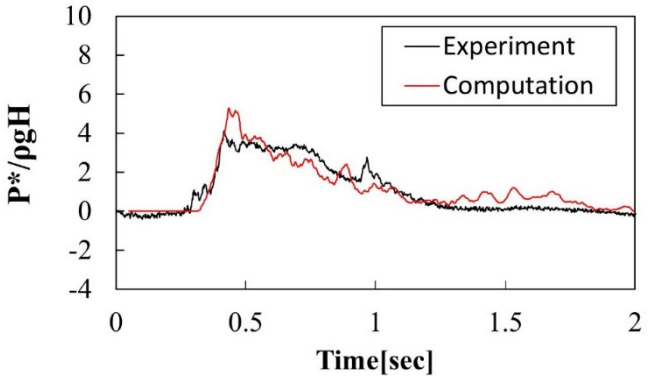
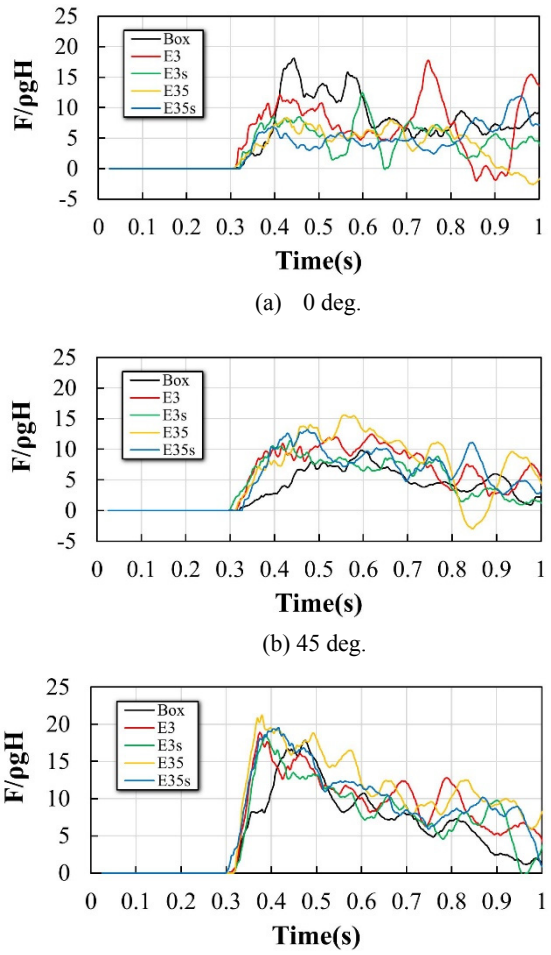


Figure 8: Comparison of time history of pressure between computational result and experimental one in Box type

In the Box shelter, the tsunami force is largest in the attack angle 0 deg. and then it is decreasing until the attack angle is close to 30 deg. The tendency is also increasing over the angle 45 deg. This is because the incoming tsunami flow into the corner of the shelter can be separated along the side face of the shelter when the angle is larger. Then the resultant tsunami force is decreasing. In over the attack angle 45 deg. the tsunami force is stronger at the side face as the projected area increasing. On the other hand the averaged tsunami force is decreasing in E3s and E35s having the curved surface on the side face. The curved surface can avoid from strongly tsunami force because the separated flow on the front face can be controlled.

To evaluate and examine effect on reduction of tsunami force, comparison of maximum tsunami force and averaged tsunami force are compared as shown in Fig.12 and 13. These data are obtained by averaging tsunami force at each attack angle. The most useful for reducing tsunami force in the maximum and averaged tsunami forces, is E35s with the largest aspect ratio on the front face and streamline on the side face among the shelters

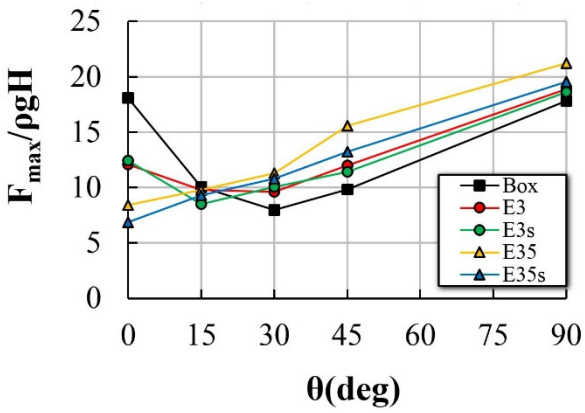


(a) 0 deg.

(b) 45 deg.

(c) 90 deg.

Figure 9: Time histories of tsunami force acting on tsunami shelter



Figire 10: Relationship between maximum tsunami force and attack angle of incident tsunami wave

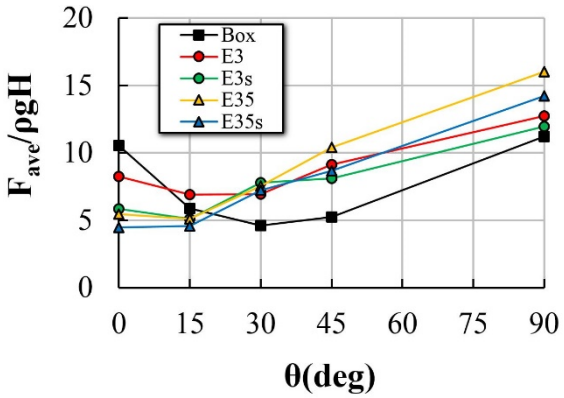


Figure 11: Relationship between averaged tsunami force and attack angle of incident tsunami

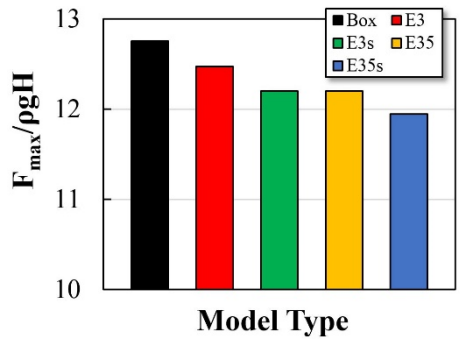


Figure 12: Comparison of maximum tsunami force

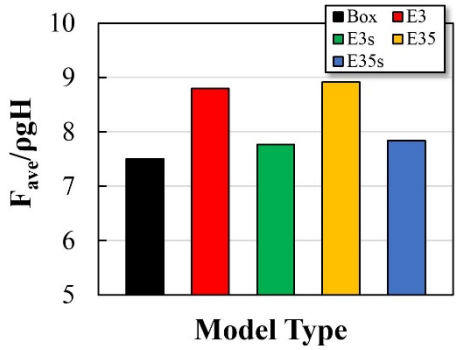


Figure 13: Comparison of averaged tsunami force

5 EXPERIMENTS

In the previous chapter, the shelter, E35s, with large aspect ratio on the front face and streamline on the side wall, is available for reducing tsunami force. The motions of the shelter, E35s with mooring are examined in experiment.

5.1 Experimental setup and conditions

Figure 14 shows overview of experimental tank (40m length × 1.2m width × 2.5m depth) and some equipments to measure tsunami force. The valiable angle 1/3 to 1/100 of the sea bottom can be set up to generate several kinds of tsunami conditions. The flat bottom as the ground was located to be available for propagating run-up tsunami. The shelter supported by four wires and pulley was located on the turn table at the bottom. The turn table can be rotated to exchange the attack angle from 0 to 45 deg. as shown in Fig.15. The tsunami force can be measured by dynamometer and tension meter connected with the wires. To reduce snap-force due to impulsive pressure, spring with constant coefficient 210N/m was employed in this mooring system. The wireless motion sensor inside the shelter can obtain accelerations of the shelter in 6 degree of freedom. The incident tsunami wave height can be measured at WG1 and the run-up wave height is also captured at WG2 located on the strandline. The incident wave period in regular wave is 2s to avoid from water distarbance caused by previous tsunami wave.

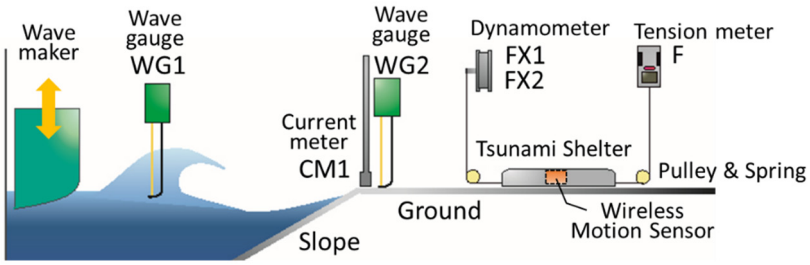


Figure 14: Experimental setup

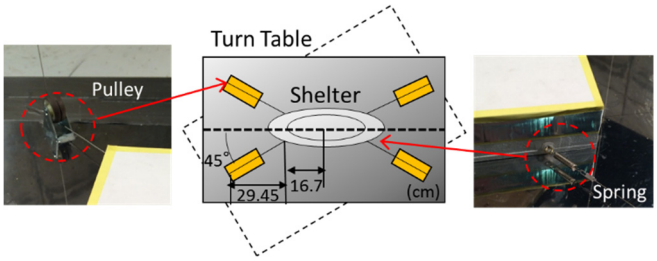


Figure 15: Mooring setup of tsunami shelter (Top view)

5.2 Shelter motions

Figure 16 and 17 show comparison of surge motion A_x/g (positive: downstream side) and pitch motion (positive: uplift of the front face) in the case of the attack angle 0 deg. The

wave height H is normalized by the shelter width d. The surge motion in E3s and E35s is lower than that in Box type. All of the values are less than 0.5g in the wave condtions, which means safety level for evacuating people. On the other hand the pitch motion is the same tendency in all wave conditions. The averaged pitch angle is approximatly 3 to 4 deg. The ellipsoid type has no advantage in pitch motion comparing with Box type. Therefore it should be necessary in practicale use to reduce the pitch motion after modifying the shleter. Figure 18 shows comparison of yaw motion in the case of the attack angle 45 deg. The yaw motion in the ellipsoid type, E3s and E35s, is lower than that in Box type. This tendency is remarkable as the wave height is larger. It can be seen that the ellipsoid type can reduce not only tsunami force but also shelter motions caused by tsunami attacking.

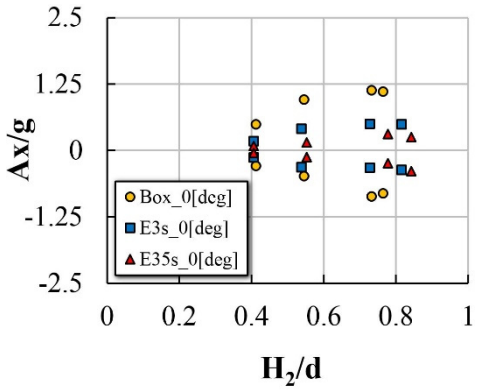


Figure 16: Comparison of surge motion (A_x/g) versus tsunami wave height (H2/d)

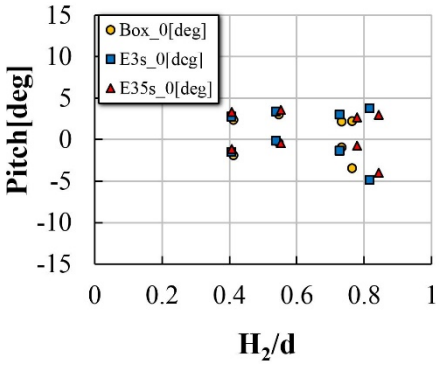


Figure 17: Comparison of pitch motion (deg.) versus tsunami wave height (H2/d)

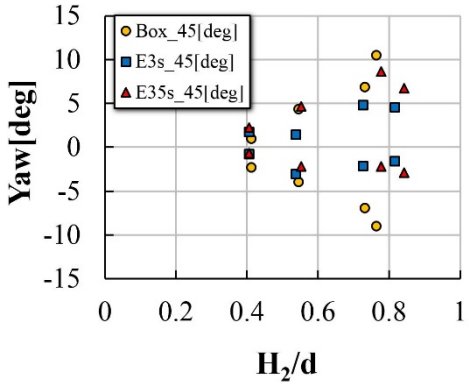


Figure 18: Comparison of yaw motion (deg.) versus tsunami wave height (H2/d)

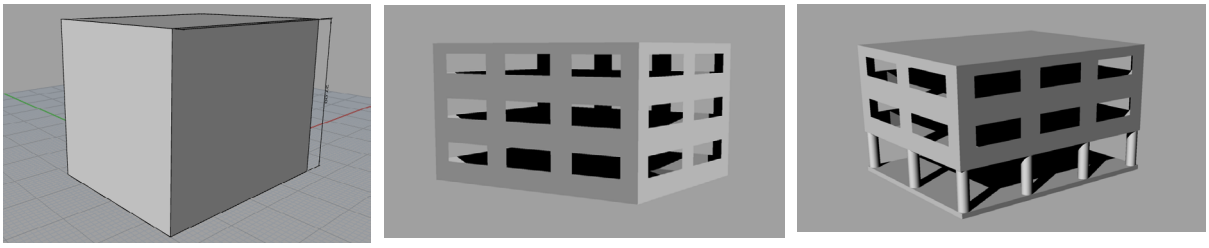
6 FURTHER RESEARCHES

One of further researches is to develop a tsunami tower building to avoid from tsunami attacking. The concept can be applied with only minor change for the existing building to reduce construction cost comparing with that for tsunami shleter. This could be a practical

measure as one tsunami option near coastal area where there are many existing building and the people living there.

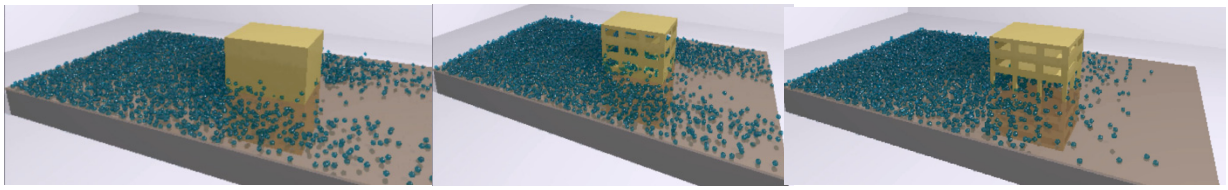
This study have proposed two modified buildings with opening parts optimized in arrangement and their ratio to the projected area and no-wall at the first floor like a piloti as shown in Fig.19 and their numerical results are shown in Fig.20. Figure 21 shows comparison of averaged tsunami force acting on the tsunami tower building. It can be seen that the opening parts and piloti are one useful minior change for the exsisting building. Their reduction rates are around 30% or more.

Figure 22 shows one more application for arrangement of building including tsunami tower or shelter to reduce tsunami force. Unpredictable tsunami force caused by the surrounding buildings and accommodations can be occured in populated zone.The snapshots are one example arrangement in coastal area. Some of the small buildings behind the larger building can escape from directly tsunami attacking. This result means that it should be considered to make appropriate arrangement of building and accomotation in populated zone near coastal line in order to reduce tsunami forcing and impulsive pressure.



(a) Normal (Type A) **(b)** Opening parts (Type B) **(c)** Piloti (Type G)

Figure 19: Tsunami tower building



(a) Normal (Type A) (b) Opening parts (Type B) (c) Piloti (Type G)

Figure 20: Tsunami tower building

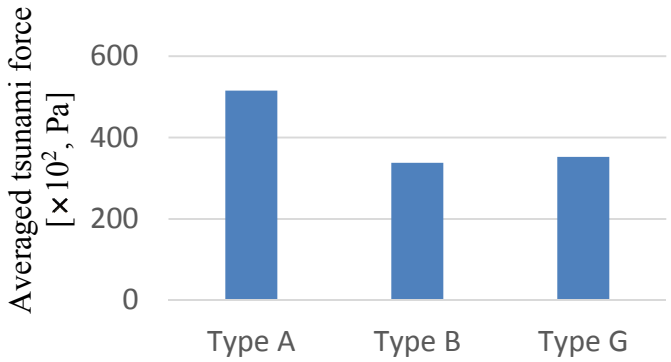


Figure 21: Comparison of averaged tsunami force acting on tsunami tower building

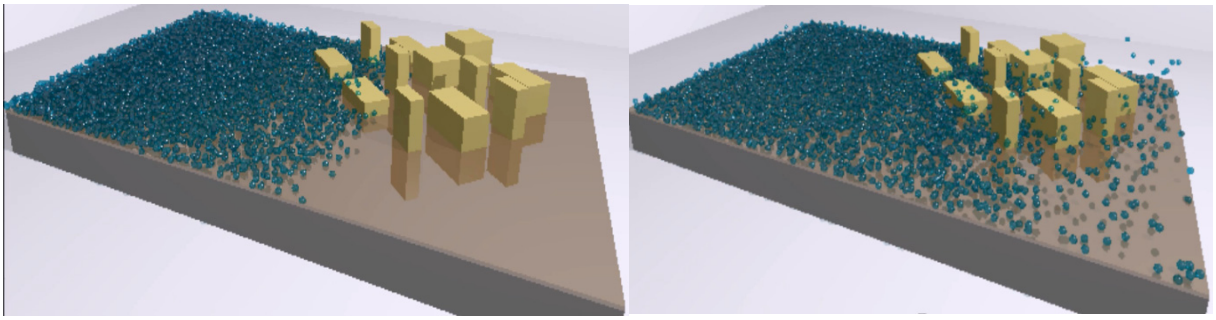


Figure 22: Snapshots of tsunami attacking on many building near coastal area

6 CONCLUSIONS

This study has proposed and developed the floating/submerged tsunami shelter to evacuate huge tsunami attacking and then has also investigated tsunami force acting on the shelter and its motions caused by tsunami impact force in particle based method and experimental work. The main results can be summarized as follows.

The ellipsoidal tsunami shelter (E35s type) can drastically reduce tsunami force and impact pressure. Moreover the surge and yaw motions in E35s with mooring system type can be also reduced comparing with Box type. In further research, the tsunami tower building with opening parts and piloti is also useful as one of tsunami evacuating options. The tsunami shelter and tower building should be arranged to reduce unpredictable tsunami force caused by a surrounding building and accommodation.

In future effort, mooring system including elastic wire will be optimized and effects on drift of floater such as wood, broken house ship and car, will be also investigated.

ACKNOWLEGEMENT

The part of the research was supported by Grant-in-Aid for Challenging Exploratory Research and adaptive and seamless technology transfer program (A-STEP) of Japan Science and Technology Agency. The authors express thanks to the supports.

REFERENCES

[1] Brahman industries, LLC, STATIM tsunami shelter system, http://www.statimshelter.com/, (2012).

[2] Government of Japan, the cabinet office, First report of disaster prediction due to the Na-Kai trough earthquake, http://www.cao.go.jp/, (2012).

[3] N.C.P corporation. "Tsunami capsule, "Noah", http://newcosmopower.com/, (2012).

[4] Mutsuda,H., M. Kamada, S. Fujii and Y. Doi. Development of Floating Tsunami Shelter with Mooring, Proceedings of the 23rd International Society of Offshore and Polar Engineers(ISOPE), pp.60-67, (2013).

[5] Mutsuda Hidemi, Masaya Kamada, Syunsuke Fujii and Yasuaki Doi. Reduction of Tsunami Force Acting on Large-sized Tsunami Shelter, The Coasts & Ports 2013 Conference, CD-R, Sydney, (2013).

[6] Gingold, R.A., and Monaghan, J.J., "Smoothed particle hydrodynamics, theory and

application to non-spherical stars, Mon. Not. Roy. Astr. Soc., Vol. 181, pp.375-389, (1977).

[7] Barraf, D.: "An introduction to physically based modeling: Rigid body simulation I ~Unconstrained rigid body dynamics~", SIGGRAPH'97 course note, D3 (1997).

[8] Mutsuda Hidemi, Yu SHIMIZU and Yasuaki Doi. Numerical Study on Interaction between Violent Wave and Structure using SPH, Particle-Based Methods, - Fundamentals and Applications, PARTICLES 2009, pp.266-269, ISBN: 978-84-96736-82-5, (2009).

[9] Iida M., Optimization of Train Nose Shape for Reducing Impulsive Pressure Wave from Tunnel Exit, Transactions of the Japan Society of Mechanical Engineers, Series B, Vol.62, No.596, pp.1428-1435 (1996) in Japanese.

[10] Yasuda, T., H. Mutsuda and N. Mizutani : Kinematics of overturning solitary waves and their relations to breaker types, Coastal Engineering, Vol.29, No.4, Elsevier, pp.317-346, (1997).

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis(Eds)

THREE-NODE ZERO-THICKNESS HYDRO-MECHANICAL INTERFACE FINITE ELEMENT FOR GEOTECHNICAL APPLICATIONS

B. CERFONTAINE*, A.C. DIEUDONNE*,†, J.P. RADU*, F. COLLIN* AND R. CHARLIER*

*Department ArGEnCo
University of Liege
B52, ULg, Liege, Belgium
e-mail: f.collin@ulg.ac.be

† FRIA - F.R.S. F.N.R.S
Brussels, Belgium

Key words: Interface, Numerical modelling, Finite elements, Offshore Engineering

Abstract. The paper presents the main features of a hydro-mechanical coupled finite element of interface. The mechanical problem accounts for the detection of contact, the development of contact pressure, shearing and relative sliding between two solids. A three-node discretisation of hydraulic problem allows the representation of fluid flow across and in the plane of the interface. The method involves a drop of pressure between each side of the interface and the inner medium. Hydro-mechanical couplings result from 1) the definition of the total pressure acting on each side of the interface according to the Terzaghi's principle; 2) the dependence of permeability on the gap opening; 3) the variation of the fluid mass stored within the gap.

1 INTRODUCTION

Interfaces play a crucial role in many fields of geotechnical engineering and engineering geology. Understanding their mechanism is necessary to deal with pile driving and design [1], the behaviour of faults in the vicinity of hydrocarbon production wells [2] and carbon dioxide geological sequestration [3], among others. In all these applications, hydro-mechanical couplings should be taken into account.
Suction caissons or bucket foundations are more and more installed as permanent foundations for offshore structures [4]. They consist of steel cylinders open towards the bottom and installed into the soil by suction. The role of interfaces is particularly crucial for their modelling, especially under pull loading [5, 6, 7].
Finite elements of interface were early developed [8] especially for the purpose of metal

forming [9]. However recent advances develop coupled finite elements taking into account fluid or gas phases [2, 10]. The main purpose of this work is to develop a finite element of interface able to reproduce the coupled hydro-mechanical behaviour in three dimensions and to apply it to the uplift modelling of a suction caisson. The element is implemented in the finite element code LAGAMINE, able to carry out fully coupled simulations [11].

2 INTERFACE ELEMENT

The finite element of interface involves two distinct but related issues: the mechanical and the flow problems. The mechanical problem tackles the detection of contact between two solids, the evolution of shearing and/or sliding. The flow problem describes fluid flows taking place inside and through the interface. These two problems are inherently interrelated. Fluid flow influences the pressure acting on each side of the interface. The mechanical opening/closing of the interface modifies its permeability and the mass of water stored inside. The extended definition of the interface element can be found in [12].

2.1 Mechanical problem

Let us consider two deformable porous media Ω^1 and Ω^2 in their current configuration in the global system of coordinates $(\mathrm{E}_1, \mathrm{E}_2, \mathrm{E}_3)$, as shown in Figure 1. In each point x_1 of the boundary Γ_c^1 where the contact between them is likely to happen, a local system of coordinate $(\mathbf{e}_1^1, \mathbf{e}_2^1, \mathbf{e}_3^1)$ can be defined such that $\mathbf{e}_1^1$ is normal to the boundary. The gap function g_N measures the distance between both solids. It is computed according to

$$\mathrm{g}_N = \left(\mathbf{x}^2 - \mathbf{x}^1\right) \cdot \mathbf{e}_1^1, \tag{1}$$

which is the closest-point projection of $\mathbf{x}^2$ onto $\mathbf{x}^1$. The gap function is generalised to each local tangential direction. However this definition has no meaning in the field of large displacements and the gap variation is defined instead [9] such that

$$\dot{\mathbf{g}} = \dot{g}_N \, \mathbf{e}_1^1 + \dot{g}_{T1} \, \mathbf{e}_2^1 + \dot{g}_{T2} \, \mathbf{e}_3^1. \tag{2}$$

When solids come into *ideal* contact, $\mathrm{g}_N = 0$, they deform and develop a contact pressure p_N since they cannot overlap each other. This condition is mathematically termed *contact constraint* or Hertz-Signorini-Moreau condition [13], it mathematically reads

$$\mathrm{g}_N \geq 0, \qquad \mathrm{p}_N \geq 0 \qquad \text{and} \qquad \mathrm{p}_N \, \mathrm{g}_N = 0. \tag{3}$$

If there is no contact, $\mathrm{g}_N > 0$ and contact pressure p_N is null. If contact takes place, $\mathrm{g}_N = 0$ and a contact pressure develops $\mathrm{p}_N \geq 0$. Shearing of the solids in contact gives birth to shear stresses (τ_1, τ_2) within the interface. They are defined in each local direction such that

$$\mathbf{t} = -\mathrm{p}_N \, \mathbf{e}_1^1 + \tau_1 \, \mathbf{e}_2^1 + \tau_2 \, \mathbf{e}_3^1. \tag{4}$$

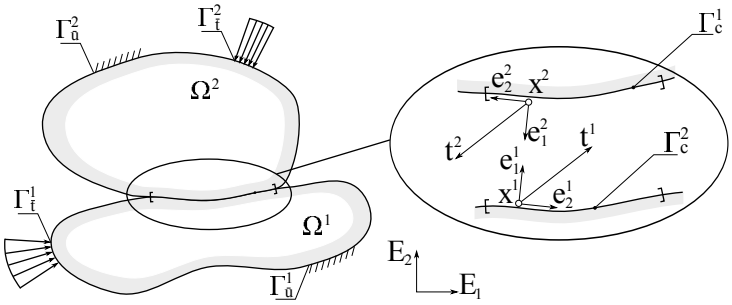


Figure 1: Statement of the mechanical problem, cross-section of the 3D problem in the $(\mathrm{E}_1, \mathrm{E}_2)$ plane: $\gamma_{\bar{u}}$, imposed displacements; $\gamma_{\bar{t}}$, imposed tractions.

The contact constraint defined by equation (3) is regularised by the penalty method. Therefore the contact is not ideal any more and interpenetration of both solids in contact is allowed, *i.e.* $\mathrm{g}_N < 0$. The evolution of normal pressure in case of contact reads

$$\dot{\mathrm{p}}_N = -\mathrm{K}_N \, \dot{\mathrm{g}}_N, \tag{5}$$

where K_N is a penalty coefficient and the minus sign ensures the contact pressure to be positive when interpenetration increases, *i.e.* $\dot{\mathrm{g}}_N < 0$.

In case of contact, both solids are either in ideal *stick* or *slip* state [13]. In the former, the relative displacement is equal to zero upon shearing. In the latter, the solids are allowed to move tangentially and the shear stress is limited to a maximum value. The transition between these states is ruled by the Mohr Criterion, such that

$$f(\mathbf{t}, \mu) = \underbrace{\sqrt{(\tau_1)^2 + (\tau_2)^2}}_{\|\tau\|} - \mu \, \mathrm{p}_N, \tag{6}$$

where μ is the friction coefficient. The stick state is also regularised by the penalty method. A small relative tangential displacement is allowed even in this case and the evolution of the shear stress is computed according to

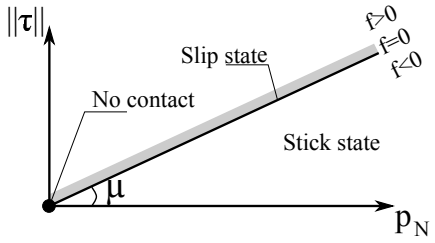
$$\dot{\tau}_i = \mathrm{K}_T \, \dot{\mathrm{g}}_{Ti}, \tag{7}$$

where K_T is a tangential penalty coefficient. Therefore the evolution of the stress state within the interface is similar to the framework of elastoplasticity. The Coulomb criterion is the yield surface described in Figure 2, the elastic state is equivalent to the stick state and plasticity to slip state. Penalty coefficients $(\mathrm{K}_N, \mathrm{K}_T)$ are similar to elastic parameters even if they are introduced as purely numerical tools.

The finites elements developed belong to the family of zero-thickness elements. These elements lie on the boundary of the solids and have no thickness. They discretise the normal constraint and the shear stress along the boundary [8, 13]. The contact pressure and the gap function are computed at each integration point of one of the two solids, according to the mortar method [9, 14, 15]. The normal contact constraint is verified in a weak sense over the element.

	No contact	Stick	Slip
p_N	$= 0$	> 0	> 0
$\|\tau\|$	$= 0$	≥ 0	$= \mu \cdot p_N$

(a) Stress state in the interface in each case.

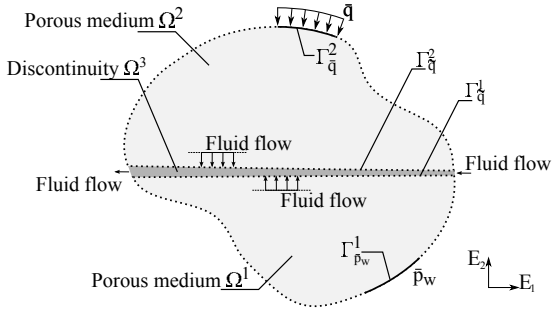


(b) Mohr-Coulomb criterion.

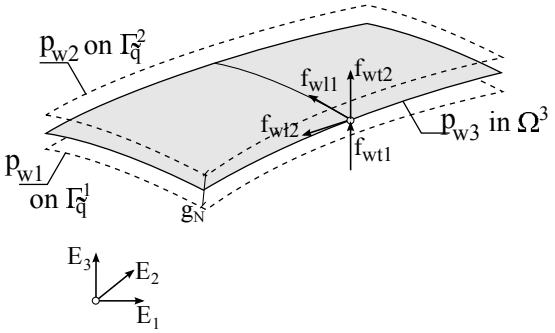
Figure 2: Analogy with elastoplasticity

2.2 Flow problem

Let us consider two solids that are porous media saturated with water. Therefore, depending on boundary conditions, fluid flow may take place within them as shown in Figure 3a. Solids Ω^1 and Ω^2 delineate a new volume Ω^3, which represents a discontinuity. Fluid flow may also exist along and through Ω^3.



(a) Definition of the flow problem (cross section of the 3D case in the (E_1, E_2) plane), porous medium, discontinuity and boundaries: $\Gamma_{\bar{q}}$, imposed fluw; $\Gamma_{\bar{p}_w}$, imposed pressure.



(b) Three-node discretisation: definition of longitudinal and transversal flows.

Figure 3

The interface is discretised according to a three-node scheme as shown in Figure 3b. The fields of water pressure are considered on each side of the interface $(\Gamma^1_{\bar{q}}, \Gamma^2_{\bar{q}})$ and inside it (Ω^3). Therefore two longitudinal and two transversal fluxes must be respectively defined in the normal and tangential local directions.
Darcy's law is assumed to represent fluid flow in the local tangential directions such that

$$\mathrm{f}_{wl(i-1)} = -\frac{k_l}{\mu_w}\left(\nabla_{\mathbf{e}_i^1} \mathrm{p}_{w3} + \rho_w \, \mathrm{g} \, \nabla_{\mathbf{e}_i^1} z\right) \rho_w \qquad \text{for} \quad i = 2, 3 \tag{8}$$

where $\nabla_{\mathbf{e}_i^1}$ is the gradient in the direction $\mathbf{e}_i^1$, z is the vertical global direction, μ_w is the dynamic viscosity of the fluid, g the acceleration of gravity, ρ_w is the density of the fluid

and k_l is the longitudinal permeability.
The transversal fluid fluxes $(\mathrm{f}_{wt1}, \mathrm{f}_{wt2})$ between each side of the interface and the inner medium are defined according to

$$\mathrm{f}_{wt1} = \rho_w \, \mathrm{T}_{w1} \, (\mathrm{p}_{w1} - \mathrm{p}_{w3}) \qquad \text{on} \quad \Gamma^1_{\tilde{q}}, \tag{9}$$

$$\mathrm{f}_{wt2} = \rho_w \, \mathrm{T}_{w2} \, (\mathrm{p}_{w3} - \mathrm{p}_{w2}) \qquad \text{on} \quad \Gamma^2_{\tilde{q}}, \tag{10}$$

where T_{w1} and T_{w2} are two transversal conductivities.

2.3 Couplings

The mechanical and flow problems are intrinsically coupled. The total normal pressure acting on each side of the interface is decomposed according to Terzaghi's principle

$$\mathrm{p}_N = \mathrm{p}'_N + \mathrm{p}_{w3}. \tag{11}$$

where p'_N is the effective pressure and p_{w3} is the water pressure inside the interface. The Coulomb criterion becomes a function of the effective pressure only.

The opening/closing of the gap g_N has two effects. Firstly it influences the permeability of the discontinuity according to the cubic law [16, 2]

$$k_l = \begin{cases} \dfrac{(\mathrm{D}_0)^2}{12} & \text{if} \qquad \mathrm{g}_N \leq 0 \\ \dfrac{(\mathrm{D}_0 + \mathrm{g}_N)^2}{12} & \text{otherwise,} \end{cases} \tag{12}$$

where D_0 is the residual hydraulic aperture. If contact holds, i.e. $\mathrm{g}_N \leq 0$, it ensures that the permeability is not null. Indeed, if the surfaces of the bodies in contact are not perfectly smooth, there is still a residual opening. Finally the total mass of fluid M_f enclosed within the interface is modified according to

$$\dot{\mathrm{M}}_f = \left(\dot{\rho}_w \, \mathrm{g}_N + \rho_w \, \dot{\mathrm{g}}_N + \rho_w \, \mathrm{g}_n \, \frac{\dot{\Gamma}_{\tilde{q}}}{\Gamma_{\tilde{q}}} \right) \Gamma_{\tilde{q}}, \tag{13}$$

where $\Gamma_{\tilde{q}}$ is the total surface along which the water flow takes place. It is assumed that the tangential displacement remains limited $\dot{\Gamma}_{\tilde{q}} \to 0$ and the fluid is incompressible $\dot{\rho}_w = 0$.

2.4 Finite element discretisation

The final discretisation of a 3D interface is represented in Figure 4. Both sides of the interface are described by classical quadrangular isoparametric finite elements. Their nodes have four degrees of freedom, three mechanical and one fluid. Inner nodes have a single fluid degree of freedom since they only describe fluid flow.

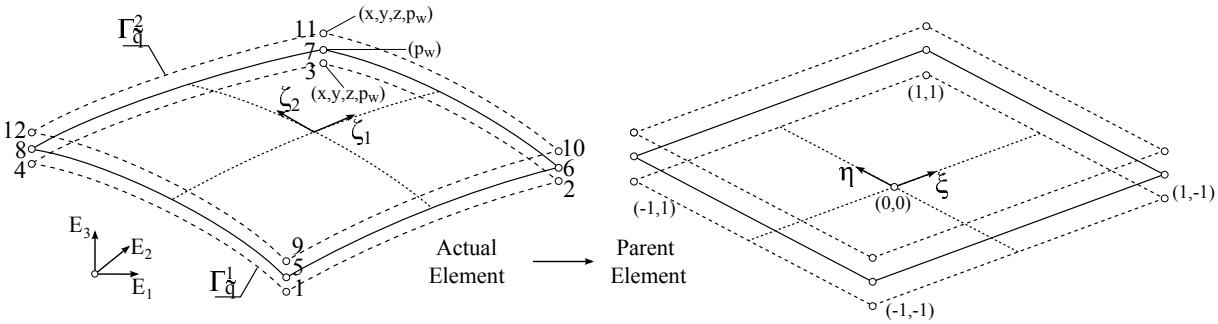
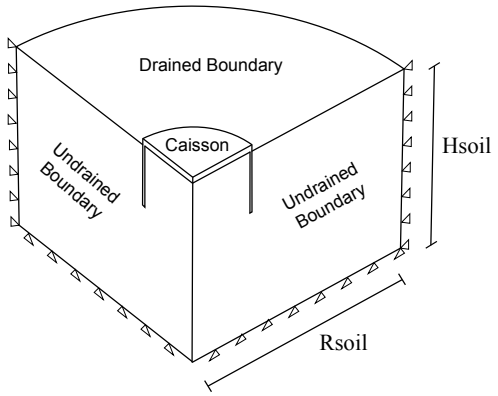


Figure 4: Discretisation of the interface into isoparametric elements. Transformation to the parent element.

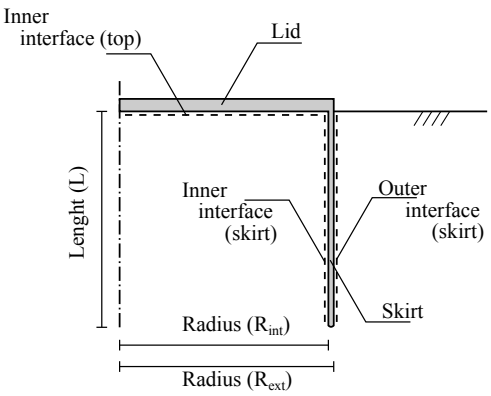
3 APPLICATION

3.1 Statement of the problem

The application consists of the pull loading of a suction caisson embedded in an elastic layer of soil. This simulation is inherently 2D but a 3D simulation is carried out in order to check the formulation of the finite element. The caisson has an outer radius of 3.9m, an inner radius of 3.8m and a skirt length of 4m as shown in Figure 5a. The thickness of its lid and skirt are respectively 0.4m and 0.1m. Interface elements are set up between the soil and the caisson as described in Figure 5b.



(a) Global sketch of the geometry of the problem.



(b) Description of the caisson.

Figure 5

The caisson is made of steel and its behaviour is elastic. The elastic soil is represented by a quarter of cylinder. Its radius is equal to 24m and its depth to 12m. The soil is assumed elastic. All material parameters are provided in Table 1. Interface elements are characterised by a friction coefficient of 0.57. The transversal conductivity is equal to zero between the inner interface and the caisson since it is impervious. This conductivity

is equal to 1.E-8 m/Pa.s between the inner interface and the soil.

Two types of uplift simulations are carried out in order to validate the formulation of the three dimensional finite element. The first simulation is *drained*, *i.e.* there is no variation of pore pressures within the soil. It occurs if the loading rate is sufficiently low and pore pressures have time to dissipate. This simulation highlights the mechanical behaviour of the interface, *i.e.* the progressive mobilisation of friction along the skirt and the soil-caisson sliding. The second simulation is partially drained, pore pressures are able to partially dissipate. It illustrates the coupled behaviour of the interface element.

Soil	E [MPa]	ν [-]	n [-]	k [m^2]	γ_s [kg/m^3]	K_0 [-]
	2E2	0.3	0.36	1.E-11	2650	1
Caisson	E [MPa]	ν [-]	n [-]	k [m^2]	γ_s [kg/m^3]	K_0 [-]
	2E5	0.3	0.36	0	2650	1
Interface	K_N [N/m^3]	K_T [N/m^3]	μ [-]	D_0 [m]	T_w [m/(Pa.s)]	
	1E10	1E10	0.57	1.E-5	1.E-8	

Table 1: Material parameters: E Young modulus, ν Poisson's ratio, n porosity, k permeability, γ_s density of solid grains, K_0 coefficient of earth pressure at rest, K_N, K_T penalty coefficients, μ friction coefficient, T_w transversal conductivity, D_0 residual hydraulic aperture.

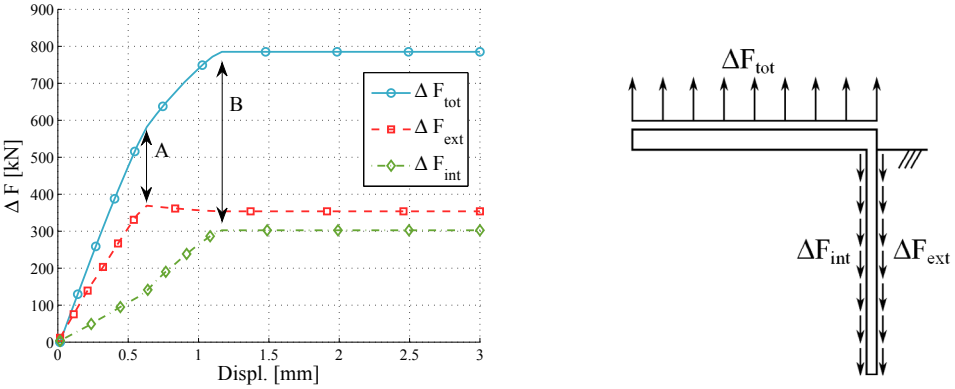
3.2 Drained simulation

During the displacement controlled simulation, the total uplifting load ΔF_{tot} is balanced by the dead weight and the shearing along the skirt (inside ΔF_{int} and outside ΔF_{ext}). Their evolution with respect to the vertical displacement is represented in Figure 6a. At the early beginning of the simulation, the evolution of ΔF_{int} and ΔF_{ext} is almost linear. Indeed, the Mohr criterion is not reached yet and the variation of the shear stress is a function of the variation of tangential displacement

$$\dot{\tau} = K_T \, \dot{g}_T. \tag{14}$$

The slopes of ΔF_{ext} and ΔF_{int} are different. Indeed, the soil plug inside the caisson moves upward with it, i.e. the caisson acts as a punch. Therefore the relative displacement is lower inside than outside.

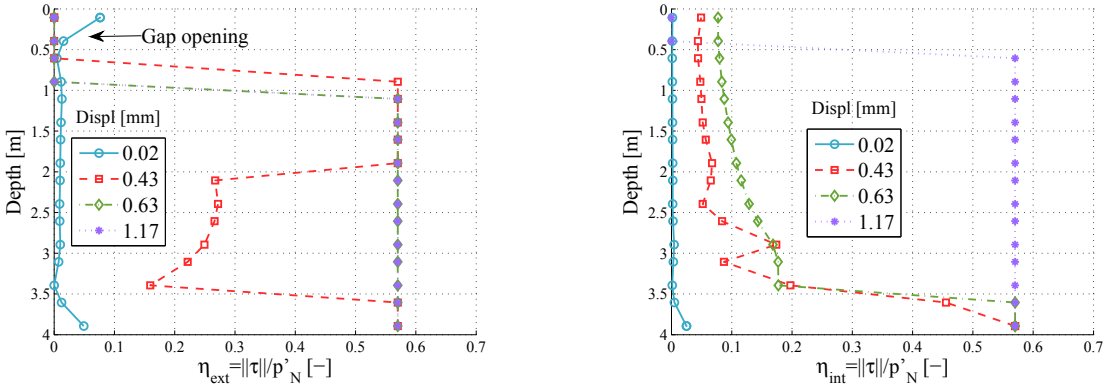
Friction is progressively mobilised outside the caisson up to point A in Figure 6a. From this point the soil and the caisson start sliding at constant shear stress, leading to a plateau in the evolution of ΔF_{ext}. The mobilised friction outside the skirt, $\eta_{ext} = (\tau/p'_N)_{ext}$, is depicted in Figure 7a. The friction coefficient is reached along the whole outer skirt after an uplift displacement of 0.63mm.
It is worth noting the opening of a gap, denoted by $\eta_{ext} = 0$, at the top of the skirt.



(a) (b)

Figure 6: Drained pull simulation of the suction caisson: ΔF_{tot} variation of total vertical load, ΔF_{ext} integral of shear mobilised outside the caisson, ΔF_{int} integral of shear mobilised inside the caisson.

Indeed, diffusion of shear stresses within the soil creates this loss of contact. This is mainly due to the elastic behaviour of the soil. Such a gap is strongly reduced if the soil has a non-linear elastoplastic behaviour [6, 17].
Similarly the outer friction is fully mobilised for a greater uplift displacement as shown in Figure 7b. The plateau is reached at point B in Figure 6a. The shear is more uniformly distributed and denotes the uplifting of the soil plug. The total applied load ΔF_{tot} reaches also a plateau since no additional friction can be mobilised. This plateau would not last for very large displacement. Indeed, the contact area decreases with increasing uplift.

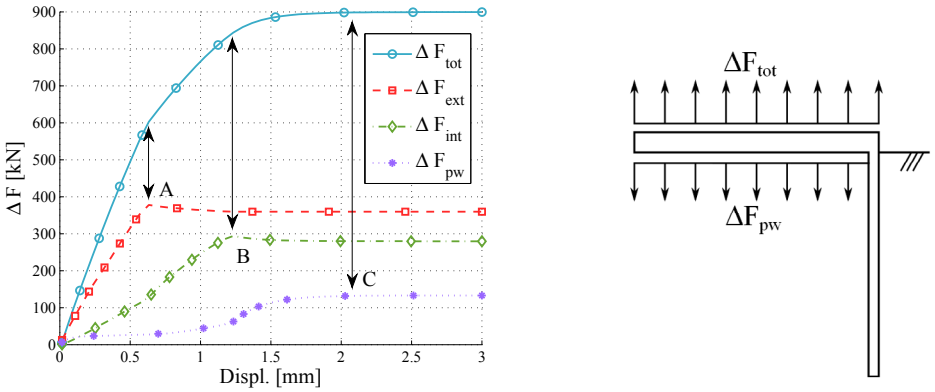


(a) Outside the caisson (η_{ext}). (b) Inside the caisson (η_{int}).

Figure 7: Drained pull simulation of the suction caisson, mobilised shear along the skirt.

3.3 Partially drained simulation

The partially drained uplifting of the caisson is illustrated in Figure 8a. In this case, the total load reached at the beginning of the plateau is greater than in the drained simulation. The suction effect involved in the installation of suction caissons is also mobilised during the uplift. The pull load creates an inverse consolidation process, where negative variations of pore pressures are generated inside the caisson. The differential of pressure between inside and outside holds the caisson. A new component of resistance is termed $\Delta \mathrm{F}_{uw}$. The distribution of variation of pore pressures, $\Delta \mathrm{p}_W$, is described in Figure 9a.



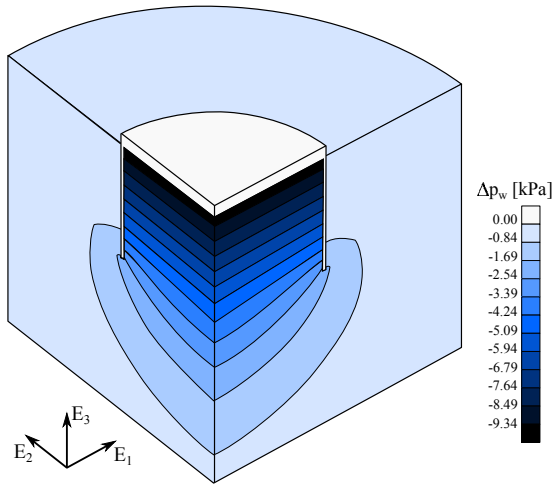
(a) (b)

Figure 8: Undrained pull simulation of the suction caisson: $\Delta \mathrm{F}_{tot}$ variation of total vertical force, $\Delta \mathrm{F}_{ext}$ integral of shear mobilised outside the caisson, $\Delta \mathrm{F}_{int}$ integral of shear mobilised inside the caisson, $\Delta \mathrm{F}_{pw}$ integral of the variation of water pressure at the top inside the caisson.

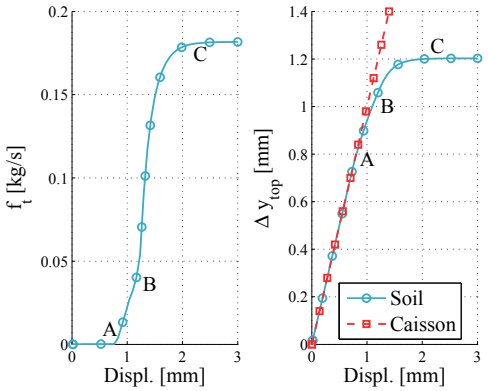
The frictional behaviour is identical to the drained simulation. Friction is fully mobilised inside and outside the caisson at points A and B in Figure 8a. The last component of resistance $\Delta \mathrm{F}_{pw}$ starts increasing strongly after that point and finally reaches a plateau. The lid of the caisson and the soil plug keep in contact up to point A, as shown in Figure 8a. After sliding takes place on the outer skirt, a gap is opened between the lid and the soil. This gap is filled by water. Therefore, a transversal flux of water holds inside the caisson. The integral of the water flux over the top surface is provided in Figure 9b. The displacement of the top soil increases upward but reaches a plateau. After a displacement of 2mm, a stationary phase takes place. The upward displacement, the transversal flux and the variation of pressure are constant. If the caisson is assumed rigid, the rate of water stored within the gap is computed according to

$$\dot{\mathrm{S}} = \rho_w \, \mathrm{v}_{up} \, \pi \, \mathrm{R}_{int}^2 / 4 = 1.89 \, 10^{-1} \, \mathrm{kg/s}, \tag{15}$$

where v_{up} = 1mm/min is the uplifting rate. The mass variation is equal to the one numerically computed as observed in Figure 9b.



(a) Variations of pore water pressure Δp_w within the soil around the caisson, at the end of the simulation.



(b) Total transversal flux between the soil and the interface inside the caisson (left); vertical displacement of the top soil inside caisson.

Figure 9

The gap opening between the skirt and the soil, observed during the drained simulation, also occurs during the partially drained one. A longitudinal flow of water takes place within this gap. This is illustrated in Figure 10 at the end of the simulation. The flow is dependent on the gap opening since the permeability depends on the cubic law. This flow reduces the efficiency of the suction caisson since it decreases the length of the drainage path. However if an elastoplastic constitutive model was used, this pipe would be strongly reduced due to plasticity effects.

4 CONCLUSIONS

The main features of a finite element of interface are presented in this work. The element is zero-thickness. The normal contact constraint is regularised by the penalty method and discretised by the mortar approach. The discretisation of the hydraulic problem is three-node, namely the fluid flow inside the interface is discretised by additional nodes. Two longitudinal and two transversal fluxes are defined in the plane and across the interface.

The crucial role of interfaces is demonstrated for the simulation of the uplifting of a suction caisson. A first drained simulation is carried out. The maximum load sustainable by the suction caisson is bounded by the maximum friction available along the skirt. The caisson starts sliding afterwards.

The partially drained behaviour illustrates the coupled role of the interface element. The

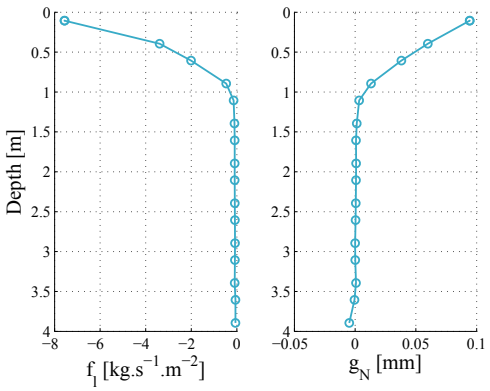


Figure 10: Relation between the longitudinal flux and the opening of a gap along the skirt, outside the caisson, end of the simulation.

inverse consolidation process taking place during the uplift generates negative variations of water pressures inside the caisson. This suction effect holds the caisson and transiently increases the total pulling load sustainable. The gap creating between the soil and the lid of the caisson is filled with water, illustrating the storage of of fluid within the interface. Longitudinal water flows also occur along the skirt due to the opening of a vertical gap. This reduces the effect of the suction caisson by reducing the length of the drainage path.

REFERENCES

[1] D. Sheng, P. Wriggers, and S.W. Sloan. Improved numerical algorithms for frictional contact in pile penetration analysis. *Computers and Geotechnics*, 33(6-7):341–354, September 2006.

[2] C. Guiducci, A. Pellegrino, J.-P. Radu, F. Collin, and R. Charlier. Numerical modeling of hydro-mechanical fracture behavior. In *Numerical models in Geomechanics*, 2002.

[3] J. Rutqvist, J.T. Birkholzer, and C.-F. Tsang. Coupled reservoir–geomechanical analysis of the potential for tensile and shear failure associated with co2 injection in multilayered reservoir–caprock systems. *International Journal of Rock Mechanics and Mining Sciences*, 45(2):132–143, 2008.

[4] G.T. Houlsby, L.B. Ibsen, and B.W. Byrne. Suction Caissons for Wind Turbines. *International Symposium on Frontiers in Offshore Geotechnics*, 75(September):94, 2005.

[5] K. Thieken, M. Achmus, and C. Schröder. On the behavior of suction buckets in sand under tensile loads. *Computers and Geotechnics*, 60:88–100, 2014.

[6] B. Cerfontaine, S. Levasseur, F. Collin, and R. Charlier. Axisymmetric transient modelling of a suction caisson in dense sand. In *Proceedings of the 8th European Conference on Numerical Methods in Geotechnical Engineering, NUMGE 2014*, volume 2, pages 1243–1248, 2014.

[7] B. Cerfontaine. *The cyclic behaviour of sand, from the Prevost model to offshore geotechnics*. PhD thesis, University of Liege, 2014.

[8] R.E. Goodman, R.L. Taylor, and T.L. Brekke. A model for the mechanics of jointed rock. *Journal of the Soil Mechanics and Foundations Division*, 94, 1968.

[9] A.M. Habraken, S. Cescotto, and Q. Banning. Contact between deformable solids: The fully coupled approach. *Mathematical and Computer modelling*, 28(4):153–169, 1998.

[10] B. Jha and R. Juanes. Coupled multiphase flow and poromechanics: A computational model of pore pressure effects on fault slip and earthquake triggering. *Water Resources Research*, 50:3376–3808, 2014.

[11] P. Gerard, R. Charlier, R. Chambon, and F. Collin. Influence of evaporation and seepage on the convergence of a ventilated cavity. *Water resources research*, 44(5), 2008.

[12] B. Cerfontaine, A.C. Dieudonné, J.P. Radu, F. Collin, and R. Charlier. 3d zero-thickness coupled interface finite element: formulation and application. *Computers and Geotechnics*, 2015 (submitted).

[13] P. Wriggers. *Computational contact mechanics*. Wiley : Chichester, second edition, 2006.

[14] M.A. Puso and T.A. Laursen. A mortar segment-to-segment contact method for large deformation solid mechanics. *Computer Methods in Applied Mechanics and Engineering*, 193(6):601–629, 2004.

[15] K.A. Fischer and P. Wriggers. Mortar based frictional contact formulation for higher order interpolations using the moving friction cone. *Computer methods in applied mechanics and engineering*, 195(37):5020–5036, 2006.

[16] R. Olsson and N. Barton. An improved model for hydromechanical coupling during shearing of rock joints. *International Journal of Rock Mechanics and Mining Sciences*, 38(3):317–329, 2001.

[17] B. Cerfontaine, F. Collin, and R. Charlier. Vertical transient loading of a suction caisson in dense sand. In *Proceedings of the 14th Int. Conference of International Association for Computer Methods and Recent Advances in Geomechanics, IACMAG 2014*, pages 929–934, 2015.

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis(Eds)

TOWARDS GOAL-ORIENTED MESH ADAPTATION FOR FLUID-STRUCTURE INTERACTION

E. GAUCI[a], F. ALAUZET[b], A. LOSEILLE[b], A.DERVIEUX[a]

[a] INRIA, Projet Ecuador, 2004 route des Lucioles - BP 93, 06902 Sophia Antipolis, France

[b] INRIA, Projet Gamma, Domaine de Voluceau, Rocquencourt, BP 105, 78153 Le Chesnay, France.

Key words: Unsteady compressible flow, Arbitrary Lagrangian Eulerian, goal-oriented mesh adaptation, anisotropic mesh adaptation, adjoint, metric.

Abstract. In order to address fluid-structure interaction, we present an *a priori* analysis for an ALE compressible flow model. This analysis is the key for an anisotropic metric-based mesh adaptation.

1 INTRODUCTION

This work takes place into a serie of works for metric-based goal-oriented mesh adaptation. Metrics are matrix fields representing a mesh thanks to the description of three (3D case) or two (2D) orthogonal vectors giving the directions of stretching of the mesh associated with the definition of the mesh size in each directions. In anisotropic mesh adaptation, the optimal metric has to be derived from the minimization of a model of the approximation error. Historically, the first model was a linear interpolation error model, leading to the so-called Hessian-based anisotropic adaptation. The optimal metric field, *i.e.* the optimal sizes and orientations distribution that will be used to govern the generation of the new mesh, is obtained as the solution of the problem of minimization of the global interpolation error. It thus depends on a computed solution, but the link with the original PDE is sometimes too weak. In contrast, goal-oriented mesh adaptation takes into account the PDE, via the introduction of an adjoint state. Extension of goal-oriented analysis to anisotropic mesh adaptation has been proposed in [15] and [11] for the steady Euler equations. An extension to transient Euler with the design of a global transient global fixed point (TFP) mesh adaptation algorithm was proposed in [7]. The TFP computes the n meshes to be used in the N_i time sub-intervals of the global time interval $]0, T[$ of the simulation. To generate the N_i sub-interval meshes, the complete state solution from time 0 to time T must first be computed as long as the complete adjoint solution from T to 0.

In order to be able, in the long term, to address fluid-structure interaction, the metric-based adaptation needs to be extended to Arbitrary-Lagrangian-Eulerian (ALE) formulations. In ALE, the mesh is moving (in fact deforming) during any time interval. An important and difficult issue is to force the mesh to stay optimally adapted while moving. In the case of a Hessian-based adaptation, the mesh needs to be adapted at any time to a time-dependent criterion, the Hessian of the unsteady sensor. But the mesh is also uniquely defined by the initial mesh and the prescribed ALE mapping. In [5, 4], the two constraints are simultaneously respected by (1) for any time mapping adaptation constraints on the initial mesh, and then (2) taking the metric-intersection of these constraints on the initial mesh. The resulting ALE-TFP adaptation has been successfully applied to several fluid-structure interaction computations. Although very efficient, this method is a Hessian-based one and inherits the Hessian-based deficiencies, which involves, as already pointed out, an insufficient accounting of the PDE itself. The purpose of the work reported here is to extend the goal-oriented formulation of [11][7] to ALE calculations. In other words, we try to build a metric-based, goal-oriented unsteady anisotropic mesh adaptation method for an ALE Euler model. In the present short paper, we concentrate on the ALE error analysis for a given time-step.

2 CONTINUOUS MESH MODEL

We propose to work in the continuous mesh framework, introduced in [9, 10]. The main idea of this framework is to model continuously discrete meshes by Riemannian metric spaces. This enables to define proper differentiable optimization [2, 6], *i.e.,* to use calculus of variations on continuous meshes. Indeed, trying to solve optimality problems manipulating discrete meshes generally leads to intractable problems in practice. This framework lies in the class of metric-based methods. A continuous mesh $\mathcal{M}$ of computational domain Ω is identified to a Riemannian metric field [8] $\mathcal{M} = (\mathcal{M}(\mathbf{x}))_{\mathbf{x}\in\Omega}$. For all $\mathbf{x}$ of Ω, $\mathcal{M}(\mathbf{x})$ is a symmetric 3×3 matrix having $(\lambda_i(\mathbf{x}))_{i=1,3}$ as positive eigenvalues along the principal directions $\mathcal{R}(\mathbf{x}) = (\mathbf{v}_i(\mathbf{x}))_{i=1,3}$. Sizes along these directions are denoted $(h_i(\mathbf{x}))_{i=1,3} = (1/\sqrt{\lambda_i}(\mathbf{x}))_{\mathbf{i=1,3}}$ and the three *anisotropy quotients* r_i are defined as: $r_i = h_i^3\,(h_1h_2h_3)^{-1}$. The diagonalization of $\mathcal{M}(\mathbf{x})$ writes:

$$\mathcal{M}(\mathbf{x}) = \sqrt[3]{\det(\mathcal{M})}(\mathbf{x})\,\mathcal{R}(\mathbf{x})\mathrm{Diag}(r_1^{-\frac{2}{3}}(\mathbf{x}), r_2^{-\frac{2}{3}}(\mathbf{x}), r_3^{-\frac{2}{3}}(\mathbf{x}))^t\mathcal{R}(\mathbf{x}), \tag{1}$$

The *complexity* $\mathcal{C}$ of a continuous mesh is the continuous counterpart of the total number of vertices:

$$\mathcal{C}(\mathcal{M}) = \int_\Omega \sqrt{\det(\mathcal{M}(\mathbf{x}))}\,\mathrm{d}\mathbf{x}.$$

Given a continuous mesh $\mathcal{M}$, we shall say, following [9, 10], that a discrete mesh $\mathcal{H}$ of the same domain Ω is a unit mesh with respect to $\mathcal{M}$, if each tetrahedron $K\in\mathcal{H}$, defined by its list of edges $(\mathbf{e}_i)_{i=1\ldots6}$, verifies:

$$\forall i\in[1,6],\quad \ell_{\mathcal{M}}(\mathbf{e}_i)\in\left[\frac{1}{\sqrt{2}},\sqrt{2}\right],\ \text{in which}\ \ell_{\mathcal{M}}(\mathbf{ab}) = \int_0^1 \sqrt{{}^t\mathbf{ab}\,\mathcal{M}(\mathbf{a}+t\,\mathbf{ab})\,\mathbf{ab}}\,\mathrm{d}t.$$

Given a smooth function u, each unit mesh $\mathcal{H}$ with respect to $\mathcal{M}$ is associated with a local interpolation error $|u - \Pi_{\mathcal{H}} u|$. In [9, 10], it is shown that all these interpolation errors are well represented by the so-called continuous interpolation error related to $\mathcal{M}$, which is expressed locally in terms of the Hessian H_u of u as follows:

$$(u - \pi_{\mathcal{M}} u)(\mathbf{x}, t) \quad = \quad \frac{1}{10} \text{trace}(\mathcal{M}^{-\frac{1}{2}}(\mathbf{x}) \, |H_u(\mathbf{x}, t)| \, \mathcal{M}^{-\frac{1}{2}}(\mathbf{x}))$$

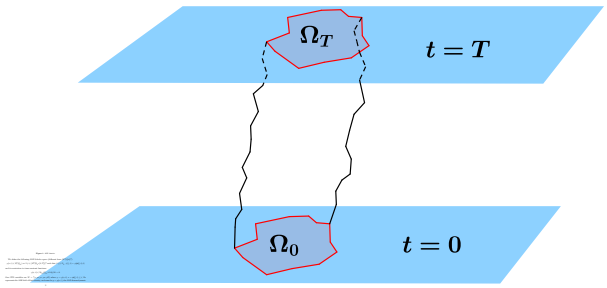
where $|H_u|$ is deduced from the Hessian H_u of u by taking the absolute values of its eigenvalues and where time-dependency notations have been added for use in the next sections.

3 ALE EULER MODEL

ALE domain and functions. Let Ω_0 be a smooth domain of $\mathbb{R}^d$. For any t in $]0, T[$ we have a mapping: $\boldsymbol{\phi} : \Omega_0 \times]0, T[\to \mathbb{R}^d$ satisfying: $\boldsymbol{\phi}(x, 0) = x$ and for which we assume that there exists a smooth deformation velocity $\dot{\boldsymbol{\phi}} = \partial \boldsymbol{\phi} / \partial t$. Let $\Omega_t = \boldsymbol{\phi}(\Omega_0, t)$.

We call Q_T the non-cylindrical union of all Ω_t for t in $]0, T[$:

$$Q_T = \bigcup_{t=0}^{t=T} \boldsymbol{\phi}(\Omega_0, t).$$



pressure, and $E = E(x,t)$ the ALE energy per unit mass. u, v, and w are the ALE Cartesian components of the ALE velocity vector $\mathbf{u}$.

The ALE-Euler fluxes write :

$$\mathcal{F}(W) = {}^t(\rho,\ \rho(\mathbf{u}-\dot{\boldsymbol{\phi}})u + p\mathbf{e}_x,\ \rho(\mathbf{u}-\dot{\boldsymbol{\phi}})v + p\mathbf{e}_y,\ \rho(\mathbf{u}-\dot{\boldsymbol{\phi}})w + p\mathbf{e}_z,\ \rho(\mathbf{u}-\dot{\boldsymbol{\phi}})E + \mathbf{u}p).$$

Let us define the variational formulation of the ALE-Euler system:

$$\text{Find } W \in \mathcal{H}^1(Q_T) \text{ such that } \forall \varphi \in \mathcal{H}^1_{cst}(Q_T), \quad (\Psi(W)\,,\,\varphi) = 0$$

$$\begin{aligned}\text{with} \quad (\Psi(W)\,,\,\varphi) \;&=\; \int_{\Omega_0} \varphi(0)(W_0 - W(0))\,\mathrm{d}\Omega \;+\; \int_0^T \frac{\partial}{\partial t}\int_{\Omega_t} \varphi W\,\mathrm{d}\Omega\,\mathrm{d}t \\ &+\; \int_0^T\int_{\Omega_t} \varphi\,\nabla\cdot\mathcal{F}(W)\,\mathrm{d}\Omega\,\mathrm{d}t \;-\; \int_0^T\int_{\partial\Omega_t} \varphi\,\hat{\mathcal{F}}(W).\mathbf{n}\,\mathrm{d}\Gamma\,\mathrm{d}t\,, \quad (2)\end{aligned}$$

where $\hat{\mathcal{F}}$ takes into account the boundary conditions. Note that the ∇ is a differentiation for constant time, identifying x-derivative and ξ-derivatives.

Time-discretized formulation. The ALE-Euler model will be advanced in time by an explicit scheme. For simplicity, we write it as a Forward-Euler time advancing:

$$\begin{aligned}&\text{Find } W \in \mathcal{H}^1(Q_T) \text{ such that } \forall \varphi \in \mathcal{H}^1_{cst}(Q_T), \quad (\Psi(W)\,,\,\varphi) = 0 \ \text{ with} \\ &(\Psi(W)\,,\,\varphi) = \int_{\Omega_0} \varphi(0)(W_0 - W^1)\,\mathrm{d}\Omega \\ &+ \sum_{n=1}^{nmax} \frac{1}{t^{n+1}-t^n}\left[\int_{\Omega_{t^{n+1}}} \varphi(.,t^{n+1})W(.,t^{n+1})\,\mathrm{d}\Omega - \int_{\Omega_{t^n}} \varphi(.,t^n)W(.,t^n)\,\mathrm{d}\Omega\right] \\ &+ \sum_{n=1}^{nmax}\int_{\Omega_{tn}} \varphi(.,t^n)\,\nabla\cdot\mathcal{F}(W(.,t^n))\,\mathrm{d}\Omega - \sum_{n=1}^{nmax}\int_{\partial\Omega_{tn}} \varphi(.,t^n)\,\hat{\mathcal{F}}(W(.,t^n)).\mathbf{n}\,\mathrm{d}\Gamma. \quad (3)\end{aligned}$$

For the sake of simplicity again, we shall not address the time discretization error in this paper. Some justification of this option for explicit time advancing can be found in [3].

Fully discrete formulation. Let us consider τ_h a finite-element triangulation (2D) or tetrahedrization (3D) of Ω. We assume that $\Omega_h^n = \Omega^n$. We consider a dual finite-volume tessellation made of median-limited cells:

$$\Omega_h^n = \bigcup_{elements} T_h = \bigcup_{ic,cells} cell_h(ic)$$

We define the following approximation space derived from the standard P_1 FEM approximation space :

$$\begin{aligned}\mathcal{V}_h \;=\; &\{\varphi \in \mathcal{H}^1(Q_T),\ \ \varphi \in \mathcal{C}^0(\bar{Q}_t),\\ &\forall\ t \in]0,T[,\ \forall \text{ element } T \in \tau_h, \varphi(.,t)|_T \text{ is affine}\\ &\forall\ n,\ 1 \le nmax,\ \forall\ i \text{ vertex} \in \tau_h,\ \varphi(\mathbf{x}(i),.) \text{ is affine in time}\}.\end{aligned}$$

We define $\mathcal{V}_{h,cst} = \mathcal{H}^1_{cst}(Q_T) \cap \mathcal{V}_h$ and the interpolation operator:

$$\Pi_h : \mathcal{H}^1(Q_T) \to \mathcal{V}_h \ \ ; \ \ \Pi_h\varphi(\mathbf{x}(i), t^n) = \varphi(\mathbf{x}(i), t^n) \ \forall \ i \text{ vertex}, \ \forall \ n, \ 1 \leq nmax.$$

The fully discrete state system which we consider (similarly to [11]) writes:

$$\text{Find } W_h \in \mathcal{V}_h \text{ such that } \forall \varphi_h \in \mathcal{V}_{h,cst}, \quad (\Psi_h(W_h) \, , \, \varphi_h) = 0 \tag{4}$$

with $\quad \forall W \in \mathcal{H}^1(Q_T), \forall \varphi \in \mathcal{H}^1(Q_T) \, ,$

$$(\Psi_h(W) \, , \, \varphi) = \int_{\Omega_0} \varphi(., t^1)(\Pi_h W_0 - \Pi_h W(., t^1)) \, \mathrm{d}\Omega$$

$$+ \sum_{n=1}^{nmax} \frac{1}{t^{n+1} - t^n} \left[\int_{\Omega_{t^{n+1}}} \Pi_h\varphi(., t^{n+1}) \Pi_h W(., t^{n+1}) \, \mathrm{d}\Omega \ - \ \int_{\Omega_{t^n}} \Pi_h\varphi(., t^n) \Pi_h W(., t^n)) \, \mathrm{d}\Omega \right]$$

$$+ \sum_{n=1}^{nmax} \int_{\Omega_{t^n}} \Pi_h\varphi(., t^n)) \, \nabla \cdot \Pi_h \mathcal{F}(W(., t^n))) \, \mathrm{d}\Omega - \sum_{n=1}^{nmax} \int_{\partial\Omega_{t^n}} \Pi_h\varphi(., t^n)) \, \Pi_h \hat{\mathcal{F}}(W(., t^n)).\mathbf{n} \, \mathrm{d}\Gamma$$

The Π_h in the initial condition and time derivative are useless for the discrete equation statement, but essential for the extension of the discrete residual to continuous functions of $\mathcal{H}^1(Q_T)$. Note that if in these initial and time terms the interpolation operator is replaced by a projection P_0 onto functions which are constant by cells, we get:

$$\int_{\Omega_{t^{n+1}}} P_0\varphi P_0 W^{n+1} \, \mathrm{d}\Omega \, \mathrm{d}t - \int_{\Omega_{t^n}} P_0\varphi P_0 W^n \, \mathrm{d}\Omega \, \mathrm{d}t \ = \ \sum_{cells,ic} \varphi(ic) \left(|meas(ic)^{n+1}| W^{n+1}(ic) - |meas(ic)^n| W^n(ic) \right)$$

in which we recognize the usual ALE finite-volume time-derivative.

4 MESH ADAPTATION HESSIAN-BASED CRITERION

4.1 Instantaneous mesh adaptation criterion

Let $s(W)$ a sensor function computed from the CFD field W at time t. It can be for example the corresponding Mach number. Starting from:

$$\|s(W) - \pi_{\mathcal{M}} s(W)\|_{\mathbf{L}^p(\Omega_h)} = \left(\int_\Omega \left(\mathrm{trace}\big(\mathcal{M}^{-\frac{1}{2}}(\mathbf{x}) |H_{s(W)}(\mathbf{x})| \mathcal{M}^{-\frac{1}{2}}(\mathbf{x}) \big) \right)^p \mathrm{d}\mathbf{x} \right)^{\frac{1}{p}} \tag{5}$$

we define as optimal metric the one which minimizes the right hand side under the constraint of a total number of vertices equal to a parameter N. After solving analytically this optimization problem, we get the unique optimal $(\mathcal{M}_{\mathbf{L}^p}(\mathbf{x}))_{\mathbf{x}\in\Omega}$ as:

$$\mathcal{M}_{\mathbf{L}^p} = D_{\mathbf{L}^p} \left(\det \left| H_{s(W)} \right| \right)^{\frac{-1}{2p+2}} |H_{s(W)} u| \ \text{ and } \ D_{\mathbf{L}^p} = N \left(\int_\Omega (\det \left| H_{s(W)} \right|)^{\frac{p}{2p+2}} \right)^{-1}, \tag{6}$$

where $D_{\mathbf{L}^p}$ is a global normalization term set to obtain a continuous mesh with complexity N and $(\det|H_u|)^{\frac{-1}{2p+2}}$ is a local normalization term accounting for the sensitivity of the $\mathbf{L}^p$ norm. In the case of an adaptation loop for solving a Partial Differential Equation, a continuous sensor function $s(W)$ is not available, but an approximate solution $s(W_{\mathcal{M}})$. In that case, the continuous interpolation error (2) is replaced by:

$$|s(W_{\mathcal{M}}) - \pi_{\mathcal{M}} s(W_{\mathcal{M}})|(\mathbf{x}) \;=\; \frac{1}{10}\text{trace}(\mathcal{M}^{-\frac{1}{2}}(\mathbf{x})\,|H_{s(W_{\mathcal{M}})}(\mathbf{x})|\,\mathcal{M}^{-\frac{1}{2}}(\mathbf{x})) \tag{7}$$

where $H_{s(W_{\mathcal{M}})}$ is an approximate Hessian of the discrete sensor, which is evaluated by the patch-recovery approximation defined in [14]. According to the continuous mesh framework, statement (6) defines directly a continuous optimal metric. In practice, solving (6) is done by approximation, *i.e.* in a discrete context with a couple (mesh, solution) denoted $(\mathcal{H}_{\mathcal{M}}, W_{\mathcal{M}})$ and iteratively through the following fixed point:

Step 1: compute the discrete state $W_{\mathcal{M}}$ on mesh $\mathcal{H}_{\mathcal{M}}$,
Step 2: compute sensor $s_{\mathcal{M}} = s(u_{\mathcal{M}})$ and optimal metric $\mathcal{M}^{opt}_{inter} = \mathcal{K}_p(H_{\mathcal{M}}(s_{\mathcal{M}}))$
Step 3: $\mathcal{M} = \mathcal{M}^{opt}_{inter}$, $\mathcal{H}_{\mathcal{M}} = \mathcal{H}_{\mathcal{M}^{opt}_{inter}}$ and go to step 1, until convergence.

4.2 A numerical example

The above method gives an optimal metric for each time of $[0, T]$. In practice, the global interval $[0, T]$ is divided into sub-intervals, $[0, T] = \cup[t_p, t_{p+1}]$ on which meshes will keep a constant topology. At each time level t^n inside $[t_p, t_{p+1}]$, we have (6) an optimal metric which is also a specification of the maximal mesh size in each direction for controlling the error to a prescribed level. This specification for t^n is therefore a constraint which we map onto the first mesh at time t_p of the subinterval. To take into account the different constraints for the different t^n of $[t_p, t_{p+1}]$, an intersection of these metrics is performed, defining the metric and the mesh at t_p, insuring that deformed mesh in $[t_p, t_{p+1}]$ satisfies the adaptation constraint. An example of computation deals with the 2D fluid-structure interaction due to the impact of a blast wave on a mobile rigid rectangular mass, cf. Figure 2.

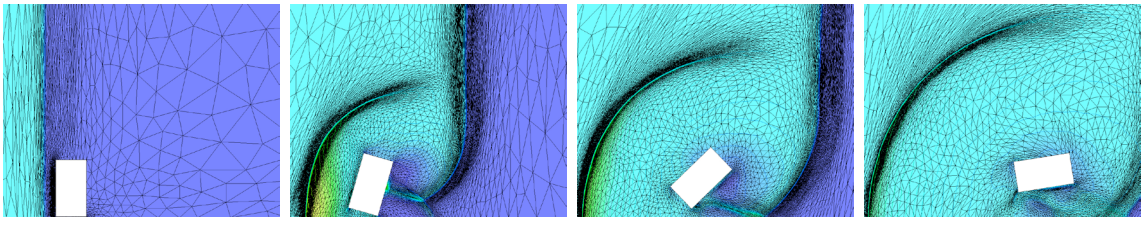


Figure 2: Application of the Fixed-Point algorithm to a blast wave inpinging an obstacle.

5 MESH ADAPTATION GOAL-ORIENTED CRITERION

According to the goal-oriented paradigm, we introduce two scalar outputs depending of the state variables (W and W_h are respectively the solutions of (3) and (5)):

$$j(W) = (g, W)_{L^2(Q_T)} \quad ; \quad \delta j = j(W) - j(W_h)$$

and we are interested by minimizing the approximation error committed on the evaluation δj of j. Following [11], [7], we formally develop δj with the help of the continuous and discrete adjoint states:

$$W^* \in \mathcal{V}, \forall \psi \in \mathcal{V}, \quad \left(\frac{\partial \Psi}{\partial W}(W)\psi, W^* \right) = (g, \psi), \tag{8}$$

$$W_h^* \in \mathcal{V}_h, \forall \psi_h \in \mathcal{V}_h, \quad \left(\frac{\partial \Psi_h}{\partial W}(W_h)\psi_h, W_h^* \right) = (g, \psi_h). \tag{9}$$

The idea is now to compute the difference of variational residual for a discrete test function:

$$(\Psi_h(W), \varphi_h) - (\Psi_h(W_h), \varphi_h) = (\Psi_h(W) - \Psi(W), \varphi_h)$$

Then assuming that W^*, $\Pi_h W^*$ and W_h^* and their gradients are close to each other:

$$\delta j \approx (\Psi_h(W) - \Psi(W), W^*) \approx (\Psi_h(W) - \Psi(W), \Pi_h W^*)\,. \tag{10}$$

The term $\Psi_h(W) - \Psi(W)$ is an *a posteriori* local error which we now evaluate.

5.1 Local error analysis

We replace in Estimation (10) operators Ψ and Ψ_h by their expressions given by Relations (3) and (5). We follow again this option. We also discard the error committed when imposing the initial condition. We finally get the following simplified error model:

$$\begin{aligned}
\delta j \approx & \sum_{n=1}^{n=nmax} \frac{1}{t^{n+1} - t^n} \int_{\Omega_{t^{n+1}}} \Pi_h W^{*,n+1} (\Pi_h W^{n+1} - W^{n+1})\, \mathrm{d}\Omega \\
& - \sum_{n=1}^{n=nmax} \frac{1}{t^{n+1} - t^n} \int_{\Omega_{t^n}} \Pi_h W^{*,n} (\Pi_h W^n) - W^n)\, \mathrm{d}\Omega \\
& + \sum_{n=1}^{n=nmax} \int_{\Omega_{t^n}} \Pi_h W^{*,n} \nabla \cdot (\Pi_h \mathcal{F}(W^n) - \mathcal{F}(W^n))\, \mathrm{d}\Omega \\
& - \sum_{n=1}^{n=nmax} \int_{\partial\Omega_{t^n}} \Pi_h W^{*,n} \, (\Pi_h \hat{\mathcal{F}}(W^n) - \hat{\mathcal{F}}(W^n)).\mathbf{n}\, \mathrm{d}\Gamma.
\end{aligned} \tag{11}$$

Integrating by parts leads to:

$$\begin{aligned}
\delta j \;\approx\; & \sum_{n=1}^{n=nmax} \frac{1}{t^{n+1}-t^n} \int_{\Omega_{t^{n+1}}} \Pi_h W^{*,n+1}(\Pi_h W^{n+1} - W^{n+1})\,\mathrm{d}\Omega \\
- & \sum_{n=1}^{n=nmax} \frac{1}{t^{n+1}-t^n} \int_{\Omega_{t^n}} \Pi_h W^{*,n}(\Pi_h W^n) - W^n)\,\mathrm{d}\Omega \\
- & \sum_{n=1}^{n=nmax} \int_{\Omega_{t^n}} \nabla \Pi_h W^{*,n} \cdot (\Pi_h \mathcal{F}(W^n) - \mathcal{F}(W^n))\,\mathrm{d}\Omega \\
- & \sum_{n=1}^{n=nmax} \int_{\partial\Omega_{t^n}} \Pi_h W^{*,n}\,(\Pi_h \bar{\mathcal{F}}(W^n) - \bar{\mathcal{F}}(W^n)).\mathbf{n}\,\mathrm{d}\Gamma. \qquad (12)
\end{aligned}$$

with $\bar{\mathcal{F}} = \hat{\mathcal{F}} - \mathcal{F}$. We observe that this estimate of δj is expressed in terms of interpolation errors of the Euler fluxes and of the time derivative weighted by continuous functions $\Pi_h W^* \approx W^*$ and $\nabla \Pi_h W^* \approx \nabla W^*$. The integrands in Error Estimation (12) contain positive and negative parts which can compensate for some particular meshes. In our strategy, we prefer not to rely on these parasitic effects and to slightly over-estimate the error. To this end, all integrands are bounded by their absolute values:

$$\begin{aligned}
\delta j \;\leq\; & \sum_{n=1}^{n=nmax} \frac{1}{t^{n+1}-t^n} \int_{\Omega_{t^{n+1}}} |\Pi_h W^{*,n+1}||\Pi_h W^{n+1} - W^{n+1}|\,\mathrm{d}\Omega \\
+ & \sum_{n=1}^{n=nmax} \frac{1}{t^{n+1}-t^n} \int_{\Omega_{t^n}} |\Pi_h W^{*,n}||\Pi_h W^n) - W^n|\,\mathrm{d}\Omega \\
+ & \sum_{n=1}^{n=nmax} \int_{\Omega_{t^n}} |\nabla \Pi_h W^{*,n}|\;|\Pi_h \mathcal{F}(W^n) - \mathcal{F}(W^n)|\,\mathrm{d}\Omega \\
+ & \sum_{n=1}^{n=nmax} \int_{\partial\Omega_{t^n}} |\Pi_h W^{*,n}|\;|(\Pi_h \bar{\mathcal{F}}(W^n) - \bar{\mathcal{F}}(W^n)).\mathbf{n}|\,\mathrm{d}\Gamma. \qquad (13)
\end{aligned}$$

5.2 Continuous error model

Working in this framework enables to write Estimate (13) in a *spatially*-continuous form, in which the Π_h are discarded, and in which the interpolation error $Id - \Pi_h$ is replaced by its continuous $Id - \pi_{\mathcal{M}}$. Then, we are interested in minimizing the following

error functional:

$$\begin{aligned}
\mathbf{E}(\mathcal{M}) &= \sum_{n=1}^{n=nmax} \frac{1}{t^{n+1}-t^n} \int_{\Omega_{t^{n+1}}} |W^{*,n+1}||\pi_{\mathcal{M}} W^{n+1} - W^{n+1}| \, \mathrm{d}\Omega \\
&+ \sum_{n=1}^{n=nmax} \frac{1}{t^{n+1}-t^n} \int_{\Omega_{t^n}} |W^{*,n}||\pi_{\mathcal{M}} W^{n} - W^{n}| \, \mathrm{d}\Omega \\
&+ \sum_{n=1}^{n=nmax} \int_{\Omega_{t^n}} |\nabla W^{*,n}| \; |\pi_{\mathcal{M}} \mathcal{F}(W^n) - \mathcal{F}(W^n)| \, \mathrm{d}\Omega \\
&+ \sum_{n=1}^{n=nmax} \int_{\partial\Omega_{t^n}} |W^{*,n}| \, |(\pi_{\mathcal{M}} \bar{\mathcal{F}}(W^n) - \bar{\mathcal{F}}(W^n)).\mathbf{n}| \, \mathrm{d}\Gamma. \qquad (14)
\end{aligned}$$

We observe that the fourth term introduces a dependency of the error with respect to the boundary surface mesh. In the present paper, we discard this term and refer to [12] for a discussion of its influence. The first term can be transformed as follows without introducing a large error:

$$\int_{\Omega_{t^{n+1}}} |W^{*,n+1}||\pi_{\mathcal{M}} W^{n+1} - W^{n+1}| \, \mathrm{d}\Omega = \int_{\Omega_{t^n}} |J_n^{n+1}|^{-1} |\tilde{W}^{*,n+1}||\pi_{\mathcal{M}} \tilde{W}^{n+1} - \tilde{W}^{n+1}| \, \mathrm{d}\Omega$$

where $|J_n^{n+1}|$ is the determinant of the transformation from Ω_n to Ω_{n+1}, and $\tilde{W}^{*,n+1}$ resp. $\tilde{W}^{n+1}$ the functions of Ω_n obtained by reverse transportation from Ω_{n+1}. Then, introducing the continuous interpolation error, we can write the simplified error model as follows:

$$\mathbf{E}(\mathbf{M}) = \sum_{n=1}^{n=nmax} \int_{\Omega} \mathrm{trace}\left(\mathcal{M}^{-\frac{1}{2}}(\mathbf{x}, t^n) \, \mathbf{H}(\mathbf{x}, t^n) \, \mathcal{M}^{-\frac{1}{2}}(\mathbf{x}, t^n) \right) \mathrm{d}\Omega \, \mathrm{d}t$$

with $\mathbf{H}(\mathbf{x}, t^n) = \sum_{j=1}^{5} \mathbf{H}_j(\mathbf{x}, t^n)$, in which

$$\begin{aligned}
\mathbf{H}_j(\mathbf{x}, t^n) &= \frac{1}{t^{n+1}-t^n} \left| |J_n^{n+1}|^{-1} \tilde{W}_j^{*,n+1}(\mathbf{x}, t) \right| \cdot \left| H(\tilde{W}_j^{n+1})(\mathbf{x}) \right| \\
&+ \frac{1}{t^{n+1}-t^n} \left| W_j^{*,n}(\mathbf{x}) \right| \cdot \left| H(W_j^n)(\mathbf{x}, t) \right| \\
&+ \left| \frac{\partial W_j^{*,n}}{\partial x}(\mathbf{x}) \right| \cdot \left| H(\mathcal{F}_1(W_j^n))(\mathbf{x}) \right| + \left| \frac{\partial W_j^{*,n}}{\partial y}(\mathbf{x}) \right| \cdot \left| H(\mathcal{F}_2(W_j^n))(\mathbf{x}) \right| \\
&+ \left| \frac{\partial W_j^{*,n}}{\partial z}(\mathbf{x}) \right| \cdot \left| H(\mathcal{F}_3(W_j^n))(\mathbf{x}) \right| \qquad (15)
\end{aligned}$$

is defined on Ω_n. Here, W_j^* denotes the j^{th} component of the adjoint vector W^*, $H(\mathcal{F}_i(W_j))$ the Hessian of the j^{th} component of the vector $\mathcal{F}_i(W)$, and $H(W_{j,t})$ the Hessian of the j^{th}

component of the time derivative of W. It should be noted that the time derivative can be estimated with more accuracy in order to avoid a large weight $(t^{n+1} - t^n)^{-1}$ which is not compensated due to the rough triangular inequality applied in the above majorations.

5.3 Spatial minimization for a fixed t

Let us assume that at time $t = t^n$, we seek for the optimal continuous mesh $\mathbf{M}_{go}(t)$ which minimizes the instantaneous error, *i.e.,* the spatial error for a fixed time t:

$$\tilde{\mathbf{E}}(\mathbf{M}(t)) = \int_\Omega \text{trace}\left(\mathcal{M}^{-\frac{1}{2}}(\mathbf{x},t)\,\mathbf{H}(\mathbf{x},t)\,\mathcal{M}^{-\frac{1}{2}}(\mathbf{x},t)\right)\,\mathrm{d}\mathbf{x}$$

under the constraint that the number of vertices is prescribed to $\mathcal{C}(\mathcal{M}(t)) = N(t)$. Similarly to [12], solving the optimality conditions provides the *optimal goal-oriented ("go") instantaneous continuous mesh* $\mathcal{M}_{go}(t) = (\mathcal{M}_{go}(\mathbf{x},t))_{\mathbf{x}\in\Omega}$ at time t defined by:

$$\mathcal{M}_{go}(\mathbf{x},t) = N(t)^{\frac{2}{3}}\,\mathcal{M}_{go,1}(\mathbf{x},t)\,, \tag{16}$$

where $\mathcal{M}_{go,1}$ is the optimum for $\mathcal{C}(\mathbf{M}(t)) = 1$:

$$\mathcal{M}_{go,1}(\mathbf{x},t) = \left(\int_\Omega (\det \mathbf{H}(\bar{\mathbf{x}},t))^{\frac{1}{5}}\mathrm{d}\bar{\mathbf{x}}\right)^{-\frac{2}{3}}\,(\det \mathbf{H}(\mathbf{x},t))^{-\frac{1}{5}}\;\mathbf{H}(\mathbf{x},t). \tag{17}$$

6 CONCLUDING REMARKS

The more complex a model is, the more necessary is the use of a mathematical method in order to control the approximation error. In fluid-structure interaction, a central challenge is the control of the error due to ALE. Several difficulties are combined: (a) the unsteadiness which implies to use several mesh topologies: we have chosen to build the method inside the Transient Fixed point algorithm which freezes the topology during time sub-intervals, (b) mesh motion, for which we use an existing elasticity model, (c) ALE error analysis. This paper concentrates on the (c) issue and proposes a formulation transforming the mesh adaptation problem into a metric-optimization problem. A proper formulation of the state equation is proposed. Several formulations of error paradigms are addressed and transformed into an algorithm involving a well-posed metric-optimization sub-problem. The preliminary results depicted here will be completed for the conference by works in progress dealing with these new formulations.

7 ACKNOWLEDGEMENTS

We thank Geraldine Olivier for fruitful discussions. This work was partly done in the MAIDESC ANR project which is supported by the French ministery of Research under contract ANR-13-MONU-0010.

REFERENCES

[1] F. Alauzet A. Loseille, A. Dervieux. Anisotropic norm-oriented mesh adaptation for compressible flows. In *53rd AIAA Aerospace Sciences Meeting*, 2015.

[2] P.-A. Absil, R. Mahony, and R. Sepulchre. *Optimization Algorithms on Matrix Manifolds*. Princeton University Press, Princeton, NJ, 2008.

[3] F. Alauzet, P.J. Frey, P.L. George, and B. Mohammadi. 3D transient fixed point mesh adaptation for time-dependent problems: Application to CFD simulations. *J. Comp. Phys.*, 222:592–623, 2007.

[4] F. Alauzet and G. Olivier. An L^p-L^∞ space-time anisotropic mesh adaptation strategy for time dependent problems. In *Proceedings of the V ECCOMAS CFD Conf.*, 2010.

[5] F. Alauzet and G. Olivier. Extension of metric-based anisotropic mesh adaptation to time-dependent problems involving moving geometries. In *49th AIAA Aerospace Sciences Meeting and Exhibit*, AIAA-2011-0896, Orlando, FL, USA, Jan 2011.

[6] V. Arsigny, P. Fillard, X. Pennec, and N. Ayache. Log-Euclidean metrics for fast and simple calculus on diffusion tensors. *Magn. Reson. Med.*, 56(2):411–421, 2006.

[7] A. Belme, A. Dervieux, and F. Alauzet. Time accurate anisotropic goal-oriented mesh adaptation for unsteady flows. *J. Comp. Phys.*, 231(19):6323 6348, 2012.

[8] M. Berger. *A panoramic view of Riemannian geometry*. Springer Verlag, Berlin, 2003.

[9] A. Loseille and F. Alauzet. Continuous mesh framework. Part I: well-posed continuous interpolation error. *SIAM J. Numer. Anal.*, 49(1):38–60, 2011.

[10] A. Loseille and F. Alauzet. Continuous mesh framework. Part II: validations and applications. *SIAM J. Numer. Anal.*, 49(1):61–86, 2011.

[11] A. Loseille, A. Dervieux, and F. Alauzet. Fully anisotropic goal-oriented mesh adaptation for 3d steady euler equations. *Journal of Computational Physics*, 229:2866–2897, 2010.

[12] A. Loseille, A. Dervieux, and F. Alauzet. Fully anisotropic goal-oriented mesh adaptation for 3D steady Euler equations. *J. Comp. Phys.*, 229:2866–2897, 2010.

[13] A. Loseille, A. Dervieux, P.J. Frey, and F. Alauzet. Achievement of global second-order mesh convergence for discontinuous flows with adapted unstructured meshes. In *37th AIAA Fluid Dynamics Conference and Exhibit*, AIAA-2007-4186, Miami, FL, USA, Jun 2007.

[14] F. Magoules. *Computational Fluid Dynamics*. CRC Press, Boca Raton, London, New York, Washington D.C., 2011.

[15] D.A. Venditti and D.L. Darmofal. Grid adaptation for functional outputs: application to two-dimensional inviscid flows. *J. Comp. Phys.*, 176(1):40–69, 2002.

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis (Eds)

APPLICATION OF HIGH-ORDER ELEMENTS FOR COUPLED ANALYSIS IN GEOMECHANICS

MINA KARDANI*, MAJIDREZA NAZEM* AND JOHN P. CARTER*

* Australian Research Council Centre of Excellence for Geotechnical Science and Engineering (CGSE)
The University of Newcastle
University Drive, Callaghan, NSW 2308, Australia
e-mail: mina.kardani@newcastle.edu.au, www.newcastle.edu.au/cgse

Key words: Coupled Problems, Geomechanics, and High-order elements.
Abstract. In this paper high-order triangular elements are implemented in the framework of the Arbitrary Lagrangian-Eulerian method for the analysis of large strain consolidation problems in geomechanics. The theory of consolidation, as well as details of the high-order elements, including cubic (10-noded), quartic (15-noded), quantic (21-noded) and sextic (28-noded) elements are discussed. The accuracy and the efficiency of high-order elements in the analysis of consolidation problems are demonstrated conducting a small deformation analysis of the soil under a strip footing as well as a large deformation analysis of a vertical cut subjected to a surcharge loading. Based on the numerical results, it is shown that high-order elements not only improve the accuracy of solution but can also significantly decrease the required computational time. It is also demonstrated that assuming identical order for displacement shape functions and the pore water pressure shape functions does not affect the stability of the time-marching analysis of consolidation nor the accuracy of the numerical predictions.

1 INTRODUCTION

The theory of linear consolidation was first proposed by Biot in 1941. However, it took several years before researchers applied the finite element method to solve Biot's consolidation equations, e.g., Sandu and Wilson (1969). The works of Christian and Boehemer (1970) and Kraus (1978) are the good examples of employing the finite element technique to solve elastic consolidation problems. Later, Small et al. (1976) proposed the first extension to Biot's theory to accommodate elastoplastic behaviour of the soil. Works of Carter et al. (1977) and Carter et al. (1979) are further extension of the theory of consolidation for elastic and elastoplastic soil subjected to large strains.

The finite element method has proven to be very effective for dealing with consolidation problems in geomechanics. However, simultaneously minimising the computational time while maximising the accuracy of numerical solution, is still a research subject of some interest. One of the ways to achieve this goal is to apply high-order elements in which increasing the order of the nodal polynomial functions may result in more accurate results in less computational time. The 15-noded triangular elements have already been applied by a few researchers to improve the solution of some well-known geotechnical problems.

Examples include study of the collapse load of an incompressible soil under strip and circular footings by Sloan and Randolph (1982), investigation of the collapse load of soil with a non-associated flow rule with small strains by De Borst and Vermeer (1984), and the consolidation problem of composite soft clay by Horpibulsuk et al. (2012). Recently, Kardani et al. (2013) compared the efficiency of high-order triangular elements such as quartic and quadratic elements by investigating the large deformation of an undrained soil under a footing. Later Kardani et al. (2014) compared the performance of the same set of high-order elements in analysing large strain coupled problems. When dealing with large deformation problem, the Arbitrary-Lagrangian-Eulerian (ALE) formulation presented by Nazem et al. (2008) has been applied to refine the mesh in order to prevent the occurrence of mesh distortion during the analysis.

In this work, the performance of 28-noded triangular elements in analysing the coupled problems of geomechanics is compared with the outcomes of 10-, 15- and 21-noded elements. The comparison will be shown by investigating the bearing capacity of the soil under a strip footing assuming small deformations as well as the stability of a vertical cutting in an undrained soil subjected to large deformations.

2 GOVERNING EQUATIONS

In geotechnical problems the deformations of the solid phase are usually coupled with the pore fluid pressures. In order to analyse these consolidation problems, the governing equations are obtained by coupling the conservation of mass and the equilibrium resulting the governing finite element equations (see Nazem *et al.* 2008):

$$\begin{bmatrix} \mathbf{K} & \mathbf{L} \\ \mathbf{L}^{\mathrm{T}} & \mathbf{0} \end{bmatrix} \begin{Bmatrix} \dot{\mathbf{u}} \\ \dot{\mathbf{p}} \end{Bmatrix} + \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{H} \end{bmatrix} \begin{Bmatrix} \mathbf{u} \\ \mathbf{p} \end{Bmatrix} = \begin{Bmatrix} \dot{\mathbf{F}}_{ext} \\ \dot{\mathbf{Q}}_{ext} \end{Bmatrix} \tag{1}$$

where $\mathbf{K}$ is the stiffness matrix, $\mathbf{L}$ represents the coupling matrix, $\mathbf{H}$ denotes the flow matrix, $\mathbf{u}$ and $\mathbf{p}$ are respectively the vectors of nodal displacements and pore water pressures, and $\mathbf{F}_{ext}$ and $\mathbf{Q}_{ext}$ represent the external force vector and the fluid supply vector, respectively. A superimposed dot in Equation (1) represents the time derivative of a variable. It is noted that the Arbitrary Lagrangian-Eulerian method proposed by Nazem *et al.* (2006) is employed for solving the large deformation problems in this study. The implicit backward Euler method is used to numerically integrate the coupled differential equation in (1). For further details, see Nazem *et al.* (2008).

3 HIGH-ORDER ELEMENTS

In this paper, high-order triangular elements including the 10-, 15, 21- and 28-noded elements are applied to discretise the problem domain. In triangular elements, the order of the polynomial shape function, p, is directly related to the number of nodes, m, according to (Dunavant, 1985)

$$m = \frac{1}{2}(p+1)(p+2) \tag{1}$$

In dealing with the coupled problem, it is generally accepted that the order of the pore

water pressure shape functions should be one degree lower than the order of the displacement shape functions. For the 6-noded elements this can easily be achieved by only considering pore pressure degrees of freedom at the corner nodes. However, with higher order elements such consideration is not feasible. To overcome the problem in this study, the same order of shape function is assumed for both displacement and pore water pressure. Later, through some numerical examples, it is demonstrated that this assumption is satisfactory. Table 1 contains the characteristic information for each type of high-order element considered, including number of nodes on each side of the element, the number of internal nodes, the polynomial order and the minimum number of quadrature points.

Table 1: Characteristic information for high-order triangular elements

Element type	No. of nodes per side	No. of internal nodes	Order of shape function	Plane strain		Axi-symmetric	
				Order of Integrand	Gauss points	Order of Integrand	Gauss points
10-noded	4	1	3	6	6	7	12
15-noded	5	3	4	8	12	9	16
21-noded	6	6	5	10	16	11	25
28-noded	7	10	6	12	25	13	37

4 NUMERICAL EXAMPLES

In this study, the accuracy and efficiency of high-order elements in tackling coupled problems is investigated by analysing two geotechnical problems. In the first example, a soil layer under a strip footing is analysed to find its undrained bearing response while assuming small deformations. In the second example, the loading of a vertical cutting is studied assuming large deformations. In both examples, the study is focused on the possibility of achieving a prescribed accuracy with fewer degrees of freedom while increasing the order of the elements.

It is noted that the high-order elements have been implemented in SNAC, a finite element program developed for analysing geotechnical problems at the University of Newcastle.

4.1 Elastoplastic analysis of a soil layer under a strip footing

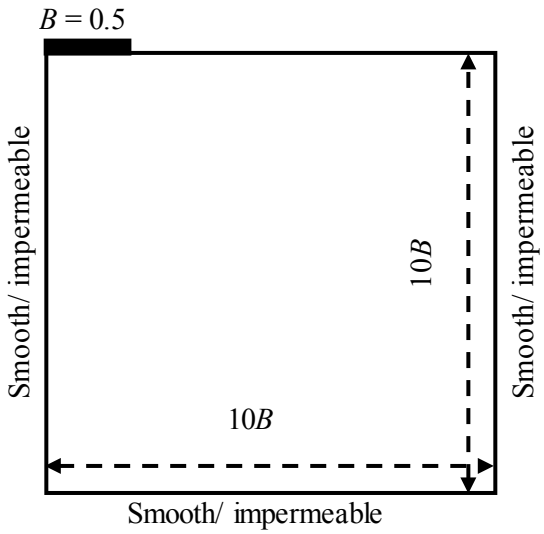
In this example, the efficiency of high-order element in dealing with geotechnical coupled problems is shown by small deformation analysis of the undrained bearing response of a soil layer under a rigid strip footing. Small (1977) showed that for a weightless soil the drained and undrained strength parameter must satisfy

$$\frac{c_u}{c'} = \frac{2\sqrt{N_\phi}}{1+N_\phi} \qquad where \qquad N_\phi = \frac{1+\sin\phi'}{1-\sin\phi'} \tag{1}$$

in which ϕ' and c' are respectively the drained friction angle and cohesion, and c_u denotes the undrained shear strength of the soil. In a coupled analysis, the excess pore water pressure will not have sufficient time to dissipate provided that the loading rate is relatively fast rate (see Small 1977). Based on Prandtl's plasticity solution, the undrained bearing capacity of soil under a strip footing is obtained by

$$q_u = N_c c_u \tag{1}$$

where $N_c=2+\pi$ represents the bearing capacity factor. The Mohr-Coulomb material model is used to predict the soil behaviour in this example. To avoid an artificial increase in shear strength of soil due to the suppression of any dilation under undrained (constant volume) conditions, it is important to assume that dilation angle is zero. The problem domain, boundary conditions and material properties are illustrated in Figure 1. Note that only the right half of the problem domain is considered in the analysis due to symmetry. Also, plane strain conditions are assumed.



Soil Properties

Drained Young's modulus, $E' = 500\ kPa$
Drained Poisson's ratio, $\nu' = 0.3$
Drained shear strength, $c' = 0.01E'$
Drained friction angle, $\phi' = 20^o$
Dilation angle, $\psi' = 0^o$
Permeability, $k = 5\times10^{-9}\ m/s$
Unit weight of the soil, $\gamma = 0\ kN/m^3$
Unit weight of the water, $\gamma_w = 10\ kN/m^3$

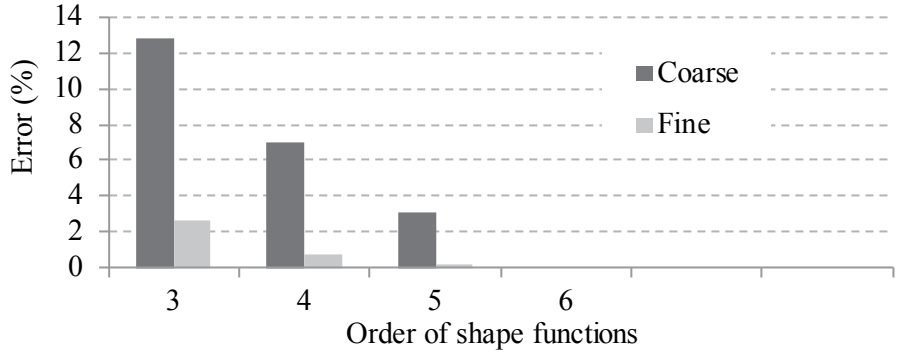
Figure 1: An undrained layer of soil under a strip footing

To provide a comparison of the efficiency of the high-order elements, two fixed grids were used to discretise the problem domain. For a meaningful comparison it is important to use more or less an identical number of degrees-of-freedom regardless of the element type. First, the problem domain is discretised by an uniform 60x60 grid which provides a relatively coarse mesh with 3721 nodal points, noting that 60 is the least common multiple of 3, 4, 5 and 6 (number of segments on one side of 10-, 15-, 21-, and 28-node elements, respectively). The second grid is relatively fine, including 120x120 uniform divisions and 14641 nodal points. Both uniform grids are used to form triangular meshes by applying 10-, 15-, 21- and 28-node elements.

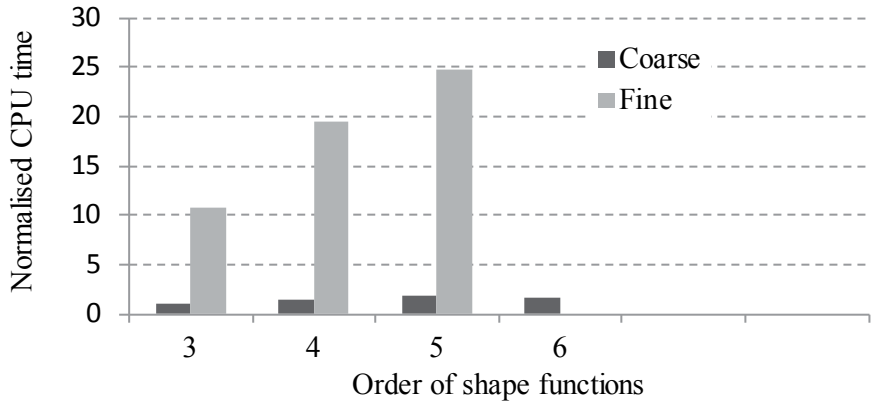
Table 2 presents the number of elements and integration points in each mesh as well as the numerical results including the error in predicted N_c and the CPU time normalised by the CPU time of the fastest analysis. The plots of error versus the order of the shape functions, as well as the plots of error versus the normalised CPU time, are presented in Figures 2a and 2b, respectively. According to Table 2, the coarse mesh of 28-noded elements provides the most accurate estimation of the undrained bearing capacity of the soil and hence there is no need to repeat the analysis using the fine mesh of 28-noded elements.

Table 2: Finite element meshes and numerical result

Element Type	Mesh	Number of elements	Number of Gauss points	N_c	Error in N_c (%)	Normalised CPU time
10-noded	Coarse	800	4800	5.80	12.84	1
	Fine	3200	19200	5.279	2.7	10.85
15-noded	Coarse	450	5400	5.50	7.0	1.4
	Fine	1800	21600	5.186	0.77	19.43
21-noded	Coarse	288	4608	5.301	3.13	1.84
	Fine	1152	18432	5.141	0.18	24.8
28-noded	Coarse	200	5000	5.140001	1.9E-7	1.6
	Fine	800	20000	-	-	-



a. Error versus order of shape functions



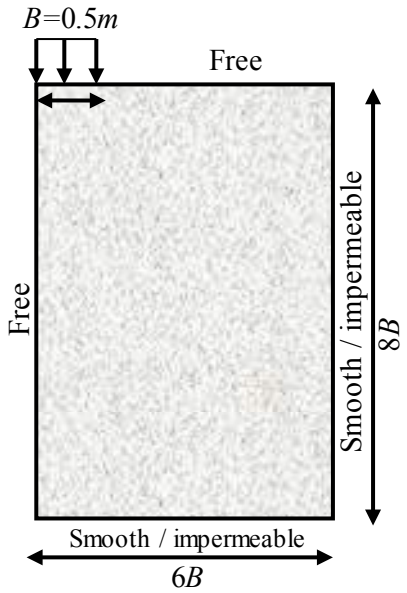
b. Normalised CPU time versus order of shape functions

Figure 2: Numerical results

Based on the result presented in Table 2 and Figure 2, the coarse mesh with cubic elements represents the fastest analysis, but the bearing capacity of the soil is considerably overestimated. On the other hand, the analysis with 28-noded elements with a CPU time just 1.66 times the fastest analysis, achieves the most accurate result. Also in this example, by comparing the analysis results of 21-noded and 28-noded elements, it is concluded that the 28-noded elements are more efficient than 21-noded elements as they not only improve the final solution but decrease the computational time.

4.2 Vertical cut

In this example, the efficiency of high-order element in dealing with coupled geotechnical problem is studied by large deformation analysing of a weightless soil under a uniform vertical load applied adjacent to a vertical cutting.

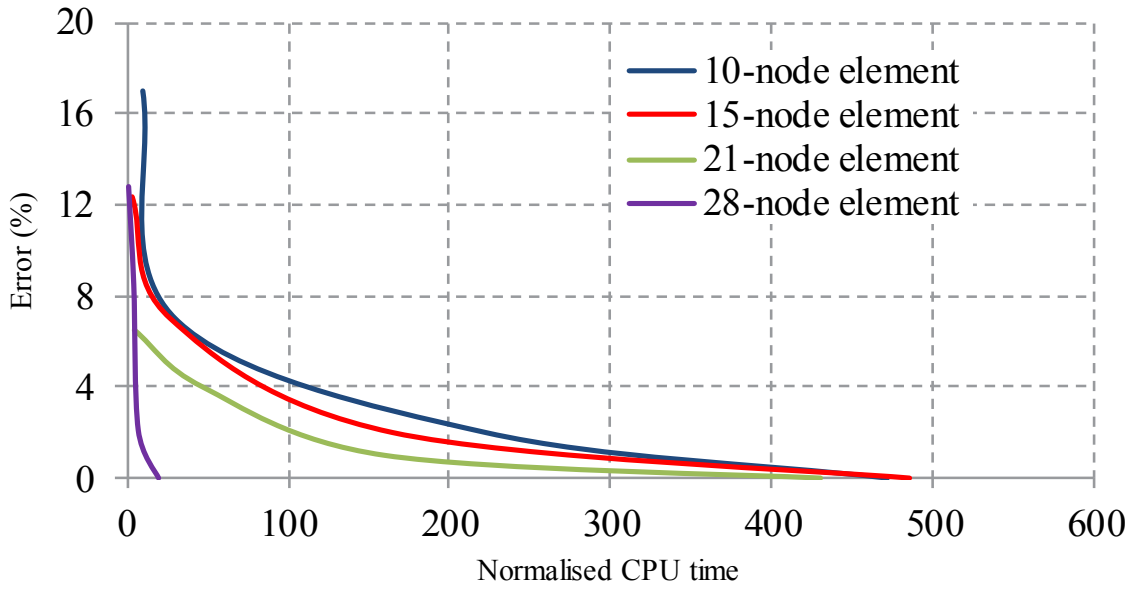


Soil Properties

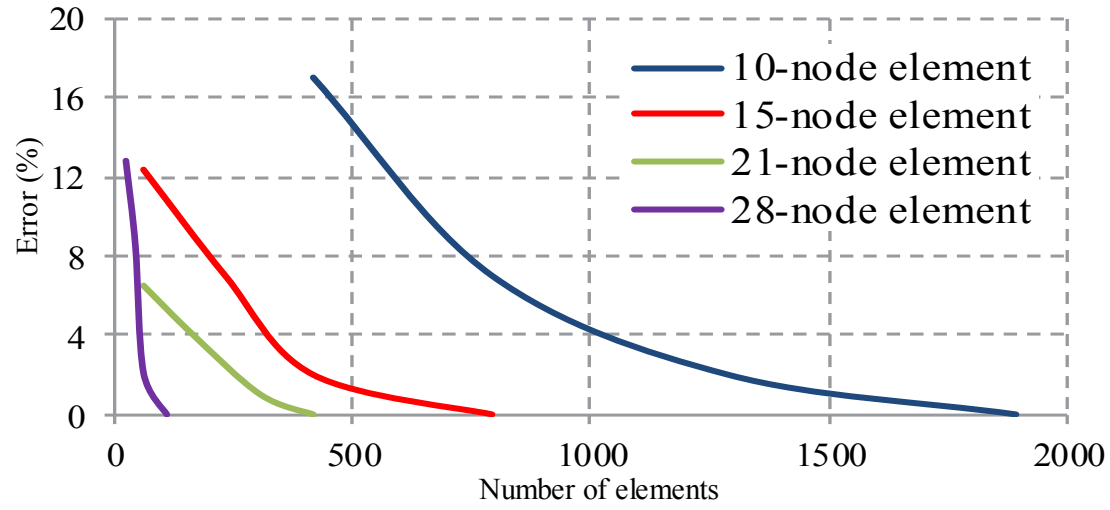
Drained Young's modulus,	$E' = 100\,kPa$
Drained Poisson's ratio,	$\nu' = 0.3$
Drained shear strength,	$c' = 5.0 kPa$
Drained friction angle,	$\phi' = 20°$
Dilation angle,	$\psi' = 0°$
Permeability,	$k = 5 \times 10^{-9}\ m/s$
Unit weight of the soil,	$\gamma = 0\ kN/m^3$
Unit weight of the water,	$\gamma_w = 10\ kN/m^3$

Figure 3: A vertical slope under vertical prescribed displacement

To model the undrained soil behaviour, the drained material properties are considered and a coupled pore water pressure-displacement analysis is carried out with a rather fast loading rate. The problem domain, material properties and boundary condition are shown in Figure 3. Based on the theoretical solution, the maximum vertical pressure which can be applied on an undrained soil is equal $2c_u$. The analyses were conducted using a trial and error strategy. For each type of high-order elements, the first analysis employed a very coarse mesh and the error was calculated. The analysis was then repeated by gradually increasing the density of the mesh, specifically in the area under the load, with the aim of approaching the exact solution. The characteristics of the discretisation as well as the error and CPU time of each analysis are summarised in Table 3. For each analysis, the errors versus number of elements as well as the error versus normalised CPU time are plotted in Figure 4. According to Figure 4a, by increasing the number of elements the error decreases, whereas the computational time increases. According to the Figure 4b, the performance of the high-order element improves as



a. Error versus number of elements in the vertical cut.



b. Error versus normalised computational time

Figure 4: Analysis results of vertical cut

the order of element is increased. The 28-noded element achieved the same accuracy of the other type of elements with the lowest computational time. This given accuracy can be achieved in the fastest time using the 28-noded element.

5 CONCLUSION

In this study, high-order finite elements are applied for coupled problems in a finite element framework, using the Arbitrary Lagrangian-Eulerian method. The accuracy and efficiency of these elements were examined by analysing two typical geotechnical problems including the undrained bearing capacity of soil under a rigid footing and the behaviour of undrained soil under a vertical pressure applied adjacent to a cutting.

For the coupled problems investigated in this study, applying high-order elements was proven to be very effective in improving the computational time. This means that increasing the order of elements leads to a decrease in the number of degrees of freedom required to achieve a given accuracy and therefore a significant drop in the computational time. Also, it was demonstrated that considering the same order of shape functions for both displacement and pore water pressure does not affect the accuracy of the numerical results. In dealing with the coupled problems studied in this paper, the 28-noded triangular elements prominently outperformed the other high-order elements.

Table 3: Descritisations in vertical cut problem

Element type	Number of Elements	Degrees of freedom	Gauss points	Error (%)	Normalised CPU time
10-node	416	5790	2496	17	9.8
	792	10938	4752	7	29.97
	1289	17697	7734	2.0	222.58
	1894	25878	11364	0	471.8
15-node	59	1533	708	12.35	3.2
	231	5733	2772	7	26.98
	416	10215	4992	2	164.78
	792	19335	9504	0	485.59
21-node	59	2367	944	6.5	4.58
	164	6333	2624	4	46.98
	305	11733	4880	1	159.79
	416	15888	6656	0	430.396
28-node	22	1317	550	12.8	1.0
	42	2415	1050	8.5	4.2
	59	3360	1457	2.1	7.0
	108	6051	2700	0	19.6

REFERENCES

[1] Sandhu, R.S., and Wilson, E.L. Finite element analysis of seepage in elastic media. *J. of Eng. Mechanics* (1969) **95**: 641-652.

[2] Christian, J.T., and Boehmer, J.W. Plain strain consolidation by finite elements. *J. of the Soil Mechanics and Foundation Division of ASCE* (1970) 1435-1457.

[3] Small, J.C., Booker, J.R., and Davis, E.H. Elasto-plastic consolidation of soil. *Int. J. of Solids and Structures* (1976) **12**: 47-76.

[4] Carter, J.P., Small, J.C., and Booker, J.R. A theory of finite elastic consolidation. *Int. J. of Solids and Structures* (1977) **13**: 467-478.
[5] Krause, G. Finite element schemes for porous media. *J. of the Eng. Mechanics Division, ASCE* (1978) **104**: 605-620.
[6] Carter, J.P., Booker, J.R., and Small, J.C. The analysis of finite elasto-plastic consolidation. *Int. J. for Num. and Analy. Meth. in Geomechanics* (1979) **3**: 107-129.
[7] Sloan, S.W., and Randolph, M.F. Numerical Prediction of collapse loads using the finite element methods. *Int. J. of Num. and Analy. Meth. in Geomechanics* (1982) **6**: 47-76.
[8] De Borst, R., and Vermeer, P.A. Possibilities and limitations of finite elements for limit analysis. Géotechnique (1984) **34**: 199-210.
[9] Nazem, M., Sheng, D., and Carter J.P. Stress integration and mesh refinement in numerical solutions to large deformations in geomechanics, *Int. J. Numer. Meth. Eng.* (2006) **65:** 1002-1027.
[10] Nazem, M., Sheng, D., Carter, J.P., and Sloan, S.W. Arbitrary Lagrangian-Eulerian method for large-strain consolidation problems. *Int. J. for Num. and Analy. Meth. in Geomechanics* (2008) **32**: 1023-1050.
[11] Horpibulsuk, S., Chinkulkijniwat, A., Cholaphatsron, A., Suebsuk, J., and Liu, M.D. Finite element analysis of consolidation behaviour of composite soft ground. *GeoCongress* (2012) 940-947
[12] Kardani, M., Nazem, M., Carter, J.P., and Abbo, A.J., The efficiency of high-order elements in large deformation problems of geomechanics. *Int. J. for Geomechanics* (2014) 10.1061/(ASCE)GM.1943-5622.0000457 , 04014101.
[13] Kardani, M., Nazem, M., Carter, J.P., and Abbo, A.J., Large strain coupled analysis of geotechnical problems using high-order elements. *Australian Geomechanics* (2014) **49**:167-174

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis (Eds)

COEFFICIENT OF CONSOLIDATION FOR SOIL – THAT ELUSIVE QUANTITY

DONG WANG*, MARK F. RANDOLPH* AND SUSAN GOURVENEC*

* Centre for Offshore Foundation Systems
The University of Western Australia
35 Stirling Highway, Perth, WA 6009, Australia
e-mails: dong.wang@uwa.edu.au, mark.randolph@uwa.edu.au, susan.gourvenec@uwa.edu.au
www.cofs.uwa.edu.au

Key words: Coefficient of Consolidation, Elastoplastic Response, Modified Cam Clay, Soil Mechanics, Stiffness.

Abstract. Although it is accepted that the coefficient of consolidation for soil is not a true material property, but reflects the net effect of permeability and compressibility, it is a very useful parameter in day to day design. Design calculations make extensive use of elastic solutions for consolidation, such as beneath a shallow foundation or around a driven pile, but an important consideration is how to measure or estimate an appropriate coefficient of consolidation to use in those solutions. Typically the quantity is determined either from laboratory oedometer tests (generally then referred to as c_v) or from field dissipation tests using a piezocone or piezoball penetrometer (generally then referred to as c_h). Since the latter form of test includes a mix of stress paths, for some of which the soil has a stiffness associated with unloading and others of which involve plastic compression, the magnitude of c_h for a given soil is typically 3 to 10 times the value of c_v from virgin compression in laboratory oedometer tests. The paper explores the relationship between c_v and c_h for different boundary value problems, within the confines of soil modelled as Modified Cam Clay, for both isotropic and anisotropic permeability. Problems range among: simulated oedometer testing, field dissipation testing and pore pressure response beneath a shallow foundation. Results of finite element analysis of this range of problems are used to develop guidelines for different classes of problem, comparing the relevant coefficient of consolidation against a benchmark c_v value associated with virgin compression in an oedometer. The normalised values of consolidation coefficient are expressed as functions of fundamental soil parameters used within Modified Cam Clay.

1 INTRODUCTION

The coefficient of consolidation, c_v, was originally presented in Terzaghi's classical one-dimensional consolidation theory to estimate foundation settlement under vertical loading. The coefficient is usually determined through laboratory oedometer or Rowe cell tests, or from in situ pore pressure dissipation tests with piezocone or piezoball. In each type of test, the value of consolidation coefficient is deduced by comparison of measured data with a theoretical response curve, generally derived from simple elastic response of the soil. In an oedometer or Rowe cell test, the one-dimensional compression of the soil sample for a given stress increment

is plotted against either the square root of time or the logarithm of time in order to deduce c_v [1]; in a field dissipation test, the excess pore pressure decay is plotted against the logarithm of time and matched to a theoretical response [2] in order to deduce an 'operative' coefficient of consolidation, usually expressed as c_h. No formal relationship between c_h and c_v has been established, although in practice the former is generally found to be significantly greater, by a factor of 3 to 10, than the latter. In both laboratory and field tests, uncertainties arise because of the need to normalise the measured responses, identifying appropriate initial and final values of the measured parameters.

In this paper, the coefficient of consolidation is interpreted from either finite strain or large deformation finite element (FE) analyses that incorporate an elastoplastic critical state soil model, Modified Cam Clay (MCC). The objective is to identify relevant 'operative' values of consolidation coefficient for different boundary value problems, devising relationships between the different values in terms of true soil parameters. Comparisons are made with experimental data obtained from laboratory and centrifuge model tests at the University of Western Australia, using normally consolidated kaolin clay, for which MCC parameters are well established. All FE simulations have been conducted in the framework of finite strain rather than small strain, using the commercial package Abaqus/Standard [3]. The soil is discretised with 8-node axisymmetric elements with reduced integration and pore pressures at 4 corner nodes (termed CAX8RP in Abaqus).

2 SOIL PROPERTIES

The properties of the kaolin clay considered in this study are listed in Table 1. The soil sample is assumed to have isotropic permeability ($k_h/k_v = 1$, where subscripts h and v indicate horizontal and vertical directions), as has been demonstrated in various centrifuge tests [4,5]. The effect of anisotropic permeability on dissipation of excess pore pressure will also be explored numerically.

According to results from Rowe cell tests [6], the coefficient of consolidation for normally consolidated conditions may be fitted as:

$$c_v = \sqrt{0.001 + 0.014\sigma'_v / p_a} \quad \text{mm}^2/\text{s} \tag{1}$$

where σ'_v is the vertical effective stress and p_a is atmospheric pressure (100 kPa). The coefficient of consolidation was determined through the square root method (Taylor's method). For normally consolidated soil under a K_0 state, the relationship between void ratio and vertical effective stress within the MCC model is

$$e = (e_N - C) - \lambda \ln\left(\frac{1+2K_0}{3}\sigma'_v\right) \tag{2}$$

where C is the distance between the virgin consolidation line (VCL) and one-dimensional normal consolidation line, and K_0 is the coefficient of earth pressure at rest. Equations 1 and 2 allow the permeability of the kaolin to be determined as a function of the void ratio e:

$$k_v = \gamma_w m_v c_v = 1.06\times10^{-5}\,\frac{\gamma_w \lambda(1+2K_0)}{(1+e)\exp\left((e_N - C - e)/\lambda\right)}\sqrt{1+\frac{0.42\exp\left((e_N - C - e)/\lambda\right)}{1+2K_0}} \quad \text{mm/s} \tag{3}$$

where γ_w is the unit weight of water and m_v is the coefficient of volume compressibility. The relationship for permeability given by Eq. 3 has been used successfully in recent numerical studies [5,7]. However, the accuracy of this equation is limited for two reasons: (i) the void ratio in a one-dimensional test varies with stress level and time during a consolidation stress increment; (ii) the coefficient of consolidation is determined from an empirical method (in this case Taylor's root time method rather than the log time method of Casagrande).

Table 1: MCC parameters for kaolin clay

Properties (after [8])	Values
Angle of internal friction, ϕ'	23°
Void ratio at $p' = 1$ kPa on virgin consolidation line, e_N	2.252
Slope of normal consolidation line, λ	0.205
Slope of swelling line, κ	0.044
Plastic compression ratio, $\Lambda = 1 - \kappa/\lambda$	0.79
Poisson's ratio, v	0.3
Submerged unit weight, γ'	6.18 kN/m^3

3 OEDOMETER TEST

To verify the effectiveness of Eq. 3, a conventional oedometer test was simulated using finite strain FE analysis. The soil sample was 75 mm in diameter and 20 mm high, with permeable top and bottom faces. The container was assumed fully smooth. Prior to one-dimensional consolidation, a small pressure of 5 kPa was applied on the soil surface to generate initial effective stresses within soil sample, with K_0 selected as 0.75 rather than the more usual estimate of 1 - sinϕ' (= 0.61). The reason is that the shape of the MCC yield envelope, assuming associated flow, automatically leads to a ratio between radial and vertical effective stress during one-dimensional compression between 0.7 ~ 0.8. The initial void ratio was thus calculated as e_0 = 1.942 and C in Eq. 2 was 0.017. The vertical stress was increased from 5 kPa to 55, 105, 205, 405 and 505 kPa. Each stage lasted 8 h, which was sufficiently long for full dissipation of excess pore pressures. The elastic part of the MCC model was described with constant Poisson's ratio of 0.3. Once the simulation was completed, the value of c_v corresponding to each loading stage was derived through the square root time method.

Figure 1 shows the variations of c_v from Eq. 1 and from the FE simulation. Agreement is reasonable, suggesting that the permeability can be quantified through Eq. 3 with satisfactory accuracy. Terzaghi's 1-D consolidation theory is based on the simplifications that the soil is homogeneous and the permeability is constant during each loading stage. Additionally, there is always some divergence of c_v obtained through different empirical methods. The good agreement highlighted in Figure 1 is probably due to the reason that both the coefficient of compressibility

$$m_v = \frac{\lambda}{(1+e)\sigma'_v} \tag{4}$$

and permeability depend on the current void ratio, but the void ratio is varied in a relatively small range for each loading stage.

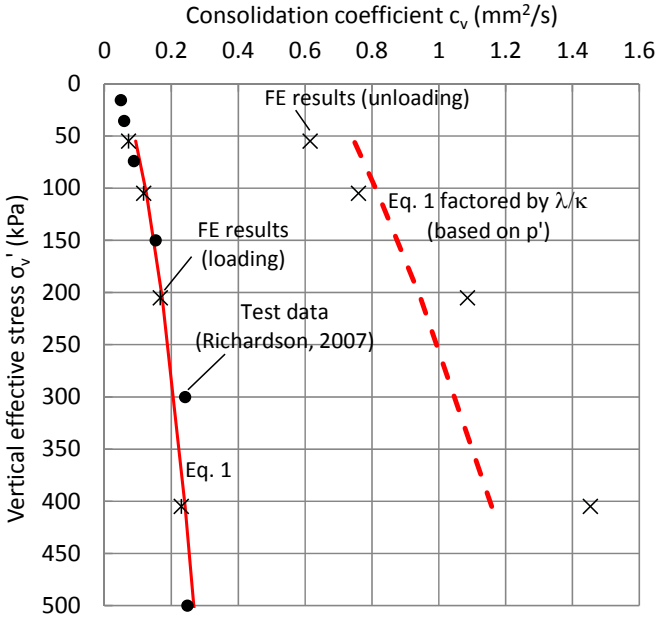


Figure 1: Coefficient of consolidation estimated by FE

The unloading phase was also simulated, reducing the vertical stress applied to the soil surface in stages to 405, 205, 105 and 55 kPa. The resulting $c_{v,unloading}$ values are shown in Figure 2, where they are compared with Eq. 1, factored by λ/κ. (Note the equation was recast to be based on p', taking account of the changing K_0 during 1-D unloading.) As may be seen, the FE data are broadly similar to the factored equation, but with a greater variation with respect to effective stress level. In fact it is difficult to 'predict' the operative consolidation coefficient for each stage of unloading, due to complex changes in the effective stress level, K_0 and void ratio that take place (non-uniformly with time through the sample), as illustrated below.

The compression and swelling curves from the loading and unloading stage are shown in e - σ'_v space in Figure 2. While the loading response matches exactly the input value of λ, the unloading response shows a gradient that gives equivalent 'κ' values (but with respect to σ'_v rather than p') that decrease from 0.032 to 0.014, compared with the input value of $\kappa = 0.044$. The discrepancy is due to the gradually increasing K_0 value during 1-D unloading, as shown in Figure 3, and indeed the $e - p'$ response matches exactly $\kappa = 0.044$. The difference in deduced values of κ obtained in either e - σ'_v or $e - p'$ space is an aspect that is generally over-looked in deriving input parameters for soil models.

Although the above discussion is somewhat pedestrian in that it covers well-known aspects of soil response, it is included in order to provide background data that illustrate the rather complex nature of the consolidation coefficient, as applied to even very simple problems. The following sections extend the study to explore 'operative' values of consolidation coefficient obtained from two different applications, consolidation around a penetrometer, and consolidation beneath a shallow foundation.

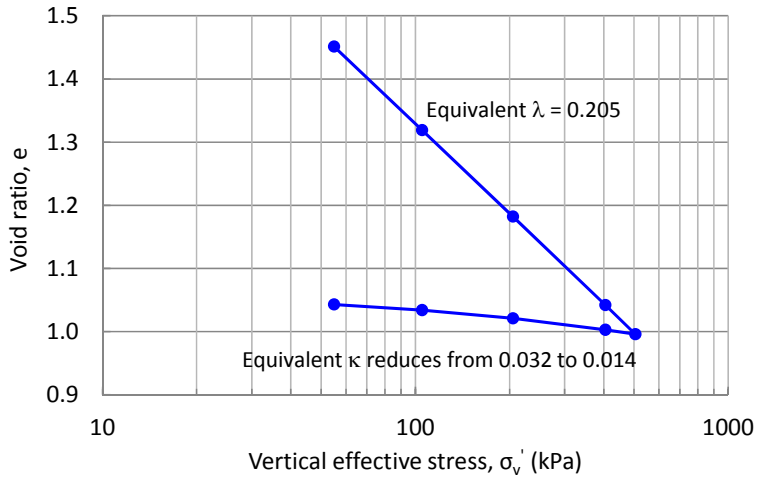


Figure 2: Compression and swelling curves from FE simulation of oedometer test

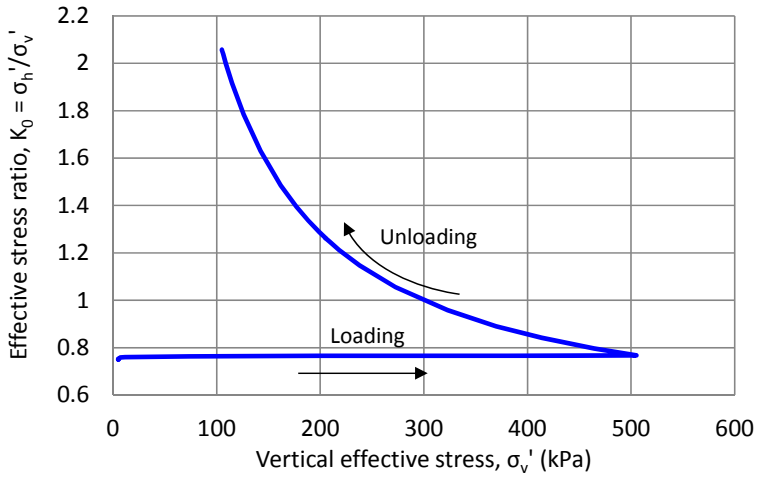


Figure 3: Effective stress ratio during numerical reproduction of compression and swelling

4 PIEZOBALL DISSIPATION TESTS IN NORMALLY CONSOLIDATED CLAY

Full-flow penetrometers such as the T-bar [9] and ball [10] have been used increasingly during the last decade, in particular in soft seabed sediments due to their advantages of larger projected area and minimal correction for overburden pressure. A ball penetrometer fitted with pore pressure transducer, a so-called piezoball [10], is of special interest as it can provide information on the soil consolidation properties by means of dissipation tests. Compared with one-dimensional consolidation laboratory test, the piezoball dissipation test is carried out in natural soil, without the need to obtain (nominally) undisturbed soil samples for later testing in the laboratory. This is especially important for offshore practice, since the cost of obtaining high-quality samples for laboratory testing are very high, and the low shear strengths in the upper few metres of the seabed pose particular practical difficulties.

In this section, the dissipation of excess pore pressure adjacent to the piezoball, following penetration, is studied using a large deformation finite element (LDFE) approach. The aim is to

establish a procedure for estimating c_h from the dissipation response for general anisotropic soil permeability. The LDFE approach is based on frequent mesh regeneration, in order to overcome soil element distortion induced by penetration of the ball. The details of the LDFE approach can be found in [11,12].

A centrifuge test [13] was reproduced to verify the reliability of the LDFE approach, assuming isotropic permeability, followed by investigation of the effect of anisotropic permeability. The model piezoball had a diameter of D_b = 15 mm and shaft diameter of d = 5 mm; the polypropylene filter element was fitted at either the mid-face (half the radius vertically from the tip of the ball probe) or the equator position; and the tests were conducted in normally consolidated kaolin at an acceleration level of 110g. The piezoball was penetrated to depth of 160 mm (estimated $\sigma'_v \sim 109$ kPa) at a rate of 1 mm/s, and then the dissipation test was conducted maintaining position of the piezoball. To simplify the analyses, avoiding major distortion of the soil surface, the piezoball in the LDFE analyses was pre-embedded at a depth of 130 mm before penetrating by 2 diameters, rather than simulating penetration from the soil surface. The shafted-piezoball was simplified as fully smooth. The coefficient of earth pressure was taken as $K_0 = 1 - \sin\phi' = 0.61$, and C in Eq. 2 was calculated as 0.048.

To explore the effect of anisotropic permeability, the ratio between horizontal and vertical permeability, $n = k_h/k_v$, was changed from 1 to 2, 5 and 10. Note that in the LDFE simulations, k_h and k_v were in terms of global Cartesian coordinates. The penetration rate used in the test, 1 mm/s, was sufficient to ensure essentially undrained conditions for n = 1. However, our trial calculations showed that the penetration phase may allow some partial consolidation for $n > 1$. The penetration rate was thus increased to 10 mm/s in the LDFE simulations. The pore pressures induced in soil with n = 1 are not affected significantly by the increase of penetrate rate, since the responses under undrained conditions are similar. Only the dissipations at mid-face of the ball are discussed here, as this position was deemed superior to the equator position for estimating c_h [7].

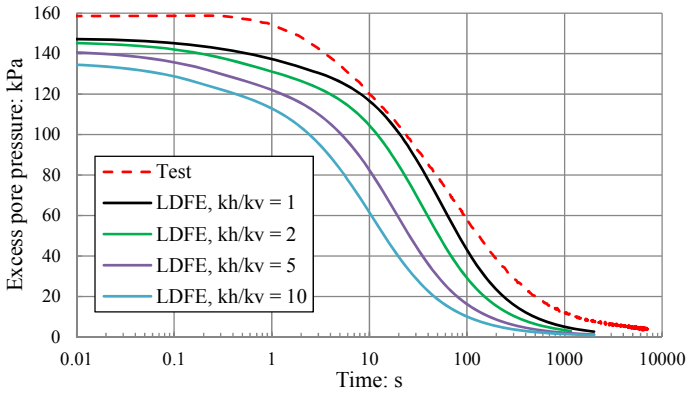
All the numerical dissipation responses are shown with the experimental data in Figure 4a. The numerical excess pore pressure at the end of penetration reduces from 147 to 135 kPa as n is increased from 1 to 10. This slight divergence suggests that nearly undrained conditions are achieved for permeability ratios no larger than 10. For soil with isotropic permeability, the measured excess pore pressure at the start of dissipation is slightly higher than the FE result. [7] suggested that the excess pore pressure is best normalised with an idealised initial excess pore pressure Δu_{ext} which is estimated using a back-extrapolation technique based on the square root of time [14]. An extensive parametric study, assuming isotropic permeability led to the suggestion that the consolidation time is best normalised as

$$T_b = \frac{c_h t}{D_b d I_r^{0.25}} \quad \text{with} \tag{5}$$

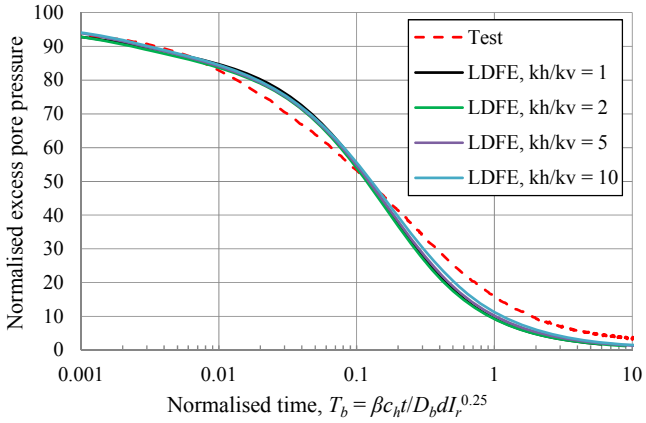
$$c_h = \frac{3(1-\nu)}{(1+\nu)}\left(\frac{\lambda}{\kappa}\right)^{\alpha} c_v = \frac{3(1-\nu)}{(1+\nu)} \frac{(1+e)p'}{\kappa^{\alpha}\lambda^{1-\alpha}} \frac{k_v}{\gamma_w} \tag{6}$$

where c_h is the operative coefficient of consolidation, I_r the rigidity index (equal to 73 here), p' the initial mean effective stress at the dissipation depth and α is a fitting parameter selected as 0.75. Note that this value of α, closer to unity than zero, reflects the stress paths followed during

the dissipation process, which tend to include a high component of (elastic) 'reloading', even for the initially normally consolidated soil conditions [7].



(a) Dimensional (time is for centrifuge model)



(b) Normalised

Figure 4: Dimensional and normalised dissipation responses at mid-face of piezoball

The normalised time from Eq. 5 is only for isotropic permeability. For permeability ratio $n > 1$, Eq. 5 may be modified as

$$T_b = \frac{\beta c_h t}{D_b d I_r^{0.25}} \tag{7}$$

with the factor β fitted as

$$\beta = \frac{n}{2} + \frac{1}{2} \tag{8}$$

The resulting normalised dissipation graphs are shown in Figure 4b. By introducing the factor β, the normalised numerical graphs with n ranging through 1 ~ 10 become nearly unique, establishing the effectiveness of the normalisation. Equation 7 indicates that the rate of dissipation at the mid-face of the piezoball depends on the average permeability in the vertical and horizontal directions. In contrast, the dissipation at the shoulder (u_2 position) of the piezocone is generally assumed to be governed by the horizontal permeability.

The unique normalised dissipation graph provided in Figure 4b can be used to estimate c_v

once the permeability ratio in a particular field is known. In situ test data for the dissipation of excess pore pressure as a function of time are interpreted through the following procedure:

(1) A nominal initial excess pore pressure, Δu_{ext}, at the start of dissipation is evaluated using the root time back-extrapolation technique [14].

(2) The dissipation time is then normalised using Eq. 7 (assisted by Eq. 8) with an assumed operative coefficient of consolidation c_h. The value of c_h is adjusted until the experimental normalised dissipation graph matches the numerical curves, such as those in Figure 4b. Note that it is necessary to assume a permeability ratio n a priori, or to undertake appropriate laboratory tests to evaluate the degree of permeability anisotropy.

(3) The coefficient of consolidation c_v at the dissipation depth of piezoball is estimated as (see Eq. 6):

$$c_v = \frac{(1+\nu)}{3(1-\nu)}\left(\frac{\kappa}{\lambda}\right)^{\alpha} c_h \tag{9}$$

A final comment relates to the application of the coefficient of consolidation to assess the degree of drainage during moving boundary problems, such as a penetration test [13,15]. The normalised velocity

$$V = \frac{vD}{c_v} \tag{10}$$

where v is the actual velocity and D is the relevant dimensions (e.g. diameter of penetrometer) may be used to quantify the degree of consolidation occurring during the motion. Generally, limits of $V < 0.3$ for fully drained conditions and $V > 30$ for undrained conditions are applied. However, it seems logical that these limits should really be expressed in terms of the same coefficient of consolidation as determined in a pore pressure dissipation test. In the case of a piezoball, that would change the limits for partial consolidation to (about) $0.1 < V' < 10$, where $V' = vD/c_h$. These limits match the experimental data of [13].

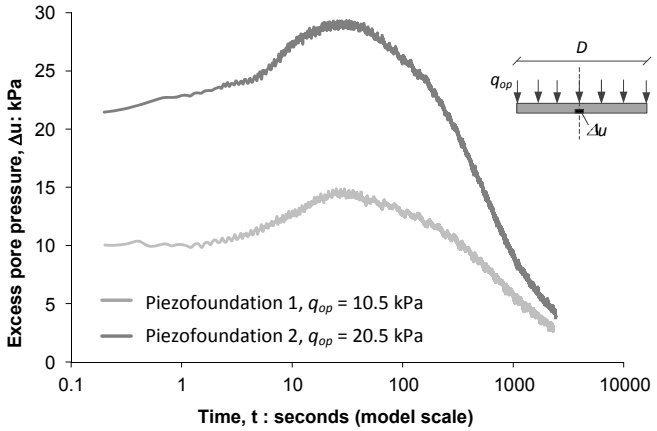
5 CHOICE OF CONSOLIDATION COEFFICIENT IN FOUNDATION DESIGN

Just as the operative coefficient of consolidation c_h for a piezocone or piezoball dissipation test is greater than the corresponding c_v value from a laboratory consolidation test, so the relevant coefficient for excess pore pressure dissipation or consolidation settlement beneath a shallow foundation may differ from c_v because of the range of stress paths followed in the underlying soil.

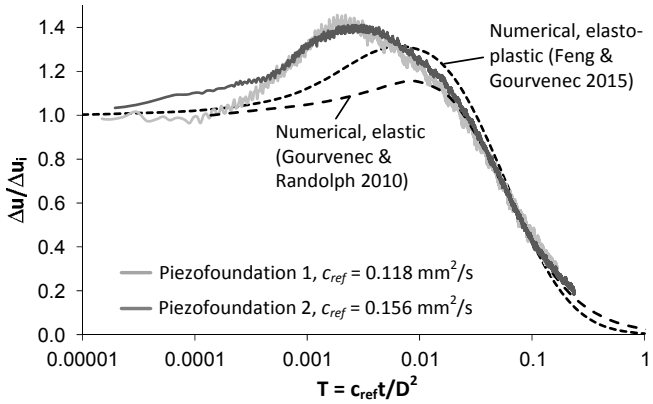
The operative coefficient of consolidation governing dissipation beneath a rigid, circular foundation resting on the surface of normally consolidated kaolin clay was investigated through centrifuge model tests [16]. The so-called 'piezofoundation' comprised a rough, rigid circular plate equipped with a pore pressure sensor at the centre of the baseplate. Two dissipation tests were carried out at constant (average) vertical stress levels of 10.5 kPa (PF1) and 20.5 kPa (PF2). The dissipation time histories from each test are shown in Figure 5a.

The piezofoundation tests exhibit the characteristic Mandel-Cryer effect with excess pore pressure increasing above the initial value of stress change during the early stage of consolidation. The Mandel-Cryer effect is a stress transfer effect significant in three-dimensional consolidation, resulting from the more rapid dissipation of excess pore pressure in

the soil near the edges of the foundation than near the centre. The resulting (greater) compression of the soil at the edges leads to a temporary (initial) transfer of total stress and hence increase in excess pore pressure in the central part of the foundation.



(a) Measured excess pore pressure dissipation



(b) Measured and numerical prediction of normalised excess pore pressure dissipation

Figure 5: Dissipation time histories for piezofoundation tests

Values of the operative coefficient of consolidation were determined by fitting the normalised measured dissipation curves to solutions based on elastic and elasto-plastic (Modified Cam Clay soil with isotropic permeability) finite element analyses of a rough, rigid, surface plate [17,18], as illustrated in Figure 5b. The excess pore pressure was normalised by the initial value and the consolidation time by a dimensionless time factor $T = c_{ref}/tD^2$, where D is the diameter of the foundation (= 40 mm). Operative values of the coefficient of consolidation, c_{ref}, of 0.12 mm^2/s and 0.16 mm^2/s for tests PF1 and PF2, with foundation loads of 10.5 kPa and 20.5 kPa respectively, were deduced in order to fit the numerically derived solutions.

Complementary piezocone tests were carried out in the centrifuge sample in order to relate the operative coefficient of consolidation for the foundation response, c_{ref}, to that measured by the piezocone, c_h. The piezocone data was normalised with the idealised initial excess pore

pressure value, Δu_{ext}, estimated using back-extrapolation based on the square root of time method [14], as applied to the piezoball results presented in Section 4. The consolidation time was normalised as a time factor, denoted T^* [2], similar to that applied to the piezoball interpretation presented in Section 4.

$$T^* = \frac{c_h t}{R^2 I_r^{0.5}} \tag{11}$$

where R is the piezocone radius and I_r is a rigidity index, taken as 88, following [19]. The normalised time histories are shown in Figure 6. Values of the coefficient of horizontal coefficient of consolidation, c_h, were extrapolated from these dissipation curves using a correlation with an established theoretical solution based on T^*_{50}, the time for 50% excess pore pressure dissipation [2]. Values range from $0.24 < c_h$ (mm^2/s) < 0.41 and are plotted as a function of vertical effective stress corresponding to the depth of each piezocone test in Figure 7, along with the operative values of the coefficient of consolidation, c_{ref}, from the foundation tests.

Additional piezcone results from other testing programmes in normally consolidated kaolin at UWA but carried out at higher stress levels [4, 20] are also shown on Figure 7, along with the prediction of c_v from c_h using Eq. 1. The ratio c_h/c_v is about 4.5, in good agreement with observations [4, 13]. The stress dependent relationships of c_{ref} and c_h can be approximated by power laws of the form given for c_v in Eq. 1. The relationships shown in Figure 7 indicate the operative coefficient of consolidation, c_{ref}, governing the response of a surface circular foundation under vertical loading, can be taken as $0.6c_h$, as measured from in situ piezocone (or piezoball) tests or $2.6c_v$, as measured through laboratory oedometer tests.

There is a curious feature of this comparison between the two numerical solutions and the experimental data. On the one hand the two numerical solutions coincide if the consolidation coefficient, c_v, from the MCC analysis is based on the virgin (plastic) 1-D compression modulus, M, so

$$c_v = \frac{kM}{\gamma_w} = \frac{k}{\gamma_w}\frac{(1+e)\sigma'_{v0}}{\lambda} \tag{12}$$

On the other hand, the deduced c_{ref} from the experimental data lies closer to the dissipation consolidation coefficient, c_h, than the value, c_v, associated with plastic 1-D compression. This apparent inconsistency needs further investigation, but again emphasises the complexity of assessing suitable value of consolidation coefficient for different boundary value problems.

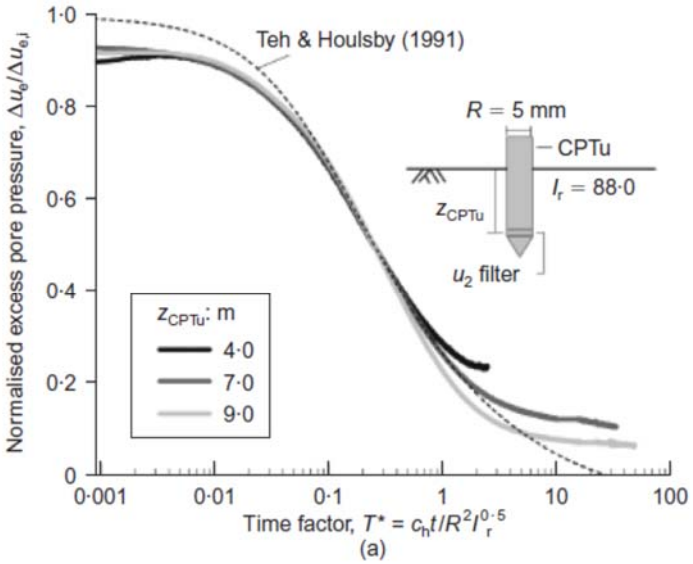


Figure 6: Normalised dissipation time histories from piezocone tests

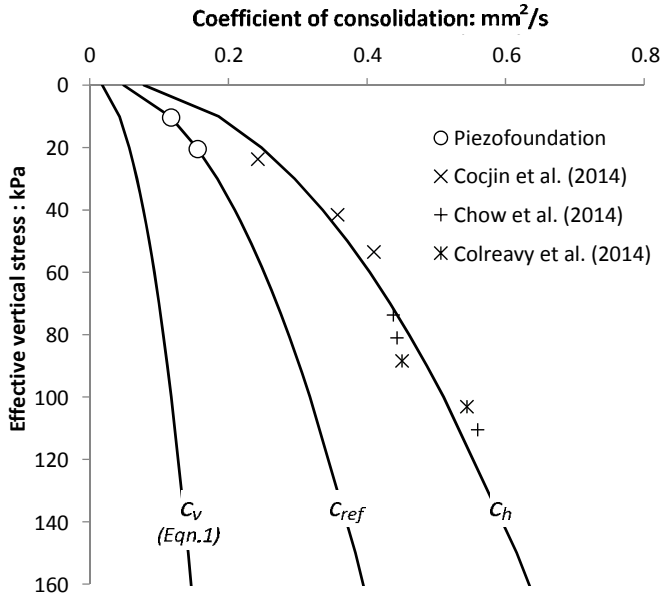
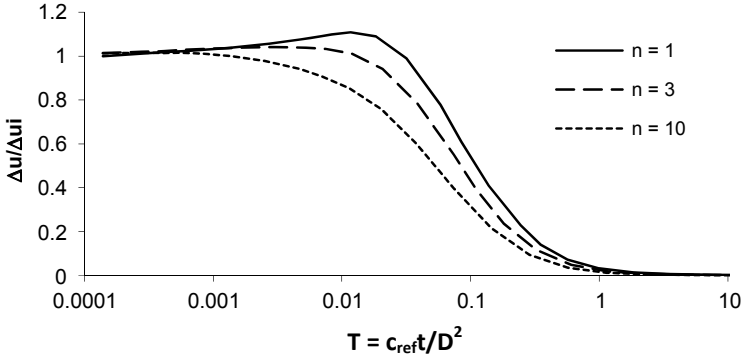
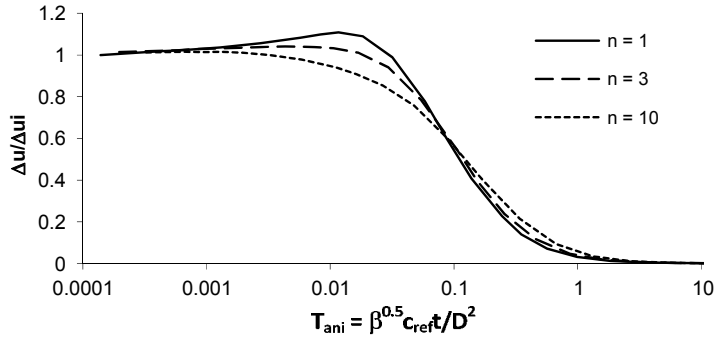


Figure 7: Coefficient of consolidation as a function of vertical effective stress.

The preceding discussion assumed isotropic permeability. Many natural soils exhibit anisotropic permeability, typically with $k_h > k_v$. The effect of anisotropic permeability has been considered previously in solutions of elastic finite element analysis [17]. Figure 8 shows selected results from the study for ratios of horizontal to vertical permeability, i.e. $n = k_h/k_v$ as defined in Section 4, of 1 (isotropic), 3 and 10. It can be seen that the Mandel-Cryer effect becomes less significant with increasing permeability anisotropy and as a result the gradient of the dissipation response becomes less steep over the central part.



(a) Time factor based on k_v



(b) Time factor based on k_h/k_v

Figure 8: Effect of permeability anisotropy on consolidation beneath a surface foundation

In the results shown in Figure 8, the permeability anisotropy ratio, n, was varied by increasing the magnitude of horizontal permeability while keeping the value of vertical permeability constant. The time for consolidation therefore reduces with increasing n due to the overall increase in the 'resultant' coefficient of (vertical and horizontal) consolidation. An adjustment can be applied to the time factor, or more strictly to the assumed coefficient of consolidation for isotropic conditions, to account for the relative increase in the resultant operative value of the coefficient of consolidation. An adjustment of the form suggested for the piezoball interpretation described in Section 4 is applied

$$T_{ani} = \frac{\beta^{0.5} c_{ref} t}{D^2} \tag{13}$$

where β is as given in Eq. 8 and c_{ref} is the operative coefficient of consolidation calculated as a function of the coefficient of vertical permeability. In the case of the shallow foundation consolidation response, an adjustment of $\beta^{0.5}$ was found to provide the best fit with the numerically derived data. Use of Eq. 13 unifies the normalised time histories for the shallow foundation consolidation response, notwithstanding the initial variations in dissipation due to the Mandel-Cryer effect (Figure 8b).

6 DISCUSSION

The time-scale of consolidation processes within soil is determined from the (potentially anisotropic) permeability and the stiffness of the soil skeleton. The two facets result in a coefficient of consolidation, which will vary from one boundary value problem to another and the extent to which the stress-strain response of the soil during consolidation comprises plastic compression or quasi-elastic 'reloading' or swelling. In clays and other fine-grained soils, the permeability of the soil is a relatively strong function of the void ratio, but may be determined through conventional laboratory testing. However, estimation of an appropriate stiffness in order to derive a coefficient of consolidation is more challenging.

Three different boundary value problems have been considered here, all analysed using Modified Cam Clay to represent normally consolidated clay, with parameters suitable for the kaolin clay used to obtain corresponding experimental data. The three problems concern: (1) 1-D consolidation and swelling; (2) pore pressure dissipation around a piezoball penetrometer; and (3) pore pressure dissipation at the centre of a circular surface foundation. For the latter two problems, the effects of anisotropic permeability have also been considered.

Comparing Figures 1 and 6, deduced values of the coefficient of consolidation are bounded by the oedometric values for plastic compression and swelling, which differ by a factor of 6 or 7 for kaolin clay, compared with the λ/κ ratio of 4.7. One of the reasons for the greater factor obtained from numerical analysis is the effect of varying K_0 (and hence ratio of p' to σ'_v) during swelling as the soil becomes increasingly more over consolidated.

The consolidation coefficient relevant for pore pressure dissipation around a penetrometer was found to be a factor of about 5 greater than the oedometric compression value (for a given vertical effective stress). Although superficially comparable with the λ/κ ratio, parametric studies varying λ and κ separately suggest a more complex relationship (Eq. 6).

Data from model tests suggest an operative coefficient of consolidation for pore pressure dissipation from beneath a shallow foundation that lies approximately mid-way between values from laboratory oedometer tests or in situ piezocone (or piezoball) dissipation tests. This appears inconsistent with finite element solutions for soil modelled either as purely elastic or as normally consolidated Modified Cam Clay, where the operative consolidation coefficient from the latter approach appeared to coincide with the oedometric compression value.

Approaches to adjust the derived operational coefficient of consolidation to account for permeability anisotropy have also been suggested. For the piezoball dissipation, the coefficient of consolidation may be adjusted by simple averaging of the permeabilities in the vertical and horizontal directions. By contrast, for dissipation beneath a shallow foundation, the operative permeability appears to vary according to the square root of the average permeability. This seems logical, given that the latter problem is more dominated by vertical flow of pore water, hence the operative coefficient of consolidation is a weaker function of the ratio k_h/k_v.

REFERENCES

[1] Whitlow, R. *Basic soil mechanics, 4th Edition*. Pearson Education, Harlow, UK (2001).
[2] Teh, C.I. and Houlsby, G.T. Analytical study of the cone penetration test in clay. *Géotechnique* (1991) 41(1): 17-34.
[3] Dassault Systèmes *Abaqus analysis user's manual*. Simulia, Providence, RI (2011).

[4] Chow, S.H., O'Loughlin, C.D. and Rolph, M.F. Soil strength estimation and pore pressure dissipation for free-fall piezocone in soft clay. *Géotechnique* (2014) 64(10): 817-827.
[5] Mahmoodzadeh, H., Randolph, M.F. and Wang, D. Numerical simulation of piezocone dissipation test in clays. *Géotechnique* (2014) 64(8): 657-666, 680.
[6] Richardson, M. *Rowe cell test on kaolin clay*. COFS internal report. The University of Western Australia (2007).
[7] Mahmoodzadeh, H., Wang, D. and Randolph, M.F. Interpretation of piezoball dissipation testing in clay. *Géotechnique* (2015) Under review.
[8] Stewart, D.P. *Lateral loading of piled bridge abutments due to embankment construction*. PhD Thesis, University of Western Australia (1992).
[9] Stewart, D.P. and Randolph, M.F. T-Bar penetration testing in soft clay. *J. Geot. Eng. Div.*, ASCE (1994) 120(12): 2230-2235.
[10] Kelleher, P.J. and Randolph, M.F. Seabed geotechnical characterisation with the portable remotely operated drill. *Proc. Int. Symp. on Frontiers in Offshore Geotechnics (ISFOG)*, Perth (2005) 365-371.
[11] Wang, D., Hu, Y. and Randolph, M.F. Three-dimensional large deformation analysis of plate anchors in uniform clay. *J. Geotech. Geoenviron. Engng*., ASCE (2010) 136(2): 355-365.
[12] Wang, D., Randolph, M.F. and White D.J. A dynamic large deformation finite element method based on mesh regeneration. *Comput. Geotech*. (2013) 54: 192-201.
[13] Mahmoodzadeh, H. and Randolph, M.F. Penetrometer testing - the effect of partial consolidation on subsequent dissipation response. *J. Geotech. GeoEnviron. Engng*., ASCE, (2014) 140(6): 04014022.
[14] Sully, J.P., Robertson, P.K., Campanella, R.G. and Woeller, D.J. An approach to evaluation of field CPTU dissipation data in overconsolidated fine-grained soils. *Can. Geotech. J.* (1999) 36(2): 369-381.
[15] Randolph, M.F. and Hope, S. Effect of cone velocity on cone resistance excess pore pressures. *Proc. Int. Symp. On Engineering Practice Performance of Soft Deposits*, Osaka (2004) 147-152.
[16] Cocjin, M.L., Gourvenec, S.M., White, D.J. and Randolph, M.F. Tolerably mobile subsea foundations – observations of performance. *Géotechnique* (2014) 64(11): 895-909.
[17] Gourvenec, S. and Randolph, M. F. Consolidation beneath circular skirted foundations. *Int. J. Geomech*. (2010) 10(1): 22-29.
[18] Feng, X. and Gourvenec, S. Consolidated undrained load-carrying capacities of mudmats under combined loading in six degrees-of-freedom. *Géotechnique* (2015) Under review.
[19] Mayne, P.W. Stress-strain-strength-flow parameters from enhanced in-situ tests. *Proc. Int. Conf. on In-Situ Measurement of Soil Prop. and Case Histories, In-Situ 2001*. Bali, Indonesia (2001) 27-47.
[20] Colreavy, C. *Private communication* (2015).

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis(Eds)

EFFECTIVE STRESS IN UNSATURATED GRANULAR MATERIALS: MICRO-MECHANICAL INSIGHTS

JÉRÔME DURIEZ*, RICHARD WAN*

*Departement of Civil Engineering
University of Calgary
Calgary, AB, T2B 1N4, Canada
e-mail: jerome.duriez@ucalgary.ca

Key words: Unsaturated material, Pendular regime, Micro-mechanics, Effective stress, Discrete Element Method

Abstract. Stress transport in an unsaturated granular material is analytically derived. As such, an effective stress tensor related to the mechanical state governed by the interganular forces within the skeleton is identified together with a capillary stress tensor that accounts for the interactions due to the liquid and gaseous phases. Using a Discrete Element Method for modelling an unsaturated granular material, this capillary stress tensor is computed along different loading paths. Knowing the applied total stress, the effective stress tensor is then readily deduced and it is shown that it describes adequately the strength of the unsaturated granular material along various loading paths for any degree of saturation.

1 INTRODUCTION

Geomaterials encountered in Civil Engineering are multiphasic systems which typically encompass solid, liquid (wetting) and gas (non-wetting) phases. Considering granular geomaterials, the solid phase is discontinuous, consisting of a collection of distinct soil particles. Voids exist between the particles, and we denote the union of the solid phase and these voids as the skeleton. Most of the classical analyses of geomaterials pertain to limiting conditions represented by the void space containing only gas (dry case) or only liquid (saturated case). However, conditions encountered in geotechnical practice involve unsaturated soils, for which these voids include both liquid and gas phases. The presence of the three phases and their mutual interactions enrich their mechanical behaviour, leading for example, to higher strengths depending on the water content. As such, we tackle here this unsaturated case. Various filling configurations of the liquid phase within the void space exist depending on the degree of saturation S_r, *i.e.* the ratio between the water volume, and the volume of the voids. Here, a *pendular regime* is invoked where the

liquid phase consists of distinct menisci between soil particles (Figure 1), corresponding to low values of degree of saturation ($S_r < 20\%$).

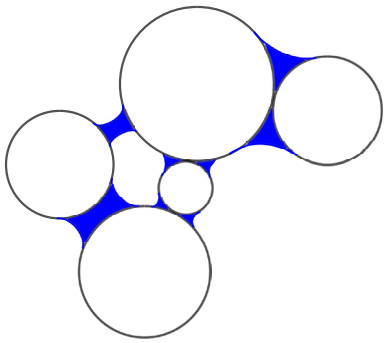


Figure 1: Unsaturated granular material in the pendular regime: water phase forms distinct menisci between contacting and some distant soil particles (idealized as spheres)

We aim here to analyze the mechanical state of unsaturated granular materials, and to investigate how it relates to skeleton behaviour. To this end, analytical derivations of stress transport in the unsaturated material are introduced in Section 2, identifying an *effective stress* related to the skeleton. Then, these newly derived analytical expressions are used in a DEM simulation of an unsaturated granular soil [1] to verify the role of effective stress in governing material strength.

2 FORCE TRANSPORT IN UNSATURATED MEDIA

2.1 Analytical derivations

In [2], the total, or macroscopic, stress for the volume V of unsaturated material depicted in Figure 1 has been computed as an average of the stresses existing in the different phases. The stresses inside the liquid or gaseous phases are readily computed as $\boldsymbol{\sigma} = u_\alpha \, \boldsymbol{\delta} \; \forall \, \vec{x} \in V_\alpha$, with V_α, $\alpha = l, g$, the volume of the phase α, u_α its uniform pressure, and $\boldsymbol{\delta}$ the identity tensor. Note that geomechanics sign convention is used troughout this paper, with compressive stresses and strains being positive.

The stress within the solid phase is computed from local stresses inside each particle, assuming they are in static equilibrium without body forces: $div(\boldsymbol{\sigma}) = \vec{0} \; \forall \, \vec{x} \in V_s^k$, with V_s^k the volume of the solid particle k. Under this assumption, the volume integrals of $\boldsymbol{\sigma}$ are classically computed from the external tractions existing on the boundaries on each particle. Taking into account the solid contact forces, as well as the loadings of the water pressure along all wetted surfaces S_l^k, of the air pressure along all dry surfaces S_g^k, and of the surface tension along all menisci contours Γ_m^k (*i.e.* the intersection of solid, liquid, and gaseous phases along the particle k), the total stress is finally obtained as:

$$\boldsymbol{\sigma} = \frac{1}{V} \sum_{cont.\, s-s} \vec{f} \otimes \vec{l} + u_g\, \boldsymbol{\delta} - (u_g - u_l)\boldsymbol{\chi} - \boldsymbol{B} \tag{1}$$

$$\boldsymbol{\chi} = \frac{V_l}{V}\boldsymbol{\delta} + \frac{1}{V} \sum_k \int_{S_l^k} R^k \vec{n} \otimes \vec{n}\, dS$$

$$\boldsymbol{B} = \frac{1}{V} \sum_k \int_{\Gamma_m^k} \vec{\gamma} \otimes R^k \vec{n}\, dl$$

In equation (1), the vectors $\vec{f}$ and $\vec{l}$ are defined for each contact between two solid grains "1" and "2": $\vec{f}$ is the contact force, as acting on the grain 2, and $\vec{l}$ is the branch vector joining the two particles centers, from 1 to 2. The R^k are the radii of the solid particles considered as spherical whereas $\vec{\gamma}\, dl$ is the incremental surface tension acting on one particle k, at one point of a meniscus contour along this particle.

This analytical model is currently being extended to take into account the interfaces between the different phases, these interfaces having their own energies and stresses [3, 4].

2.2 Discussion

In equation (1), the first term in the expression of $\boldsymbol{\sigma}$ corresponds to the Love-Weber formula [5, 6]. It is related to the skeleton behaviour, as it will be shown in Section 3. Thus, it is denoted as *effective stress*.

In addition to this effective stress, two other stress-like tensors arise from the derivations, neglecting the air pressure $u_g\, \boldsymbol{\delta}$. They correspond to the mechanical interactions due to the presence of both liquid and gaseous phases. For this reason, their sum is denoted as *capillary stress*.

One term of the capillary stress is proportionnal to the matric suction $u_g - u_l$, with a symmetric tensor $\boldsymbol{\chi}$ depending on the distribution of the water phase around the solid particles. This tensor is isotropic, *i.e.* spherical, if and only the above-mentioned distribution is isotropic. The water phase distribution being to a certain extent controlled by the distribution of grain contacts, a deviatoric nature for $\boldsymbol{\chi}$ may arise during a deviatoric loading, due to induced anisotropy [7].

A second term, denoted $\boldsymbol{B}$, describes the mechanical action on the solid grains of the surface tension existing at the liquid-gas interface. In case of a monosized granular material, with an isotropic distribution of liquid phase, this tensor vanishes [2]. In other cases, this tensor is a non-null deviatoric one since, assuming a null contact angle:

$$tr(\boldsymbol{B}) = \frac{1}{V} \sum_k \int_{\Gamma_m^k} R^k \vec{n}.\vec{\gamma_{lg}}\, d\Gamma = 0 \tag{2}$$

According to equation (1), the sum of these two terms, $(u_g - u_l)\boldsymbol{\chi} + \boldsymbol{B}$, is equal to the difference between the effective stress $1/V \sum \vec{f} \otimes \vec{l}$ and the total one $\boldsymbol{\sigma}$. Conversely,

for a known distribution of the water phase in an unsaturated granular material under a total stress $\boldsymbol{\sigma}$, equation (1) can be applied to determine the effective stress related to the granular skeleton: $\boldsymbol{\sigma} + (u_g - u_l)\boldsymbol{\chi} + \boldsymbol{B}$. Note that for the saturated case, equation (1) reduces to Terzaghi's equation.

3 COMPARISON WITH DISCRETE SIMULATIONS

3.1 The DEM model

Simulations of an unsaturated granular material are performed using the code Yade [8], that is based on the Discrete Element Method [9]: the model describes the mechanical state of a collection of discrete elements that correspond to the soil particles. In addition to the contact forces between touching elements, the model includes capillary forces acting on elements. Introducing these capillary forces allows us to simulate the liquid bridges (menisci) of the pendular regime [1].

The contact between particles is governed by frictional elastic-plastic contact laws. Repulsive normal forces act when spheres get closer and they are linearly computed from the relative normal displacement. Contact forces include tangential forces too, computed linearly from the relative tangential displacement up to a plastic threshold obeying Coulomb friction law. This part of the model depends on three micro-parameters: two local stiffnesses, and one microscopic (inter-particle) friction angle. Values used in [1] are retained.

Suction controlled simulations are performed based on an algorithm solving the Laplace equation that describes pendular menisci [1]. From the surface tension, the suction $u_g - u_l$, and features of the numerical packing (radii and distance between elements), a distribution of menisci results directly from the algorithm. Menisci characteristic features are then derived, such as filling angles, volumes, and associated forces acting on grains. Assuming a wetting loading path (increasing S_r), menisci are created between touching particles, and exist between these particles as long as a solution is found for Laplace equation. In particular, menisci still exist when the particles do not touch anymore, leading to distant attractive forces.

Different loading paths may be simulated, using periodic and non-periodic boundary conditions. Periodic and non-periodic triaxial paths ($\sigma_I = \sigma_{yy} > \sigma_{II} = \sigma_{xx} = \sigma_{III} = \sigma_{zz} = cst$) are considered, as well as periodic simple shear tests ($\partial v_x/\partial y = cst$, $\sigma_{yy} = \sigma_{zz} = cst$), see Figure 2. Periodic simulations rely on the homogeneous strain of an unit cell including the sample (see [10] for details). Non-periodic simulations involve rigid frictionless boundaries whose movements impose strain onto the numerical sample.

Periodic and non-periodic simulations were performed on two similar numerical samples. The two samples, one periodic and another non-periodic, include 20,000 spherical particles obeying an uniform size distribution with $D_{max}/D_{min} = 3$ and $D_{50} = 45.3 \pm 0.5\,\mu m$. The porosity of the two samples under 1 kPa isotropic pressure is 0.36 ± 0.01, leading to a dilatant behaviour for confining pressures in the range [1; 20 kPa].

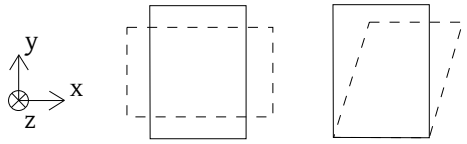


Figure 2: Triaxial (axisymmetric) compression: left, and simple shear: right

3.2 Discrete results for the dry case

The failure of the granular material is first investigated in the dry case, without introducing any capillary forces in the model. Limit stress states for the two samples along various loading paths are depicted in Figure 3 according to the first and second invariants of $\boldsymbol{\sigma}$: the mean stress $p = (\sigma_I + \sigma_{II} + \sigma_{III})/3$ and the deviatoric stress $q = \sqrt{3/2} * ||\boldsymbol{\sigma} - p\,\mathbf{1}||$ ($q = \sigma_I - \sigma_{III}$ for the triaxial paths). Since the plastic limit criterion of granular materials also depends on the third invariant, deviatoric stresses for simple shear paths are corrected according to the current Lode angle, θ, values, assuming a Lade criterion [11]. As such, direct comparison is possible between the various loading paths, despite different θ values ($\theta = 0°$ for triaxial compression, while, here, $\theta \in [22°; 30°]$ for simple shear). As is evidenced in Figure 3, all stress limit states obey a common Mohr-Coulomb (MC) criterion, with a macroscopic friction angle $\phi = 29°$.

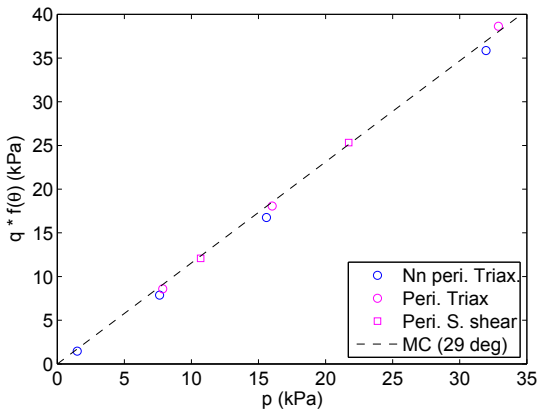


Figure 3: Stress limit states according to the mean and deviatoric stresses, p and q. Deviatoric stresses for simple shear tests are corrected to be directly compared to values of triaxial tests, despite the difference in Lode angle θ

3.3 Discrete results for the unsaturated case

The model is then applied to the unsaturated conditions. Six non-periodic triaxial loadings are simulated for different confining pressures, in {5;10} kPa, and suction values, in {10;50;100;300} kPa, see Table 1. Associated degrees of saturation values are low (below 10%) ensuring a pendular regime.

Table 1: Initial degrees of saturation (%) for the simulated unsaturated triaxial loadings

		$u_g - u_l$ (kPa)			
		10	50	100	300
σ_{xx} (kPa)	5	-	-	0.32	0.045
	10	8.7	1.0	0.32	0.046

Classically, limit stress states in unsaturated conditions correspond to higher deviatoric stresses, not obeying anymore to the Mohr-Coulomb determined from the dry case (Figure 4).

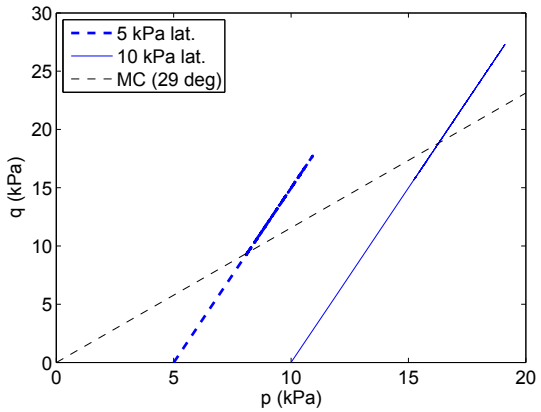


Figure 4: Stress paths for two unsaturated triaxial loadings with 100 kPa suction, and different lateral pressures

Indeed, due to the presence of liquid and gaseous phases, the effective stresses related to the solid phase are now different from the total stresses (Figure 5). All stress tensors being axisymmetric around $\vec{y}$ axis (see Figure 2) during triaxial loadings, it is sufficient to consider the mean (one third of the trace) and deviatoric ($yy - xx$) components of each tensor. Both water and gaseous phases induce mainly a change in mean stresses. The deviatoric component of the capillary stress tensor is here small, so that there is not much of a difference between deviatoric effective and total stresses. This is related to the initial isotropy of the DEM packing (see [2] for a study of inherent anisotropic packings), which is consistent with initial null values of the deviatoric stress. However, because of the induced anisotropy at the solid contact scale, a deviatoric feature for the capillary stress tensor is induced by the loading (see also the Figures 7 and 8) and this deviatoric feature increases with suction (Figure 6).

Among the two terms forming the capillary stress tensor, $(u_g - u_l)\boldsymbol{\chi}$ is rather predominant with respect to $\boldsymbol{B}$ (compare Figures 7 and 8). However, the relative magnitudes of these two terms depend on the suction. Indeed, the deviatoric component of $(u_g - u_l)\boldsymbol{\chi}$

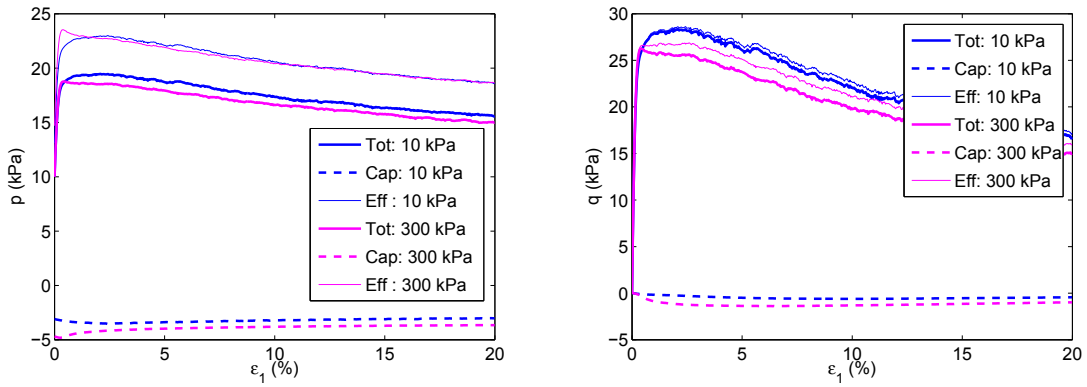


Figure 5: Mean (left) and deviatoric (right) components of the different stress tensors during two unsaturated triaxial loadings, under 10 kPa lateral pressure, and two different suctions. Capillary stress tensor is $(u_g - u_l)\boldsymbol{\chi} + \boldsymbol{B}$, and effective stress tensor is $\boldsymbol{\sigma} - (u_g - u_l)\boldsymbol{\chi} - \boldsymbol{B}$

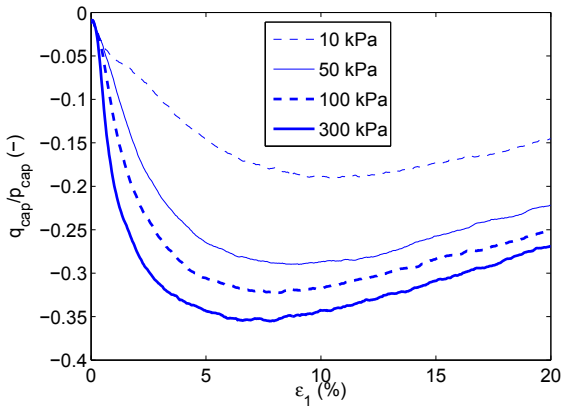


Figure 6: Deviatoric feature of the capillary stress tensor $(u_g - u_l)\boldsymbol{\chi} + \boldsymbol{B}$ for triaxial loadings under 10 kPa lateral pressure and different suctions

increases considerably from initial small values with suction (Figure 7) whereas on the other hand the deviatoric component of $\boldsymbol{B}$ (*i.e.* its norm) decreases significantly from large values with suction (Figure 8). This is because, for increasing suction, both degree of saturation and filling angles decrease in the granular material, leading to lower menisci contour, and lower values for the components of $\boldsymbol{B}$. In the end, for high suction values, $(u_g - u_l)\boldsymbol{\chi}$ is the main contributor to the capillary deviatoric stresses. On the other hand, for low suctions, deviatoric capillary stresses arise mainly thanks to $\boldsymbol{B}$.

Periodic unsaturated simple shear simulations, with $\sigma_{xx} = \sigma_{yy} = 10$ kPa, are also performed for different suction values. Corresponding degrees of saturation (for the initial states) are detailed in Table 2. The same trends concerning the relative magnitudes of these two terms are observed during simple shear than triaxial conditions (Figures 9 and 10).

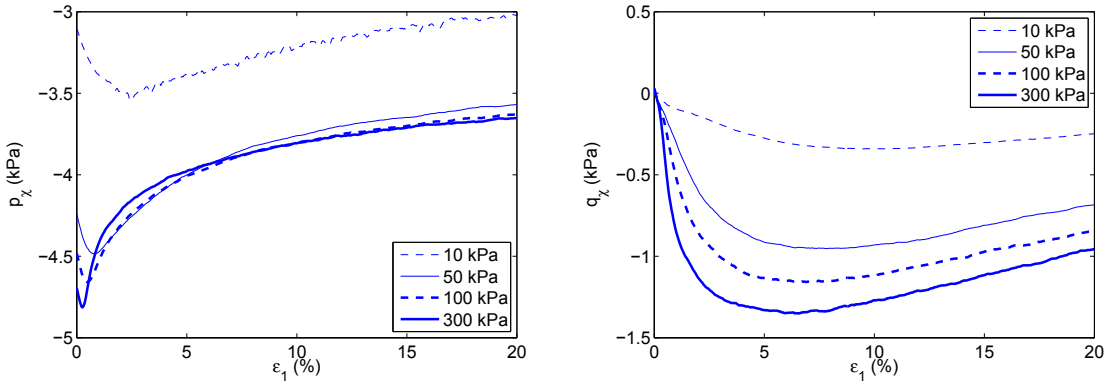


Figure 7: Mean (left) and deviatoric (right) components of $(u_g - u_l)\boldsymbol{\chi}$ for triaxial loadings under 10 kPa lateral pressure and different suctions

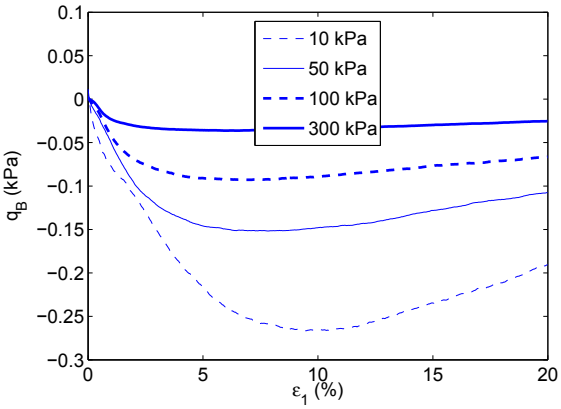


Figure 8: Deviatoric component of $\boldsymbol{B}$ for triaxial loadings under 10 kPa lateral pressure and different suctions (remind that $tr(\boldsymbol{B}) = 0$)

Table 2: Initial saturation ratios (%) for the unsaturated simple shears

$u_g - u_l$ (kPa)			
10	50	100	300
10.1	1.2	0.37	0.054

3.4 A comprehensive plastic limit criterion for dry and unsaturated cases

For the various loading paths simulated in unsaturated conditions, effective stresses $\boldsymbol{\sigma'}$ related to the skeleton are deduced from the total and capillary stresses. Note that the DEM model offers a direct computation of effective stresses [1], which is not done here. As in the previous Section 3.2, limit stress states are determined and plotted in the (p', q')

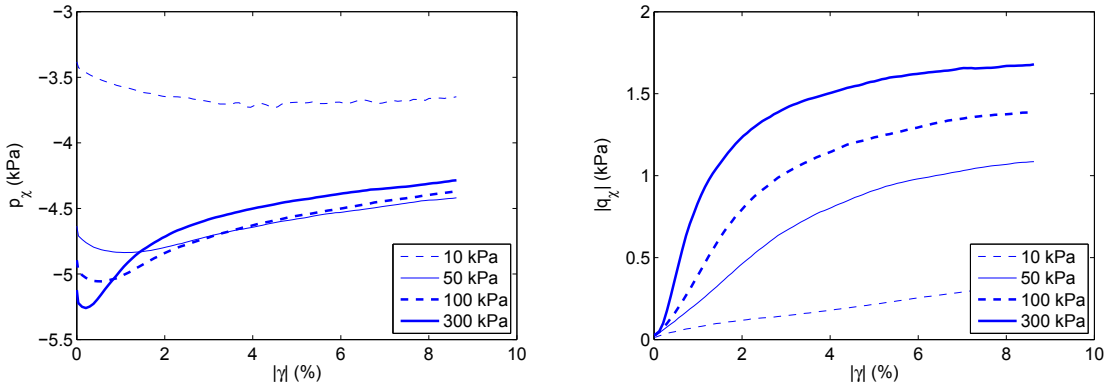


Figure 9: Mean (left) and deviatoric (right) components of $(u_g - u_l)\boldsymbol{\chi}$ for simple shear loadings with different suctions

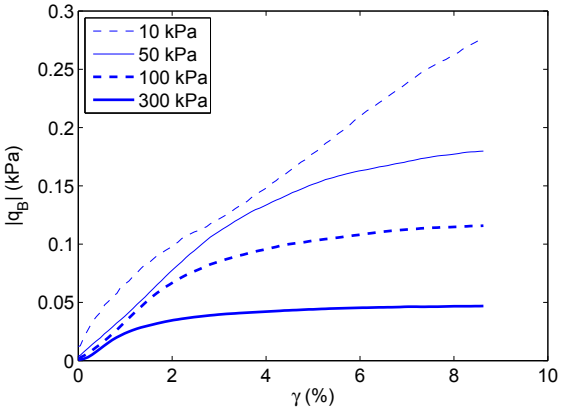


Figure 10: Deviatoric component of $\boldsymbol{B}$ for simple shear loadings with different suctions

plane (Figure 11). The Mohr-Coulomb criterion determined from the dry case in Section 3.2 is obeyed with an acceptable dispersion. This validates the relevancy of the effective stress derived in equation (1) to express the strength of a granular material, whatever its saturation ratio.

4 Conclusion

Analytical derivations of stress transport in an unsaturated granular media identified an effective stress related to the mechanical state governed by intergranular forces within skeleton, and a capillary stress arising from the mechanical interactions due to the liquid and gaseous phases. This capillary stress depends on the packing and on the liquid phase distribution inside the material. This stress is non-spherical for anisotropic liquid distribution, as it has been evidenced in case of loading-induced anisotropy. This deviatoric feature of the capillary stress increases with matric suction.

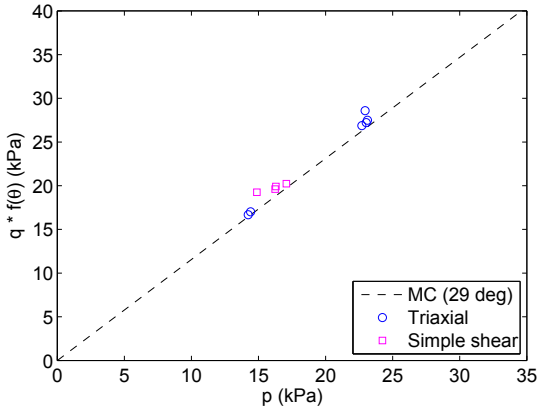


Figure 11: Limit stress states for unsaturated cases along different loading paths, compared with the dry Mohr-Coulomb criterion deduced from the dry case

In fact, two tensorial terms accounting for the capillary actions can be distinguished, with one deviatoric in case of a null contact angle. Relative magnitudes of these two terms depend on the suction.

Determining such capillary stress allows us to identify the effective stress tensor. It is shown that this effective stress tensor is an adequate variable to express the strength of an unsaturated granular material, along various loading paths irrespective of the degree of saturation.

REFERENCES

[1] L. Scholtès, B. Chareyre, F. Nicot, and F. Darve. Micromechanics of granular materials with capillary effects. *International Journal of Engineering Science*, 47(1):64 – 75, 2009.

[2] R. Wan, S. Khosravani, and M. Pouragha. Micromechanical analysis of force transport in wet granular soils. *Vadose Zone Journal*, 13(5):1–12, 2014.

[3] Norman R. Morrow. Physics and thermodynamics of capillary action in porous media. *Industrial & Engineering Chemistry*, 62(6):32–56, 1970.

[4] Xavier Chateau and Luc Dormieux. Micromechanics of saturated and unsaturated porous media. *International Journal for Numerical and Analytical Methods in Geomechanics*, 26(8):831–844, 2002.

[5] A.E.H. Love. *A treatise on the mathematical theory of elasticity.* Cambridge University Press, Cambridge, 1927.

[6] J. Weber. Recherches concernant les contraintes intergranulaires dans les milieux pulvérulents. *Bulletin de liaison des Ponts et Chaussées*, 20:1–20, 1966.

[7] Masanobu Oda, Siavouche Nemat-Nasser, and Junichi Konishi. Stress-induced anisotropy in granular masses. *Soils and Foundations*, 25(3):85–97, 1985.

[8] V. Smilauer, E. Catalano, B. Chareyre, S. Dorofeenko, J. Duriez, A. Gladky, J. Kozicki, C. Modenese, L. Scholtès, L. Sibille, J. Stránský, and K. Thoeni. *Yade Documentation*. The Yade Project, 1st edition, 2010. http://yade-dem.org/doc/.

[9] P.A. Cundall and O.D.L. Strack. A discrete numerical model for granular assemblies. *Géotechnique*, 29:47–65, 1979.

[10] F. Radjai and C. Voivret. Periodic boundary conditions. In F. Radjai and F. Dubois, editors, *Discrete-element Modeling of Granular Materials*. ISTE-Wiley, 2011.

[11] Poul V. Lade and James M. Duncan. Elasto-plastic stress-strain theory for cohesionless soil. *Journal of the Geotechnical Engineering Division*, 101(10):1037–1053, 1975.

SEISMIC RESPONSE ANALYSIS BY SUBLOADING SURFACE MODEL

KENICHIRO MIYASHITA*, KOICHI HASHIGUCHI†, AND SHIGERU SATO*

* Pacific Consultants Corporation (PCKK)
7-5,Sekido 1-chome, Tama-shi,Tokyo,Japan
Email: kenichirou.miyashita@os.pacific.co.jp - Web page: http://www.pacific.co.jp

† Dept. Dependent and Optimum Design, Joining and Welding Research Institute
Mihogaoka 11-1, Ibaragi-shi, Osaka 567-0047, Japan
Email: khashi@kyudai.jp

Key words: Seismic Response, liquefaction, Elastoplasticity.

1 INTRODUCTION

A lot of disaster by liquefaction have been reported in area along the shore of Japan. In particular, liquefaction has occurred in the wide area in the Great East Japan Earthquake of 2011. Various approaches for the liquefaction analysis have been proposed up to present. Among these approaches, the subloading surface model is formulated in the framework of the plasticity model and thus it is expected to provide a highly pertinent simulation of cyclic loading behaviour of materials. Further, the explicit constitutive equation of soils has been formulated to describe the cyclic loading behaviour with the cyclic mobility [1]. In this study, the validity of the liquefaction analysis by the subloading surface model is examined by comparing the simulation by the subloading surface model with the actual record for the acceleration wave in the ground surface to the input of the actual data of the acceleration wave in the soil ground base. The actual data used in the simulation was recorded in the Kushiro earthquake in 1993. [2]

2 SUBLOADING SURFACE MODEL

Subloading surface model is the elastoplasticity model that considering plastic strain rate induced by the rate of stress inside the yield surface. Its basic concept and equations for subloading surface model is indicated in the following.

2.1 Normal-yield and subloading surfaces

The subloading surface is introduced which always passes through current stress point and has similar shape to the yield surface in subloading surface model, which the yield surface is renamed the *normal-yield surface*. In subloading surface model, the plastic strain rate generates by a change of not only normal-yield surface but also R, which is the ratio of the size of the subloading surface to that of normal-yield surface. R is called as the *normal-yield ratio.* The state R=0 corresponds to the null stress state in which a purely elastic deformation behavior

occurs, the state $0<R<1$ to the subyield state, and the state $R=1$ to the nomal-yield state in which the stress lies on the normal-yield surface for which the plastic strain rate has been formulated in conventional plasticity.The yield surface and the subloading surface are described as Eq.(1) and Eq.(2).

$$f(\boldsymbol{\sigma},\boldsymbol{\beta}) = F(H) \tag{1}$$
$$f(\bar{\boldsymbol{\sigma}},\boldsymbol{\beta}) = RF(H) \tag{2}$$

where

$$\bar{\boldsymbol{\sigma}} \equiv \boldsymbol{\sigma} - \bar{\boldsymbol{\alpha}} \tag{3}$$
$$\bar{\boldsymbol{\alpha}} = (1-R)\boldsymbol{c} \tag{4}$$

$\boldsymbol{\sigma}$ is the Cauchy stress, $\boldsymbol{\beta}$ is the roationl-hardening variable, H is the isotropic hardening variable. $\boldsymbol{c}$ is the elastic-core, i.e the center of similarity of subloading surface to the normal-yield surface. $\bar{\boldsymbol{\alpha}}$ is the similar point in the subloading surface to origin. They are illustrated on the (p,q) plane in Fig.1. where p is mean effective pressure, q is the stress difference between the vertical and horizontal directions.

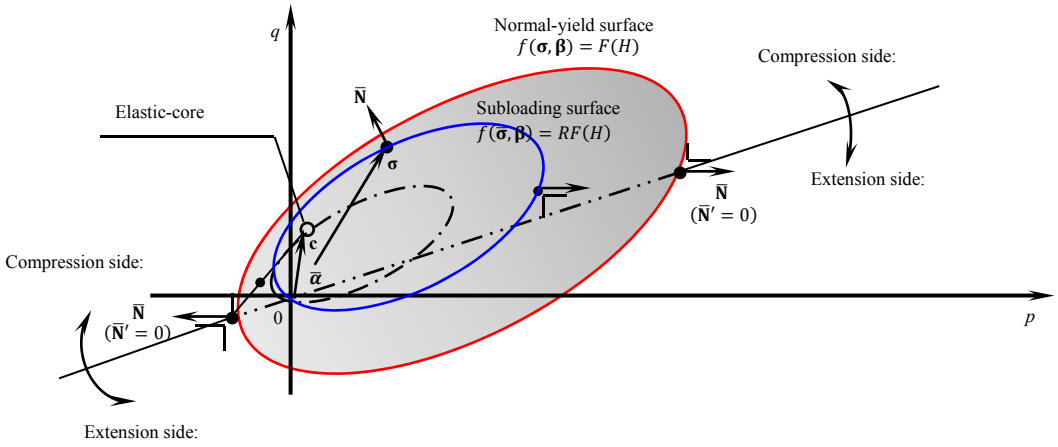


Figure 1: Rotated normal-yield, subloading and similarity-center surfaces in the(p,q)plane

The evolution rule of R is given as Eq.(5), which is based on assumption that as the subloading surface approaches the normal-yield surface, R does not increase gradually , and when the subloading surface corresponds to normal-yield surface, the rate of normal-yield ratio is 1.0.

$$dR = U(R)\|\boldsymbol{D}^{\boldsymbol{p}}\| = \lambda U(R) \tag{5}$$

Where U is monotonically deceasing function of R (See Eq.(6)), $\boldsymbol{D}^{\boldsymbol{p}}$ is plastic strain rate, λ is the magnitude of plastic strain rate. By this modeling, the subloading surface model can express the natural stress-strain curve that the plastic strain rate increase smoothly as stress approaches the normal-yield surface.

$$\left.\begin{aligned} U(R) &= u\cot(\frac{\pi}{2}R) \\ U &= +\infty \text{ for } R = 0 \\ U &= 0 \text{ for } R = 1 \end{aligned}\right\} \tag{6}$$

2.2 The strain rate by the subloading surfaces model

The strain rate by the subloading surface is indicated in the following. The partial derivatives of the Eq.(2) is shown below.

$$\frac{\partial f(\bar{\boldsymbol{\sigma}},\boldsymbol{\beta})}{\partial \bar{\boldsymbol{\sigma}}} : d\boldsymbol{\sigma} - \frac{\partial f(\bar{\boldsymbol{\sigma}},\boldsymbol{\beta})}{\partial \bar{\boldsymbol{\sigma}}} : d\bar{\boldsymbol{\alpha}} + \frac{\partial f(\bar{\boldsymbol{\sigma}},\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} : d\boldsymbol{\beta} = dRF + RdF \tag{7}$$

Eq.(7) is rewritten as below

$$\bar{\boldsymbol{N}} : d\boldsymbol{\sigma} - \bar{\boldsymbol{N}} : \left\{ d\bar{\boldsymbol{\alpha}} + \frac{dF}{F}\bar{\boldsymbol{\sigma}} + \frac{dR}{R}\bar{\boldsymbol{\sigma}} - \frac{1}{RF}\left(\frac{\partial f(\bar{\boldsymbol{\sigma}},\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} : d\boldsymbol{\beta}\right)\bar{\boldsymbol{\sigma}} \right\} = 0 \tag{8}$$

Where

$$\bar{\boldsymbol{N}} \equiv \frac{\partial f(\bar{\boldsymbol{\sigma}},\boldsymbol{\beta})}{\partial \boldsymbol{\sigma}} \Big/ \left\| \frac{\partial f(\bar{\boldsymbol{\sigma}},\boldsymbol{\beta})}{\partial \boldsymbol{\sigma}} \right\| \tag{9}$$

Further, assume the associated flow rule

$$\boldsymbol{D}^{\boldsymbol{p}} = \lambda \bar{\boldsymbol{N}} \tag{10}$$

Let the translation rule of elastic-core and the evolution rule of rotational hardening be given as [1]

$$\mathrm{d}\boldsymbol{c} = \bar{c}\lambda\left(\frac{\bar{\boldsymbol{\sigma}}}{R} - \frac{\boldsymbol{c}}{\chi}\right) + \left(\frac{dF}{F} - \frac{1}{\chi F}\frac{\partial f(\boldsymbol{c},\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} : d\boldsymbol{\beta}\right)\boldsymbol{c} \tag{11}$$

$$\mathrm{d}\boldsymbol{\beta} \equiv \lambda \boldsymbol{b} \tag{12}$$

Where χ , $\bar{c}$ are material parameter. Substituting Eq.(10),Eq.(11),Eq.(12) into Eq.(8) yields

$$\bar{\boldsymbol{N}} : d\boldsymbol{\sigma} - \lambda M^{p} = 0 \tag{13}$$

where

$$M^{p} \equiv \bar{\boldsymbol{N}} : \{ \frac{dF}{dH}\frac{h}{F}\bar{\boldsymbol{\sigma}} + \frac{U(R)}{R}(\boldsymbol{\sigma} - \boldsymbol{c}) + \bar{c}(1-R)\left(\frac{\bar{\boldsymbol{\sigma}}}{R} - \frac{\boldsymbol{c}}{\chi}\right) - \frac{1}{RF}\left(\frac{\partial f(\bar{\sigma},\beta)}{\partial \boldsymbol{\beta}} : \boldsymbol{b}\right)\bar{\boldsymbol{\sigma}} - \frac{1-R}{\chi F}\left(\frac{\partial f(\boldsymbol{c},\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} : \boldsymbol{b}\right)\boldsymbol{c} \} \tag{14}$$

$$h \equiv \frac{dH}{\lambda} \tag{15}$$

Then the magnitude of plastic strain rate is given by

$$\lambda = \frac{\bar{\boldsymbol{N}} : d\boldsymbol{\sigma}}{M^{p}} \tag{16}$$

The strain rate is given by

$$\boldsymbol{D} = \boldsymbol{E}^{-1}\colon d\boldsymbol{\sigma} + \lambda\overline{\boldsymbol{N}} \tag{17}$$

Where $\boldsymbol{D}$ is the strain rate, $\boldsymbol{E}$ is young's modulus.

2.3 Description of Cyclic mobility by subloading surface model

Cyclic mobility occurring in the liquefaction in sands is a peculiar phenomenon exhibiting a butterfly-shaped stress loops and a S-shaped stress-strain loops under undrained cyclic loading. In elastoplasticity, Considering the plastic potential different from yield surface, the method that the elastic shear modulus *G* is increased by recovery of the effective pressure is often used as the method of expressing cyclic mobility. However, the physical explanation of the dilatancy is indefinite, because this method considers the 2 kinds of the dilatancy.

The subloading surface model expresses cyclic mobility by formulating the rate of isotropic hardening/softening variable *H* that influence of the deviatoric plastic strain rate is incorporated.

$$h = -tr\overline{N} + \mu_d\left\|\overline{N'}\right\| \frac{\{\frac{\chi_d}{p+\zeta F}\}^{a-1}}{\{\frac{\chi_d}{p+\zeta F}\}^{a-1+b}} \tag{18}$$

$$\chi_d \equiv \frac{\|\sigma\prime\|}{M_d} \tag{19}$$

$$M_d \equiv \frac{14\sqrt{6}sin\phi_d}{(3-sin\phi_d)(8+cos3\theta_d)} \tag{20}$$

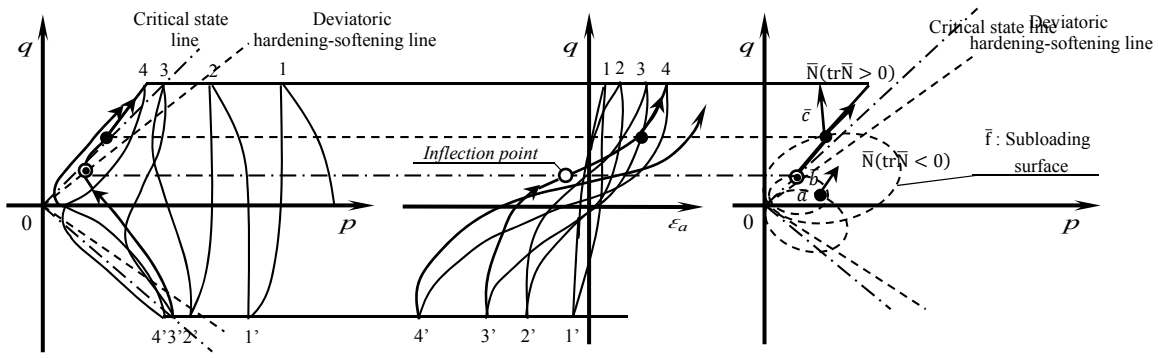
$$cos3\theta_d \equiv \sqrt{6}tr\boldsymbol{\tau}^3 \tag{21}$$

$$\boldsymbol{\tau} \equiv \frac{\boldsymbol{\sigma'}}{\|\boldsymbol{\sigma'}\|} \tag{22}$$

where μ_d,ϕ_d,a,b and ζ are material constants. The hardening and the softening are induced outside and inside, respectively, conical surface $\|\boldsymbol{\sigma'}\| = M_d(p+\zeta F)$. The deviatoric hardening rate depends nonlinearly on the modified stress ratio $\chi_d/(p+\zeta F)$.

By this method, the subloading surface model is able to express cyclic mobility without considering 2 kinds of the dilatancy. The method of expressing cyclic mobility in detail is indicated in the following.A butterfly-shaped stress-strain loops in cyclic mobility is illustrated in Fig.2(b). This phenomenon can be simulated by subloading surface model as fellows: The deviatoric stress varies under a high effective pressure in the initial stage of cyclic loading so that the plastic volume contraction is induced leading to a denser arrangement of sand particles. To keep the volume constant, elastic volume expansion is induced by the decrease of effective confining pressure under undrained condition. After the effective pressure decreases as represented at the point $\bar{a}$ in Fig.2(c), the deviatoric stress increases over the critical state line by the deviatoric hardening for $\|\boldsymbol{\sigma'}\|/p > tan\phi_d$ and reaches the dense state(tr$\overline{N} > 0$) causing the plastic volume expansion so that the effective pressure increases responding to the elastic volume contraction in order to keep the volume constant.As the dense state with the deviatoric hardening because of $\|\boldsymbol{\sigma'}\|/p > M_c > M_d$ proceeds, the effective stress rise up at almost constant effective stress ratio as represented at the point $\bar{c}$ in Fig.2(c). Consequently, the effective stress path goes up straightly from the origin in the (p,q) plane. The nomal-yield

surface expands markedly so that the strain rate decreases gradually in this process. Then, the $q-\varepsilon_a$ curve gets warped to the upper as shown in this Fig.2(b). where ε_a is the vertical strain.



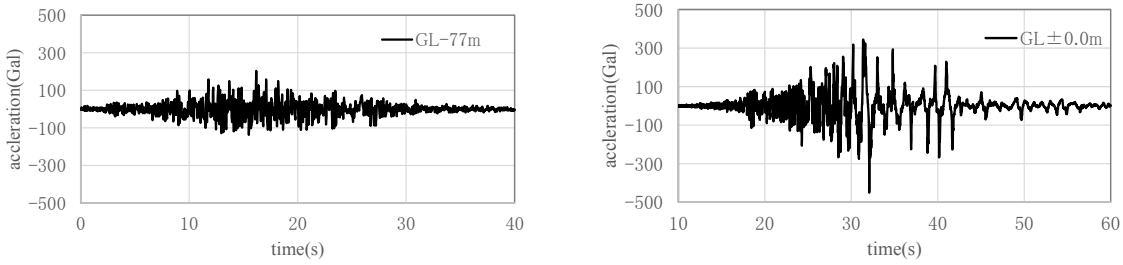
(a)Stess path (b)Stress-strain relation (c) Variations of normal-yield and subloading surfaces

Figure 2: Phenomenon in cyclic mobility

3 THE SIMULATION OF THE KUSHIRO EARTHQUAKE IN 1993

3.1 The acceleration histories of observation

For the Kushiro earthquake, the observation acceleration histories are obtained at ground level GL±0.0 and underground GL -77m at the observation point KUSHIRO G near-shore. KUSHIRO G is a strong-motion observation point managed by a Japanese incorporate administrative agency called the Port and Airport Research Institute, which discloses the observation results on the WEB. Fig. 3 indicates the disclosed wave form observed at KUSHIRO G. The greater acceleration of GL ± 0.0 than of GL -77m indicates that the acceleration was amplified by the resonance with the ground. Also, the wider time interval between peaks of GL±0.0 than of GL -77m, and furthermore, the spike form acceleration wave form after 30 seconds are considered it is caused by the cyclic mobility.



(a)GL-77.0m (a)GL0.0m

Figure 3: Observation acceleration histories

3.2 Ground condition and the material parameters

The speed of s-wave propagation through medium by PS logging and unit weight volume of the ground at observation point KUSHIRO G are also disclosed. The following indicates these values and the material parameters set based on these values.

Table 1: Ground condition and the material parameters

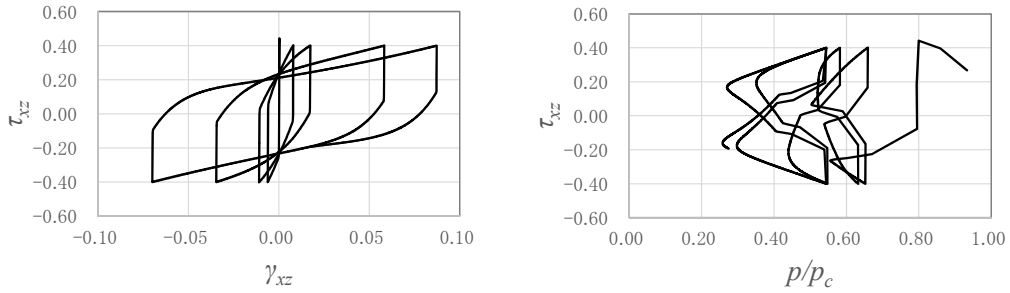
level			layer		ρ (t/m^3)	Vs (m/s)	μ_d	φ_d	a	b	u	c	F_0
0	～	-2	layer1	sand	1.8	146	1	28	3	20	10	20	353
-2	～	-5	layer2	sand	1.8	146	1	28	3	20	10	20	933
-5	～	-13	layer3	sand	1.95	355	1	28	3	20	10	20	1882
-13	～	-24	layer4	sand	1.9	357	1	28	3	20	10	20	3542
-24	～	-37	layer5	sand	2	324	1	28	3	20	10	20	5718
-37	～	-52	layer6	sand	2	324	1	28	3	20	10	20	8383
-52	～	-64	layer7	sand	2	337	1	28	3	20	10	20	10952
-64	～	-77	layer8	sand	2	337	1	28	3	20	10	20	13331

Where F_0 is the parameter of the size of normal-yield surface, V_s is s-wave propagation through medium.

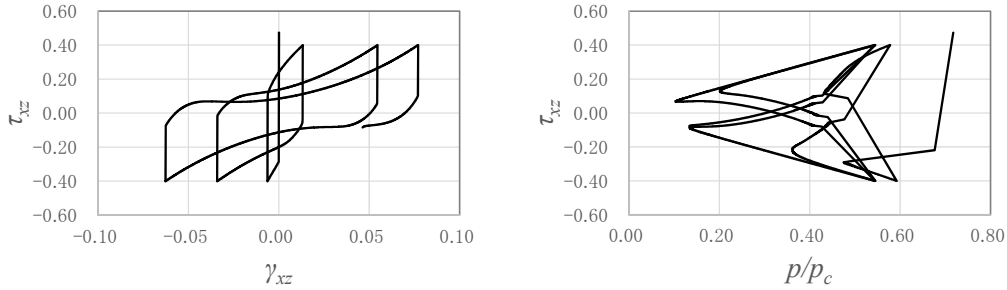
3.3 Simulations of a simple shear test

Prior to run the seismic simulations, simulations of a simple shear test in horizontal direction considering layer2 and 3 shown in the Table 1 were practiced to confirm the reproducibility of the cyclic mobility by the subloading surface model. The stress-strain relationship and stress path obtained by the simulations are shown in Fig. 4. Where, p_c is an initial mean effective pressure, τ_{xz} is a shear stress in horizontal direction, γ_{xz} is a shear strain in horizontal direction. *FLAC3D* based on the explicit dynamic relaxation method is adopted in this simulation.

Fig.4 indicates subloading surface model can well reproduce the cyclic mobility as the stress-strain loops are butterfly-shaped. Also, the stress path well expresses the tracks of mean effective pressures that gradually decrease by the deviatoric stress and then increase linearly in consequence of dilatancy.



(a)layer2



(b)layer3

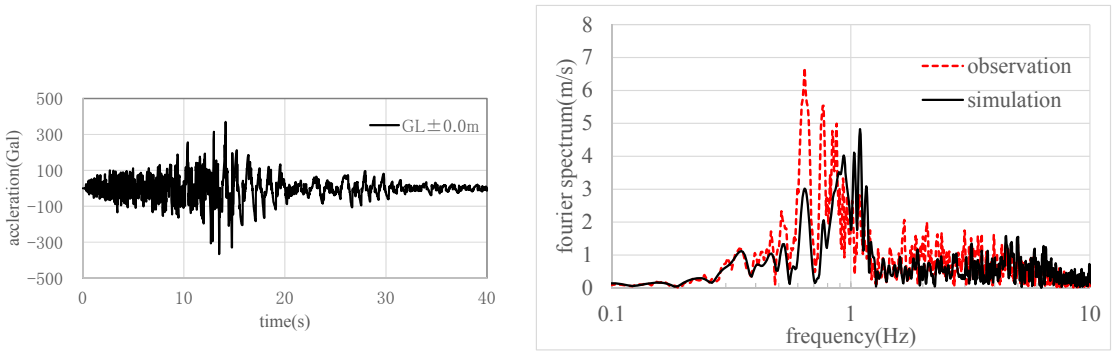
Figure 4: Simulation of a simple shear test in horizontal direction

3.4 Simulation results

A simulation model was set up considering the horizontal ground between GL±0.0 to GL-77m and run giving the observed wave form at GL-77m.The acceleration history for surface of layer is shown in Fig 5(a). The fourier spectrum of the observed wave form and simulation result is shown in Fig 5(b).

The maximum acceleration obtained by the simulation almost fits the observed data, and the trend that the frequency is prolonged after the occurence of the maximum acceleration is well reproduced. The simulated model shows larger fourier spectrum at higher frequencies compared to the observed wave data, which indicates the model evaluates the ground somewhat harder than the actual situation.

The stress-strain relationship at Layer2 is shown in Fig 6. The cyclic mobility of layer2 is not as explicit as the simple sear test. The observed wave data shows elasticity before the occurence of maximum acceleration, and this study assumed large size of yeilding surface to reproduce the behaviour. It is considered, as the result, the plastic strain became smaller and it disturbed the non-occuremce of the cyclic mobility.



(a)acceleration history for surface (b)fourier spectrum

Figure 5:Results of Simulation

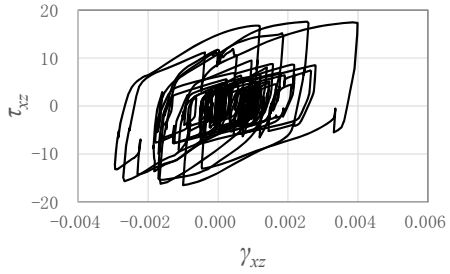


Figure 6:Stress-Strain relationship

4 CONCLUSIONS

- For the simulations of a simple shear test in horizontal direction, the subloading surface model was able to express the cyclic mobility.
- Cyclic mobility of the simulated model was not explicit, but the maximum acceleration obtained by the simulation almost matched the observed data. Also, the trend that the frequency is prolonged after the occurrence of the maximum acceleration was well reproduced.

ACNOWLEDGEMENT

The seismic waves adopted to the study was sourced from an observation point KUSHIRO G installed by a Japanese incorporate administrative agency called the Port and Airport Research Institute.

REFERENCES

[1] K. Hashiguchi, “Elastoplasticity Theory”, Second Edition, Lecture Notes in Applied and Computational Mechanics, Springer (2013).

[2] Strong-motion Seismograph Networks : http:// www. kyoshin. bosai.go.jp /kyoshin/ docs/ kyoshin_index.html

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis (Eds)

THERMO-PORO-MECHANICAL COUPLED PROCESSES DURING THERMAL PRESSURIZATION AROUND NUCLEAR WASTE REPOSITORY

M N. VU[*], D. SEYEDI[*] AND G. ARMAND[†]

[*] Andra, R&D Division
1-7 rue Jean Monnet, 92290 Châtenay-Malabry, France
e-mail: minh-ngoc.vu@andra.fr, darius.seyedi@andra.fr

[†] Andra, R&D Division, Meuse/Haute-Marne Underground Research Laboratory
RD960, 55290 Bure, France
email: gilles.armand@andra.fr

Key words: THM modeling, thermal pressurization, sensibility, nuclear waste reposity, COx.

Abstract. This paper investigates the thermo-hydro-mechanical behavior of Callovo-Oxfordian claystone, a potential host formation for prospective nuclear waste disposal in France. Thermal pore pressure appears in low permeability soils and rocks due to the difference between the thermal expansion coefficients of water and the argillaceous skeleton, as well as the low permeability of the media and the its relative rigidity, which prevent dissipation of the fluid pressure. Coupled thermo-hydro-mechanical numerical analyses have been carried out to enhance the understanding of the Callovo-Oxfordian claystone behavior subjected to heat emitted from radioactive waste that diffuses through the near-field rock to the far-field. In this view, the "thermal pressurization coefficient", defined as the increase of pore pressure due to 1°C increase of temperature, was calculated. This coefficient depends on the nature of the rock, i.e the thermo-poro-mechanical parameters such permeability, Biot's coefficient, rigidity, thermal conductivity as well as their anisotropies. Finally, the effect of parameters' variability on the thermal pressurization coefficient is discussed through a sensibility analysis.

1 INTRODUCTION

Clays and claystones are studied as potential host formation for the disposal of exothermal high activity radioactive waste at great depth in various countries such as the Callovo-Oxfordian claystone (COx) in France, Boom Clay in Belgium, and Opalinus Clay in Switzeland.

During the exploitation stage of deep radioactive geological disposal, exothermic reactions of waste provoke thermal perturbation within the repository environment. Heat emitted from radioactive waste diffuses through the near-field rock to the far-field. Therefore, the host rock is subjected to an increase of temperature (up to about 80°C). The temperature rise in a low permeability porous medium such as COx claystone, generates pore pressure increase essentially due to the difference between the thermal expansion coefficients of water ($\sim 10^{-4} K^{-}$

[1]) and the one of the argillaceous rock skeleton (~1.28 $10^{-5}K^{-1}$) ([1],[2]). This phenomenon has clearly been observed in undrained laboratory tests conducted on COx ([3]), Boom clay ([4]) and Opalinus clay ([5]). The physical phenomenon is more complicated in reality, where the pore pressure occurs due to following principle factors: difference of thermal dilatation coefficient between water and solid skeleton, fluid dissipation controlled by the permeability, the deformation affected by the rock rigidity and the geological structure, as well as the boundary conditions. Further understanding on the thermal effects in clays has been gained from various in situ heating tests performed in Underground Research Laboratory (URL): CACTUS, ATLAS and CERBERUS tests in the Hades URL (Belgium) ([7],[8]); HE-D test in Mont Terri URL (Switzeland) ([9],[10]) and TER, TED and ALC tests in Meuse/Haute Marne URL (France) ([11],[12],[13]). All these field tests showed that the sedimentary clays react to heat propagation with a raise of pore pressure and related mechanical effects. The measured results and numerical interpretations have led to better characterisation of THM properties of these clay formations ([14],[15],[16],[17]).

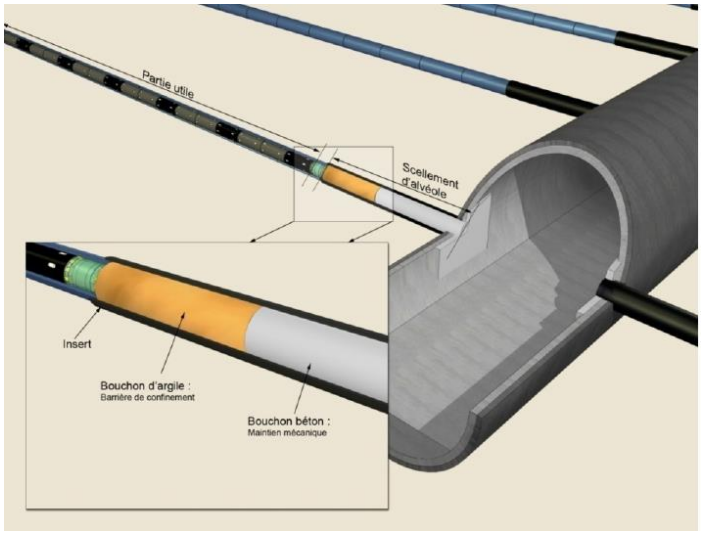


Figure 1: HLW area within nuclear waste storage benchmark concept.

This paper focuses on the analysis of coupled THM processes in COx claystone formation that would result from the development of a high level activity waste (HLW) repository. In the reference concept, the HLW area consists of parallel horizontal cells (i.e., micro-tunnels) excavated from the access gallery (Figure 1). Regarding the periodicity of the micro-tunnels and their lengths, a 2D plane strain model is used for study of the THM behaviour of this area that consists in a plan perpendicular to the cell axis at its middle length. A thermal calculation is performed beforehand in order to simulate the heat emitted from the radioactive waste container as an increase of temperature applied on the cell wall in the THM modeling using the Code_Aster ([6]). Numerical simulations allow also to estimate pore pressure increase and to calculate $\Delta P/\Delta T$, which can be defined as a thermal pressurization coefficient. In fact, this coefficient differs from the classical one determined in undrained conditions. The theoretical analysis shows the dependency of this calculated thermal pressurization on the rate of temperature increase, the stress state and the nature of the rock, i.e., the thermo-poro-

mechanical parameters such permeability, Biot's coefficent, rigidity, thermal conductivity, etc ([2]). As for many sedimentary rocks, the THM parameters of the COx present a significant variability. Experiments show also the dependency of the thermal pressurization on the anisotropies of Young's modulus, of permeability and thermal conductivity tensors. A sensibility analysis is carried out to enhance the understanding of the effect of this variability on the thermal pressurization coefficient.

2 THEORETICAL FORMULATIONS

The theoretical equations used in this work are those of fluid saturated porous media with THM coupling, as described in Coussy ([18]). The coupled THM problem requires a solution that verifies simultaneously the following balance equations:

Momentum balance

$$\nabla.\boldsymbol{\sigma} + \rho \boldsymbol{F}^m = 0 \tag{1}$$

Fluid mass balance

$$\frac{\partial m_w}{\partial t} + \nabla.\boldsymbol{M}_w = 0 \tag{2}$$

Energy balance

$$\boldsymbol{M}_w \boldsymbol{F}^m + \Theta = h_w^m \frac{\partial m_w}{\partial t} + \dot{Q} + \nabla.(h_w^m \boldsymbol{M}_w) + \nabla.\boldsymbol{q} \tag{3}$$

where $\boldsymbol{\sigma}$ (Pa) is the total stress tensor; ρ (kg.m^{-3}) the total homogenized specific mass; $\boldsymbol{F}^m$ (N.kg^{-1}) the mass force density; m_w (kg.m^{-3}) the mass content of water; $\boldsymbol{M}_w$ (kg.m^{-2}.s^{-1}) the water mass flow; h_w^m (J.kg^{-1}) the specific enthalpy of water; Q (J.m^{-3}) the non-convective heat; q (J.m^{-2}.s^{-1}) the heat flow and Θ (W.m^{-3}) the heat source.

In porous medium theory, the total stress tensor is decomposed into two contributions: effective stress $\boldsymbol{\sigma}'$ and Biot'stress $\boldsymbol{\sigma}_p$=-$\boldsymbol{B}p_w$ with $\boldsymbol{B}$ is the Biot's tensor and p_w the pore pressure:

$$\boldsymbol{\sigma} = \boldsymbol{\sigma}' + \boldsymbol{\sigma}_p \tag{4}$$

This preliminary study is focused on the effect of temperature increase in the micro-tunnel on the far field behavior. Therefore, the linear thermo-elastic is assumed for rock skeleton everywhere during the considered time which is consistent far away from micro-tunnel with in situ observations around boreholes. The relation between effective stress $\boldsymbol{\sigma}'$ and total strain $\boldsymbol{\varepsilon}$ is written in the following incremental form:

$$d\boldsymbol{\sigma}' = \boldsymbol{C}:(d\boldsymbol{\varepsilon} - \boldsymbol{\alpha} T) \tag{5}$$

where **C** is the drained elasticity tensor that depends on Young modulus and Poisson ratios, T the temperature and $\boldsymbol{\alpha}$ the tensor of thermal dilatation coefficient.

Considering the Eurelian porosity ϕ, its variation is derived as follows:

$$d\phi = \boldsymbol{B}:d\boldsymbol{\varepsilon} - \phi d\varepsilon_v - 3\boldsymbol{\alpha}_\phi dT - \frac{dp_w}{\boldsymbol{M}_\phi} \tag{6}$$

where ε_v=tr($\boldsymbol{\varepsilon}$) is the volumetric strain, $\boldsymbol{\alpha}_\phi$ the differential dilatation coefficient:

$$\alpha_\phi = \frac{1}{3}(\boldsymbol{B} - \phi\boldsymbol{\delta}) \tag{7}$$

and $\boldsymbol{M}_\phi$ the Biot's modulus:

$$\boldsymbol{M}_\phi = (\boldsymbol{B} - \phi\boldsymbol{\delta}) : \boldsymbol{S}^s : \boldsymbol{\delta} \tag{8}$$

with $\boldsymbol{\delta}$ the identical tensor and $\boldsymbol{S}^s$ the compliance tensor of the solid skeleton.

Fluid diffusion and heat conduction are governed by Darcy's and Fourier's laws respectively:

$$\boldsymbol{q} = -\boldsymbol{\lambda}\nabla T \tag{9}$$

$$\boldsymbol{M}_w = \frac{\rho_w \boldsymbol{k}_{int}}{\mu_w}(-\nabla p_w + \rho_w \boldsymbol{F}^m) \tag{10}$$

where $\boldsymbol{\lambda}$ the thermal conductivity tensor, μ_w the fluid dynamic viscosity and $\boldsymbol{k}_{\text{int}}$ (m^2) the intrinsic permeability tensor.

3 THERMO-HYDRO-MECHANICAL MODELING

According to the reference concept of HLW nuclear waste area (**Figure 1**), THM plane strain model is considered. The geometry model consists in a vertical cut, perpendicular to the cell's axis at its middle length, from the surface to 1000m of depth. The left side of the model passes a cell center, while the right side goes over the mean between two cells. The geological strata are presented in the Table 1. The HLW cells are located within the Callovo-Oxfordian UA layer. The temperature evolution due to heat generation from the waste canisters is considered as thermal load in the THM modeling. Thanks to the model symmetry, thermal flow, fluid flow and displacements are imposed to be nil on its left and right sides. These boundary conditions are also applied to the bottom of model. Concerning the initial conditions, the temperature and the pore pressure change linearly from the surface down to the lower side. The vertical stress is equal the ground weight. The horizontal stress in the plan is equal to the vertical stress while the horizontal stress, in the direction perpendicular to considered plan, is 1.3 times more than the vertical stress

Table 1: Geological strata

Layer	Depth (m)
Kimmerdgien	202
Oxfordian Carbonate	485
Callovo-Oxfordian USC	536
Callovo-Oxfordian UA	635
Dogger	1000

In the following, the results of numerical simulations will be analyzed at a small zone located at the middle distance between two parallel cells at the level of the cell center.

3.1 Thermal pressurization coefficient

The THM simulations allow determining the temperature and interstitial pressure

evolutions at the considered point. Three different thermal loads applied on the cell wall are considered (**Figure 2**). The first one exhibits a fast increase of temperature during first five year and then a rapid decrease of temperature until 200 years followed finally by a gradual attenuation. The second and third ones present also a fast temperature increase for first twenty years but continued afterwards by a progressive decrease until 10000years. They have the same pattern but different in temperature magnitude. **Figure 3** displays the evolution of the pore pressure as function of temperature at the considered zone.

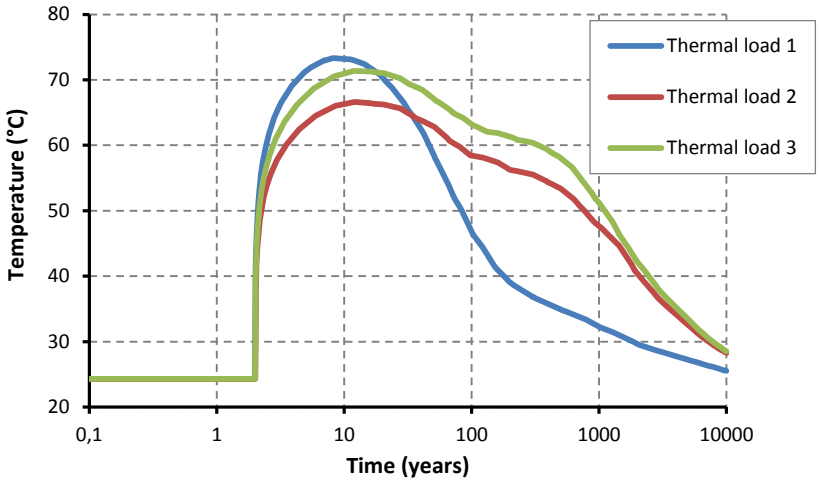


Figure 2: Thermal temperature imposed on the cell wall

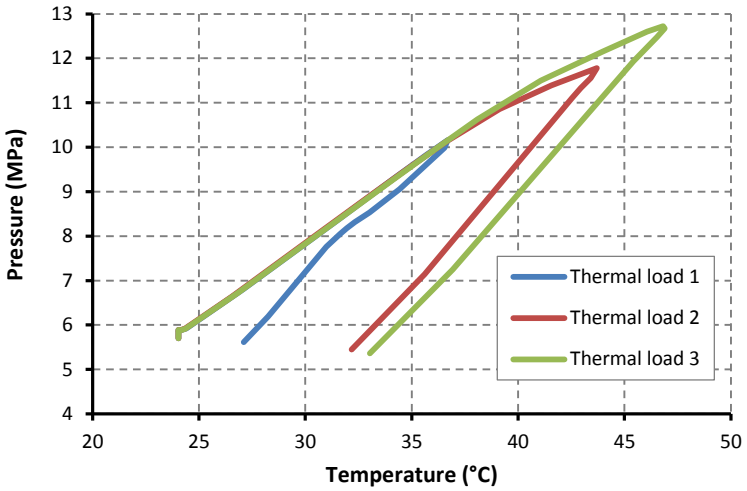


Figure 3: Evolution of pressure versus temperature at considered point.

Two different behaviors can be distinguished during the heating period. Firstly, an almost linear relation between the pore pressure and the temperature is observed when the temperature reaches to about 32°C. Secondly, nonlinearity between these parameters is observed. This observation is translated by a little change of thermal pressurization coefficient before 32° of heating followed by a progressive decrease of this parameter (**Figure 4**). This

phenomenon is also remarked in undrained heating test on Rothbach rock ([2]). Over the cooling time, the decrease of pore pressure is linear regarding to the temperature, i.e., the thermal pressurization coefficient is constant. This coefficient is greater for heating phase than cooling one in the considered configuration. As a conclusion, thermal pressurization coefficient depends on the temperature, as well as the thermal load path. The variation of this coefficient calculated in this work is found in the range obtained by in situ heating test carried out on COx at Meuse/Haute Marne URL ([12],[17]).

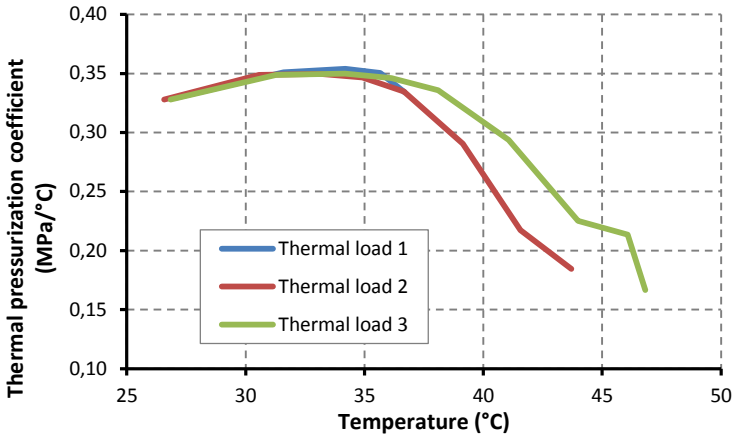


Figure 4: Evolution of thermal pressurization temperature versus temperature during the thermal loading period at considered point.

3.2 Sensitivity analyses

In this section the influence of different parameters on the pore pressure is investigated. A sensitivity analysis has been performed to examine the effect of THM parameters and, in this way, to improve the understanding of the system subject to THM perturbations. To do so, a potential range of variation of parameters is considered. The results of the sensitivity analysis presented here focus on the potentially most important parameters: Young's modulus, Biot's coefficient and intrinsic permeability for the case of thermal load 1 presented in **Figure 2**.

As many bedded formation, Callovo-Oxfordian claystone exhibits transverse isotropic properties. The value of Young's modulus controls the strain in an elastic solid, thus influences also the magnitude of pore pressure since the volumetric strain of the porous media accommodates partially the expansion of water due to the temperature increase. Sensitivity analyses have been performed on this parameter by varying the vertical component of the rigidity tensor, while keeping constant other parameters including the anisotropy ratio between Young's modulus. As showed in **Figure 5**, the Young modulus has a significant effect on the pore pressure increase. The pore pressure variation pattern is unchanged but the increase of Young's modulus leads to an increase of pore pressure magnitude.

The thermally induced increase of pore pressure in porous media is coupled with a decrease in effective stress. Hence, the Biot's coefficient plays its role. Homand et al [19] showed experimentally that this coefficient of COx material depends on the confining stress. The analyses have been realized for three values of Biot's coefficient 0.4, 0.6, 1.0 and the

results of which are presented in **Figure 6**. A decrease of Biot's coefficient results in an increase of pore pressure even if the form of the curve remains unchanged. This observation is quite consistent to undrained configuration.

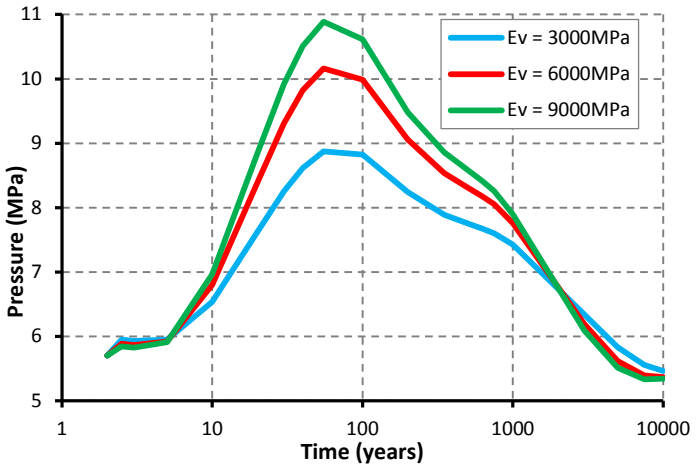


Figure 5: Effect of Young's modulus variation on pore pressure evolution at the considered zone.

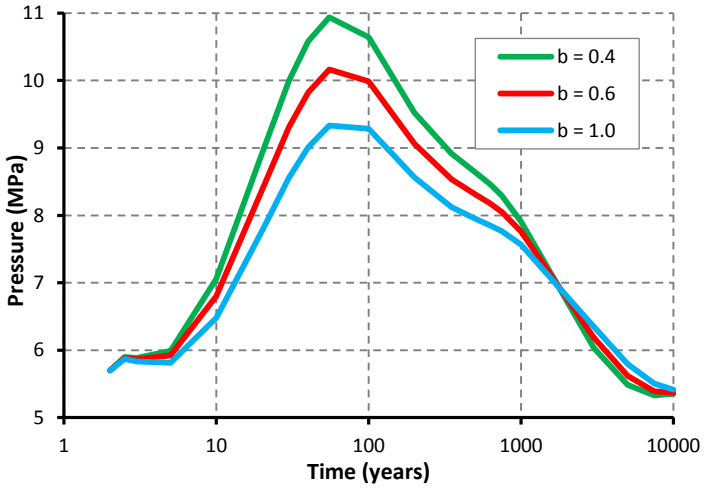


Figure 6: Effect of Biot coefficient variation on pore pressure evolution at the considered zone.

The mechanism underlying the thermo-hydraulic behavior in the clay is a competition between the generation of pore pressure due to the differential thermal dilatation of fluid and skeleton solid and the dissipation of pore pressure, whose rate is essentially controlled by fluid permeability. The effect of the intrinsic permeability on the pore pressure variation is shown in **Figure 7**. A lower value of permeability prevents the fluid dissipation and thus results in a higher pore pressure regardless a higher permeability. Moreover, a lower permeability delays the pressure peak comparing to a higher permeability. This phenomenon is hidden by logarithm scale.

The sensitivity analyses have also carried out on other parameters namely thermal expansion of solid gains, thermal conductivity, anisotropy ratio of Young's modulus and

anisotropy ratio of permeability. The effect of these parameters on the pore pressure is much less significant that three parameters presented above. Moreover, the sensitivity analyses exposed in **Figure 5-Figure 7**show a rather equivalent effect of Young's modulus, Biot's coefficient and intrinsic permeability on the pore pressure variation. This is not the case of in situ heating test observation and interpretation ([14]) where the effect of permeability on generated pore pressure is much more considerable than two others parameters. This difference may be due to the boundary condition. The considered zone is located on the right side of model where the fluid flow and heat flow are nil, i.e. undrained in the horizontal direction. This partially undrained condition may make the effect of permeability less significant. Moreover, the effect of the most influent parameters on the pressure increase depends also to the considered range of the variability. Thus, the obtained results do not represent the intrinsic influence of each parameter, but its influence for the range of its variability.

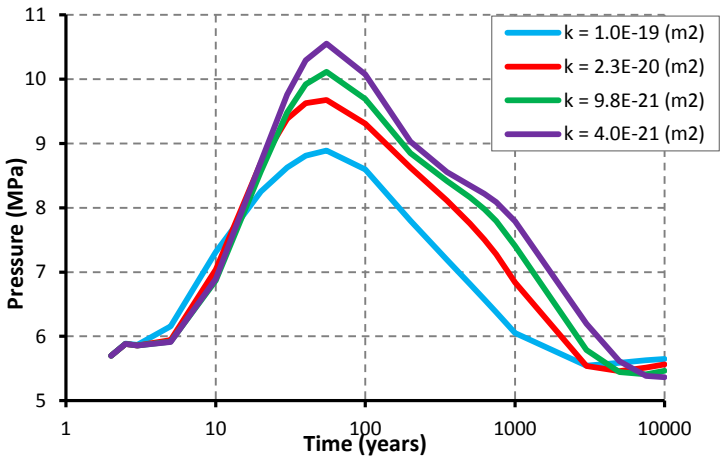


Figure 7: Effect of intrinsic permeability variation on pore pressure evolution at the considered zone.

4 CONCLUSIONS

Thermo-hydro-mechanical coupled process has been investigated in the radioactive high level activity waste repository context. The heat emitted from the waste provokes the pore pressure increase due to the differential expansion of water and solid skeleton. The thermal pressurization coefficient at a particular point, defined as the pore pressure increase induced by 1°C of temperature increase at this point, is calculated for three different thermal load paths. The results show the dependency of this coefficient on the temperature, the thermal load paths, as well as heating-cooling state. A fast increase in temperature yields to an almost constant thermal pressurization coefficient in the low permeability porous medium.

Sensitivity analyses have been also carried out to enhance the understanding of the effect of rock properties on the thermally induced pore pressure in the considered configuration. The significant influences are observed by the variation of Young's modulus, Biot's coefficient and intrinsic permeability. The considerable effect of intrinsic permeability, that controls the fluid dissipation, is not remarked in this study regarding to other works in the literature due to

different boundary conditions. This shows that the THM process depends strongly on the boundary conditions. This study emphasize that Young's modulus, Biot's coefficient and intrinsic permeability play a major role on the overpressure due to heating meaning that it is important to well characterize those parameters and to reduce their incertitude in order to perform reasonable THM calculations. Further calculations will be performed with more complex model of the COx behavior to also study the effect of temperature in the near field.

REFERENCES

[1] Campanella R.G. and Mitchell J.K. Influence of temperature variations on soil behavior. *Jnl Soil Mech. Fdn Div., Am. Soc. Civ. Engrs* (1968) **94**:709-734.

[2] Ghabezloo S. and Sulem J. Stress dependent thermal pressurization of a fluid-saturated rock. *Rock mechanics and Rock engineering* (2009) **42**:1-24.

[3] Mohajerani, M., Delage, P., Sulem, J., Monfared, M., Tang, A.M. and Gatmiri, B. A laboratory investigation of thermally induced pore pressures in the Callovo-Oxfordian Claystone. *International Journal of Rock Mechanics and Mining Sciences* (2012) **52**:112-121.

[4] Monfared, M., Delage, P., Sulem, J., Mohajerani, M. and Tang, A.M. Thermal pressurization in Boom clay. *Proceedings of the 18th International Conference on Soil Mechanics and Geotechnical Engineering*.

[5] Monfared, M., Delage, P., Sulem and J., Mohajerani, M. A laboratory investigation on thermal properties of the Opalinus claystone Rock mechanics and rock engineering (2012) **44** (6), 735-747.

[6] EDF R&D: http://www.code-aster.org.

[7] Bernier, F. and Neerdael, B. Overview of in situ thermomechanical experiments in clay: concept, results and interpretation. *Engng Geol.* (1996) **41**(1–4), 51–64.

[8] De Bruyn and D. Labat, S. The second phase of ATLAS: the continuation of a running THM test in the HADES underground research facility at Mol. *Engng Geol.* (2002) **64**(2–3), 309–316.

[9] Kull, H., Jockwer, N., Zhang, C.L., Wileveau, Y. and Pepa, S.: Measurement of thermally induced pore-water pressure increase and gas migration in the Opalinus claystone at Mont Terri. *Physics and Chemistry of the Earth* (2007) **32**, 937-946.

[10] Jobmann, M. and Polster, M. The response of Opalinus claystone 1 due to heating: A combined analysis of in situ measurements, laboratory investigations and numerical calculations. *Physics and Chemistry of the Earth* (2007) **32**, 929-936.

[11] Wileveau Y. and Su K. In situ thermal experiments carried out in Opalinus claystone and Callovo-Oxfordian claystones by ANDRA experiment set-up and measurement results, *Clay in natural and engineered barriers for radioactive waste confinement*, Conf. Lille, 17-20 Sept, 2007.

[12] Conil, N., Armand, G., Garitte, B., Jobmann, M., Jellouli, M., Filippi, M., De La Vaissière, R. and Morel, J. In situ heating test in Callovo-Oxfordian claystone : measurement and interpretation. *Clay in natural and engineered barriers for radioactive waste confinement*, Conf. Montpellier, 22-25 October, 2012.

[13] Morel, J, Bumbieler, F., Conil, N. and Armand, G. Feasibility and behavior of a full scale disposal cell in a deep clay layer, *EUROCK 2013, Wroclaw, 21-26 September 2013*.

[14] Gens, A., Vaunat, J., Garitte, B. and Wileveau, Y. In situ behaviour of a stiff layered clay subjected to thermal loading: observations and interpretation. *Géotechnique* (2007) **57**(2), 207–228.

[15] Filippi, M. and Conil N., Numerical determination of callovo-oxfordian claystone thermal characteristics: TED experiment. *Clay in natural and engineered barriers for radioactive waste confinement*, Conf. Montpellier, 22-25 October, 2012.

[16] Garitte, B., Gens, A., Vaunat, J., Armand, G. and Conil, N. Analysis of the temperature and pore water pressure field in the TED heating experiment. *Clay in natural and engineered barriers for radioactive waste confinement*, Conf. Montpellier, 22-25 October, 2012.

[17] Garitte, B., Vaunat, J., Gens, A. and Armand G. 3D Interpretation of an in situ heating test in the Callovo-Oxfordian mudstone. *Clays in Natural and Engineered Barriers for Radioactive Waste Confinement*, Conf. Nantes, 29th March - 1st April 2010

[18] Coussy O *Poromechanics* (2004), John Wiley & Sons, The Atrium, Southern Gate, Chichester, England.

[19] Homand, F., Shao, J., Giraud, A., Auvray, C. and Hoxha, D. Pétrofabrique et propriétés mécaniques des claystones. *Comptes Rendus Geosciences* (2006) **338**(12-13), 882–891.

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis(Eds)

ANALYSIS OF TIME FILTERS USED WITH THE LEAPFROG SCHEME

Yong Li* and Catalin Trenchea†

*Department of Mathematics, 301 Thackeray Hall,
University of Pittsburgh, Pittsburgh, PA, 15260, US
e-mail: yol34@pitt.edu, http://www.pitt.edu/~yol34/

†Department of Mathematics, 301 Thackeray Hall,
University of Pittsburgh, Pittsburgh, PA, 15260, US
e-mail: trenchea@pitt.edu, http://www.math.pitt.edu/~trenchea/

Key words: Numerical weather prediction, Leapfrog method, Time filters

Abstract. We present the linear analysis of recent time filters used in numerical weather prediction. We focus on the accuracy and the stability of the leapfrog scheme combined with the Robert–Asselin–Williams filter, the higher-order Robert–Asselin type time filter, the composite-tendency Robert–Asselin–Williams filter and a more discriminating filter.

1 Introduction

The *leapfrog* (LF) time-stepping scheme emerged, from the early years of numerical weather prediction, as the method of choice and is still popular for a number of reasons. Perhaps the most important attribute of the leapfrog scheme is that it preserves exactly the amplitude of a pure oscillation. The dissipative characteristics of other time integration schemes are generally too strong, while the absence of computational damping of leapfrog scheme is especially desirable for long-time integrations. Another feature of the leapfrog method is efficiency, namely, it evaluates the right-hand side of the meteorological tendency equations only once per time step, in contrast with most other schemes. The leapfrog scheme applied to a generic differential equation

$$\frac{du}{dt} = F(u)$$

is given by

$$u^{n+1} = u^{n-1} + 2\Delta t F(u^n),$$

where Δt is the time step and u^n is the approximated solution at time $t^n = n\Delta t$.

The leapfrog method is a three-time-level scheme, and when applied to a simple set of linear differential equations, it generates two modes of motion. One is the *physical* mode, which contributes to the true solution, while the other one is the *computational* mode, which is merely artificial and has no relation to the differential equations that are being integrated. The computational mode of the leapfrog scheme is undamped in linear problems, meaning that it preserves the amplitude in each time step. In nonlinear problems, however, the nonlinear terms introduce couplings between the physical and computational modes which may amplify the computational mode. In short-time simulations of weather and climate, the growth of the computational mode is generally hard to detect, but when long-time integrations are considered, the computational mode dominates the solution.

One possible approach to control the leapfrog scheme's computational mode is to periodically use a two-time-level scheme, e.g., a Matsuno step after every 11 leapfrog steps [21]. The idea is to reset the amplitude to zero periodically, so it never becomes large enough/problematic. Another technique is to use different explicit time-stepping schemes, e.g., the second-order Adams-Bashforth method [18], the third-order Adam-Bashforth [8], the leapfrog-trapezoidal method [14, 33] or the Magazenkov method [19].

The ubiquitous strategy in atmospheric models, for controlling the leapfrog scheme's computational mode, is the non-intrusive implementation of a time filter after each leapfrog time step. Robert [25] designed such a filter, which Asselin [3] analyzed and proved to effectively damp the computational mode of the leapfrog scheme. This time filter is referred to as the Robert–Asselin (RA) filter. The RA-filtered leapfrog scheme is defined by

$$\begin{aligned} v^{n+1} &= u^{n-1} + 2\Delta t F(v^n), \\ u^n &= v^n + \frac{\nu}{2}\left(v^{n+1} - 2v^n + u^{n-1}\right), \end{aligned}$$

where v and u denote the unfiltered (provisional) and filtered (definitive) variables, respectively. The dimensionless parameter $\nu \in [0, 1]$ determines the strength of the filter.

The accuracy and stability properties of the RA filter were investigated in [4, 12, 6, 10, 26, 7, 24, 5, 28, 13]. Currently, the RA filter is used in operational numerical weather prediction models, atmospheric general circulation models for climate simulation, ocean general circulation models, and models of fluids in rotating annulus laboratory experiments, etc. A comprehensive list of atmospheric models with RA filter can be found in [28]. Unfortunately, the RA filter also damps the physical mode. As a result, the formal second-order accuracy of the leapfrog scheme is reduced to first order, and can degrade the accuracy of model simulations. Therefore, physical quantities (e.g., energy) conserved by the time-continuous equations are not necessarily conserved by time-discretized equations when the filter is used.

Because the RA-filtered leapfrog scheme is widely used in legacy codes for atmospheric models, non-intrusive and simple-to-implement improvements of RA appear attractive, in order to avoid the significant programming undertaking. Williams [28] proposed a

modification of the RA filter, which combined with the leapfrog scheme is

$$\begin{aligned} w^{n+1} &= u^{n-1} + 2\Delta t F(v^n), \\ u^n &= v^n + \frac{\nu\alpha}{2}(w^{n+1} - 2v^n + u^{n-1}), \\ v^{n+1} &= w^{n+1} - \frac{\nu(1-\alpha)}{2}(w^{n+1} - 2v^n + u^{n-1}), \end{aligned}$$

where w, v, and u denote the unfiltered, once filtered, and twice filtered variables, respectively. The parameter ν is as in the RA filter, and the new dimensionless parameter $\alpha \in [0.5, 1]$. Linear analysis shows that ν plays a role in controlling the computational mode of the leapfrog scheme, while α is the remedy in restoring accuracy. The filter is now referred to as the Robert–Asselin–Williams (RAW) filter. It reduces the negative impact of the RA filter on the physical mode and increases the numerical accuracy to second order, at the price of a slight instability. The filter has been implemented and studied in [27, 1, 23, 32, 20, 31, 13], and its behavior in implicit-explicit (IMEX) integrations was analyzed in [29].

Later, Williams [30] proposed two methods for further improving the RAW-filtered leapfrog scheme. The first algorithm is a combination of the RAW filter with a composite-tendency leapfrog (CTLF) scheme:

$$\begin{aligned} w^{n+1} &= u^{n-1} + 2\Delta t[\gamma F(v^n) + (1-\gamma)F(w^n)], \\ u^n &= v^n + \frac{\nu\alpha}{2}(w^{n+1} - 2v^n + u^{n-1}), \\ v^{n+1} &= w^{n+1} - \frac{\nu(1-\alpha)}{2}(w^{n+1} - 2v^n + u^{n-1}), \end{aligned}$$

where γ is a real number. A more discriminating filter takes the form $(1, -4, 6, -4, 1)$ instead of $(1, -2, 1)$, and the scheme is

$$\begin{aligned} w^{n+1} &= u^{n-1} + 2\Delta t[\gamma F(v^n) + (1-\gamma)F(w^n)], \\ u^n &= v^n + \nu\alpha(w^{n+1} - 4v^n + 6u^{n-1} - 4u^{n-2} + u^{n-3}), \\ v^{n+1} &= w^{n+1} - \nu(1-\alpha)(w^{n+1} - 4v^n + 6u^{n-1} - 4u^{n-2} + u^{n-3}). \end{aligned}$$

Both methods are computationally more demanding since they require two tendency calculations per time step, which is the most expensive component of contemporary atmosphere and ocean models. Nevertheless, the improvements to the amplitude accuracy are considerable, especially, the latter. The increased accuracy may allow a longer time step for the same error tolerance, tending to offset the increased expense. The RAW-filtered leapfrog scheme is analyzed in [17], and its behavior in IMEX integrations is studied in [2].

Recently, Li and Trenchea [15] proposed a higher-order Robert–Asselin (hoRA) type time filter.

$$v^{n+1} = u^{n-1} + 2\Delta t F(v^n),$$
$$u^n = v^n + \frac{\beta}{2}\left(v^{n+1} - 2v^n + u^{n-1}\right) - \frac{\beta}{2}\left(v^n - 2u^{n-1} + u^{n-2}\right),$$

where the dimensionless parameter $\beta \in [0, 0.4]$. Under the same computational cost as RAW, the hoRA-filtered leapfrog scheme exhibits third-order accuracy. Compared with the third-order Adams-Bashforth method, the hoRA-filtered leapfrog scheme is almost as accurate, stable and efficient, yet easily implementable in legacy codes. A study of the filter in IMEX integrations was conducted in [16].

In the sequel we present the linear analysis for the leapfrog scheme combined with the aforementioned time filters, focusing on the accuracy and stability.

2 Linear analysis of the leapfrog scheme combined with time filters

We now derive the stability condition, amplitude, phase-speed, and the consistency errors. These properties are illustrated by analyzing solutions to the pure oscillation equation (see e.g., [8, 9])

$$\frac{du}{dt} = i\omega\, u, \tag{1}$$

where i is the imaginary unit, and ω a real constant. Define the *amplification factor* A as the ratio of the approximate solution at two adjacent time steps, $A = u^{n+1}/u^n$. The amplification factor yields information on two quantities of interest: the amplitude and the relative phase change per time step. Specifically, A can be expressed in modulus-argument form $A = |A|e^{i\theta}$, where

$$|A| = \sqrt{Re(A)^2 + Im(A)^2}, \quad \theta = tan^{-1}(Im(A)/Re(A)).$$

For the true solution to the oscillation equation (1), the *exact amplification factor* $A_{\rm e} = e^{i\omega\Delta t}$ has unity magnitude and phase change $\omega\Delta t$ over a time interval Δt. The amplitude errors are defined as the difference between the magnitude of the approximate amplification factor $|A|$ and the correct value of unity. When $|A| = 1$, the scheme is *neutral*, if $|A| < 1$, the scheme is *damping* (indicating stability), and if $|A| > 1$, it is *amplifying* (instability). The relative phase change or the phase speed, on the other hand, is measured by the ratio of the phase change of the numerical scheme per time step divided by the phase change of the true solution over the same time interval, and is denoted by $R = \theta/\omega\Delta t$. The phase-speed errors are defined as the difference between the phase speed R and the unity over a time interval Δt. When $R > 1$, the method is *accelerating*, and if $R < 1$, it is *decelerating*. Unlike the amplitude, the phase change does not influence the stability of the numerical solution. Instead, the phase errors accumulate and can become large over a long time period of integration.

2.1 The hoRA-filtered leapfrog scheme

The hoRA-filtered leapfrog (LF-hoRA) scheme [15] applied to (1) is

$$v^{n+1} = u^{n-1} + 2i\omega\Delta t v^n, \tag{2}$$

$$u^n = v^n + \frac{\beta}{2}\left(v^{n+1} - 2v^n + u^{n-1}\right) - \frac{\beta}{2}\left(v^n - 2u^{n-1} + u^{n-2}\right). \tag{3}$$

The system of equations (2)-(3) is equivalent to the following linear multistep method:

$$u^{n+1} - 2\beta u^n - (1-2\beta)u^{n-1} = i\omega\Delta t(2u^n - 3\beta u^{n-1} + \beta u^{n-2}). \tag{4}$$

2.1.1 Consistency errors, amplitude errors and phase-speed errors

Using Taylor expansion, the local truncation error of (4) is shown to be

$$\tau_n(\Delta t) = \frac{2-5\beta}{6}(i\omega\Delta t)^2 u'(t^n) + \frac{11\beta}{12}(i\omega\Delta t)^3 u'(t^n) + \mathcal{O}[(i\omega\Delta t)^4].$$

Thus, the LF-hoRA scheme is second order in general, and third order when $\beta = 0.4$.

Formula (4) yields the following equation for the amplification factor:

$$A^3 - 2(\beta + i\omega\Delta t)A^2 + (3\beta i\omega\Delta t - 1 + 2\beta)A - \beta i\omega\Delta t = 0. \tag{5}$$

Equation (5) has three roots, one is the physical mode denoted A_p, and the other two are computational modes. Since computational modes are well-controlled by the filter, we focus on the amplitude and phase-speed errors for the physical mode. A series expansion for $|A_p|$ in powers of $\omega\Delta t$ yields the amplitude error as follows:

$$|A_p| - |A_e| = |A_p| - 1 = \frac{\beta(2\beta-3)}{8(1-\beta)^2}(\omega\Delta t)^4 + \mathcal{O}[(\omega\Delta t)^6].$$

The amplitude error after taking a single time step scales as $(\Delta t)^4$, hence it is of order $(\Delta t)^3$ over $T/\Delta t$ time steps. The phase-speed error is

$$R_p - 1 = \frac{\arg(A_p)}{\omega\Delta t} - 1 = \frac{2-5\beta}{12(1-\beta)}(\omega\Delta t)^2 + \mathcal{O}[(\omega\Delta t)^4].$$

The phase speed of the physical mode is fourth-order accurate when $\beta = 0.4$ and second order otherwise.

2.1.2 Stability analysis

To determine the maximum $\omega\Delta t$ for which all numerical amplification factors of the LF-hoRA scheme are non-amplified, we use the *root locus curve* method (see e.g., [11]). The characteristic equation of (4) is

$$\zeta^3 - 2\beta\zeta^2 - (1-2\beta)\zeta - z(2\zeta^2 - 3\beta\zeta + \beta) = 0,$$

where ζ denotes the points on the unit circle, i.e., $\zeta = e^{i\theta}$ for $\theta \in [0, 2\pi]$, and $z \in \mathbb{C}$. The curve z is called the *root locus curve*. In our case $z = i\omega\Delta t$ lies on the imaginary axis, and consequently θ satisfies

$$\cos\theta = 1 \text{ or } \cos\theta = \beta - \frac{1}{2}, \quad \text{and hence} \quad z = 0 \text{ or } z = \pm i\frac{\sqrt{\frac{3}{4} + \beta - \beta^2}}{1 + \frac{3}{2}\beta - \beta^2},$$

which indicates the intersections of the root locus curve with the imaginary axis in the complex plane. Thus, the stability of the LF-hoRA scheme is provided by

$$\omega\Delta t \leq \frac{\sqrt{\frac{3}{4} + \beta - \beta^2}}{1 + \frac{3}{2}\beta - \beta^2}, \quad 0 < \beta \leq 0.4.$$

2.2 The RAW-filtered composite-tendency leapfrog scheme

Notice that the RA-filtered leapfrog (LF-RA) scheme is recovered when $\alpha = 1$ in the RAW-filtered composite-tendency leapfrog (CTLF-RAW) scheme, while LF-RAW scheme is a special case of CTLF-RAW when $\gamma = 1$. For this reason, it suffices to analyze CTLF-RAW (refer to [30] for more details). The scheme applied to (1) is

$$w^{n+1} = u^{n-1} + 2i\omega\Delta t(\gamma v^n + (1-\gamma)w^n), \tag{6}$$

$$u^n = v^n + \frac{\nu\alpha}{2}(w^{n+1} - 2v^n + u^{n-1}), \tag{7}$$

$$v^{n+1} = w^{n+1} - \frac{\nu(1-\alpha)}{2}(w^{n+1} - 2v^n + u^{n-1}). \tag{8}$$

The three dimensionless parameters in the scheme are ν, α, and γ, where ν corresponds to the classical Robert–Asselin filter parameter, α partitions the RAW filter displacements between the n'th and $(n+1)$'th time levels, and γ specifies the weighting coefficients for the composite tendency. Although previous work [30] assumed $0 \leq \gamma \leq 1$, here we allow γ to vary outside this range.

The system of equations (6)-(8) is equivalent to the following linear multistep method:

$$\begin{aligned} &u^{n+1} - \nu u^n - (1-\nu)u^{n-1} \\ &= i\omega\Delta t\Big((2 - \nu\gamma(1-\alpha))u^n + \nu(2\gamma + \alpha - 2 - 2\alpha\gamma)u^{n-1} + \nu(1-\alpha)(1-\gamma)u^{n-2}\Big). \end{aligned} \tag{9}$$

2.2.1 Consistency errors, amplitude errors and phase-speed errors

The local truncation error of (9) is

$$\begin{aligned} \tau_n(\Delta t) = \left(\frac{1}{2} - \alpha\right)\nu(i\omega\Delta t)u'(t^n) + \frac{1}{6}\left(2 - \nu(7 - 9\alpha) + 6\nu\gamma(1-\alpha)\right)(i\omega\Delta t)^2 u'(t^n) \\ + \frac{\nu}{24}(25 - 28\alpha - 24\gamma + 24\alpha\gamma)(i\omega\Delta t)^3 u'(t^n) + \mathcal{O}(\Delta t^4). \end{aligned}$$

The scheme is generally first-order accurate[1] if $\nu \neq 0$, and second order if $\alpha = 0.5$, as noted by Williams [30]. Further, the method becomes third order if $\alpha = 0.5$ and $\gamma = (5\nu - 4)/(6\nu)$. This third-order scheme would require $\gamma < 0$ if $\nu < 4/5$. Finally, the scheme exhibits fourth-order accuracy if $\alpha = 0.5, \nu = -8$, and $\gamma = 11/12$. This case is of no practical interest because the negative value of ν forces the computational mode to be amplified.

Remark 2.1 *The LF-RA scheme is first-order accurate. The LF-RAW is firsr order in general, and second order when* $\alpha = 0.5$.

The amplitude error of CTLF-RAW is given by

$$|A_p| - 1 = \frac{\nu(1-2\alpha)}{2(2-\nu)}(\omega\Delta t)^2 + \mathcal{O}\left[(\omega\Delta t)^4\right],$$

yielding first-order amplitude accuracy, independent of γ. Since LF-RA recovers when $\alpha = 1$, its amplitude is therefore first order. When $\alpha = 1/2$, the quadratic term vanishes and the amplitude error becomes

$$|A_p| - 1 = \frac{\nu(4\gamma - 3 + \nu - \nu\gamma)}{4(2-\nu)^2}(\omega\Delta t)^4 + \mathcal{O}\left[(\omega\Delta t)^6\right],$$

which implies the third-order amplitude accuracy. The fourth-order term now depends on γ. Specifically, CTLF-RAW is amplifying when $\gamma > (3-\nu)/(4-\nu)$, and is damping if $\gamma < (3-\nu)/(4-\nu)$. Recall that LF-RAW is recovered when $\gamma = 1$, hence it is unstable when $\alpha = 1/2$. When $\gamma = (3-\nu)/(4-\nu)$, the amplitude error is fifth-order accurate:

$$|A_p| - 1 = \frac{\nu}{4(4-\nu)(2-\nu)^2}(\omega\Delta t)^6 + \mathcal{O}\left[(\omega\Delta t)^8\right].$$

However, the coefficient of the sixth-order term is always positive, implying a slight instability of the scheme. The phase-speed error, when $\alpha = 1/2$, is

$$R_p - 1 = \frac{6\nu\gamma + 4 - 5\nu}{12(2-\nu)}(\omega\Delta t)^2 + \mathcal{O}[(\omega\Delta t)^4].$$

The phase speed is fourth-order accurate if further $\gamma = (5\nu - 4)/(6\nu)$.

2.2.2 Stability analysis

Using the similar technique as in Section 2.1.2, we derive the stability condition for the CTLF-RAW scheme. First, the time step condition for LF-RA is

$$\omega\Delta t \leq \sqrt{\frac{2-\nu}{2+\nu}}.$$

[1]If $\nu = 0$, the scheme is generally second order, but then the filter is inactive and the computational mode is uncontrolled. For this reason, $\nu \neq 0$ is not considered throughout the article.

The LF-RAW is stable under the following condition:

$$\omega\Delta t \leq \frac{1}{\alpha}\sqrt{\frac{(2-\nu)(2\alpha-1)}{2-\nu+2\alpha\nu}}, \quad \alpha \in [1/2, 1].$$

For the CTLF-RAW method, it is of more interest when $\alpha = 1/2$ since the scheme is at least second-order accurate. The stability condition in this case is given by

$$\omega\Delta t \leq \frac{2}{(1-\gamma)(4-\nu)}\sqrt{\frac{(3-\nu)-(4-\nu)\gamma}{1+\nu(1-\gamma)}}, \quad \gamma \leq (3-\nu)/(4-\nu).$$

2.3 The composite-tendency leapfrog scheme with more discriminating filter

Applied to equation (1), the scheme [30] is

$$w^{n+1} = u^{n-1} + 2i\omega\Delta t(\gamma v^n + (1-\gamma)w^n), \tag{10}$$
$$u^n = v^n + \nu\alpha(w^{n+1} - 4v^n + 6u^{n-1} - 4u^{n-2} + u^{n-3}), \tag{11}$$
$$v^{n+1} = w^{n+1} - \nu(1-\alpha)(w^{n+1} - 4v^n + 6u^{n-1} - 4u^{n-2} + u^{n-3}), \tag{12}$$

which is equivalent to the following linear multistep method:

$$\begin{aligned} &u^{n+1} - \nu(4+3\alpha)u^n - (1-7\nu-\nu\alpha)u^{n-1} - \nu(4-3\alpha)u^{n-2} + \nu(1-\alpha)u^{n-3} \\ &= 2\left(1-\nu(1-\alpha)\gamma\right)u^n - \nu\left(8(1-\alpha)(1-\gamma)+12\alpha\right)u^{n-1} \\ &+ \nu\left(12(1-\alpha)(1-\gamma)+8\alpha\right)u^{n-2} - \nu\left(8(1-\alpha)(1-\gamma)+2\alpha\right)u^{n-3} + 2\nu(1-\alpha)(1-\gamma)u^{n-4}. \end{aligned} \tag{13}$$

2.3.1 Consistency errors, amplitude errors and phase-speed errors

The local truncation error of (13) is

$$\begin{aligned} \tau_n(\Delta t) = &\frac{1-\nu(1+2\alpha)}{3}(i\omega\Delta t)^2 u'(t^n) + \frac{\nu(3-5\alpha)}{3}(i\omega\Delta t)^3 u'(t^n) \\ &+ \frac{181\nu - 308\alpha\nu - 120\nu\gamma + 120\alpha\nu\gamma - 1}{60}(i\omega\Delta t)^4 u'(t^n) + \mathcal{O}[(i\omega\Delta t)^5]. \end{aligned}$$

Theoretically, the scheme could be third-order accurate if $\alpha = (1-\nu)/(2\nu)$, and even higher-order accurate for appropriate values of the parameters which set zero the coefficients of the higher-order terms. However, the root condition is not satisfied in this case. To see this, set $\omega = 0$ and write (13) in terms of the amplification factor A:

$$A^4 - \nu(4+3\alpha)A^3 - (1-7\nu-\nu\alpha)A^2 - \nu(4-3\alpha)A + \nu(1-\alpha) = 0. \tag{14}$$

It turns out that when $\alpha = (1-\nu)/(2\nu)$, equation (14) has the root $A = 1$ with multiplicity two, violating the root condition (see e.g., [22]). Indeed, the numerical solution grows

linearly in time, while is supposed to be constant. Thus, the scheme is second-order accurate.

The amplitude error is

$$|A_p| - 1 = -\frac{\nu(1-2\alpha)}{2(1-\nu-2\alpha\nu)}(\omega\Delta t)^4 + \mathcal{O}[(\omega\Delta t)^6],$$

which, by setting $\alpha = 1/2$, becomes

$$|A_p| - 1 = \frac{\nu(5-8\gamma-9\nu+14\nu\gamma)}{8(1-2\nu)^2}(\omega\Delta t)^6 + \mathcal{O}[(\omega\Delta t)^8].$$

Further, the sixth-order term vanishes when $\gamma = (5-9\nu)/(2(4-7\nu))$ and gives the seventh-order amplitude error:

$$|A_p| - 1 = -\frac{5\nu(4-13\nu+11\nu^2)}{32(1-2\nu)^2(4-7\nu)}(\omega\Delta t)^8 + \mathcal{O}[(\omega\Delta t)^{10}].$$

The phase-speed error, in this case, is second order:

$$R_p - 1 = \frac{1}{6}(\omega\Delta t)^2 + \mathcal{O}[(\omega\Delta t)^4].$$

2.3.2 Stability

As shown in the consistency error analysis, the composite-tendency leapfrog scheme with the more discriminating filter is second-order accurate regardless of the parameters. Nevertheless, the amplitude exhibits the highest-order of accuracy when $\alpha = 1/2$ and $\gamma = (5-9\nu)/(2(4-7\nu))$. For this reason, we only consider the stability for the chosen values of the parameters. Applying the root locus curve technique, the time step condition for this scheme is

$$\omega\Delta t \leq \sqrt{1-\left(\frac{8-45\nu+55\nu^2}{12-20\nu}\right)^2}\,\frac{8(4-7\nu)(2-5\nu+5\nu^2)}{(4+25\nu-55\nu^2)(16-68\nu+105\nu^2-55\nu^3)}. \tag{15}$$

3 Conclusions

The development of accurate and efficient time-stepping schemes is an important key in improving the fidelity of the numerical simulations for weather and climate, and is still an active area of research.

We surveyed the recent progress on time filters, a post-processing non-intrusive technique which improves accuracy and stability, and uses legacy codes in a black-box manner. We focus on time filters used in conjunction with the leapfrog scheme, the most commonly employed time-stepping scheme in the weather and climate community. Specifically, we present the accuracy and stability analysis of RA, RAW, hoRA, and the more discriminating filtered leapfrog/composite-tendency leapfrog schemes. The properties of these methods are summarized in Table 1, an addendum to the comparison Table 2.2 in [9].

Method	Order	Amplitude	Phase speed	Maximum $\omega\Delta t$
LF-RA	1	$1-\frac{\nu}{2(2-\nu)}p^2$	$1+\frac{1+\nu}{3(2-\nu)}p^2$	$\sqrt{\frac{2-\nu}{2+\nu}}$
LF-RAW	1 or 2	$1-\frac{\nu(2\alpha-1)}{2(2-\nu)}p^2+\mathcal{O}(p^4)$	$1+\left(\frac{(1-\nu(1-\alpha))(2-\alpha\nu)}{(2-\nu)^2}-\frac{1}{3}\right)p^2$	$\frac{1}{\alpha}\sqrt{\frac{(2-\nu)(2\alpha-1)}{2-\nu+2\alpha\nu}}$
LF-hoRA	2 or 3	$1-\frac{\beta(3-2\beta)}{8(1-\beta)^2}p^4$	$1+\frac{2-5\beta}{12(1-\beta)}p^2+\mathcal{O}(p^4)$	$\frac{\sqrt{\frac{3}{4}+\beta-\beta^2}}{1+\frac{3}{2}\beta-\beta^2}$
CTLF-RAW	2 or 3	$1+\frac{\nu(4\gamma-3+\nu-\nu\gamma)}{4(2-\nu)^2}p^4+\mathcal{O}(p^6)$	$1+\frac{6\nu\gamma+4-5\nu}{12(2-\nu)}p^2+\mathcal{O}(p^4)$	$\sqrt{\frac{4((3-\nu)-(4-\nu)\gamma)}{(1+\nu(1-\gamma))[(1-\gamma)(4-\nu)]^2}}$
CTLF-D	2	$1-\frac{5\nu(4-13\nu+11\nu^2)}{32(1-2\nu)^2(4-7\nu)}p^8$	$1+\frac{1}{6}p^2$	Formula (15)

Table 1: Comparison between the leapfrog scheme combined with time filters. The amplitude, phase speed, and time step limitations are those associated with the application of each scheme to the oscillation equation (1). For brevity, the more discriminating filtered composite-tendency leapfrog scheme is abbreviated by CTLF-D. We denote $p=\omega\Delta t$, and amplitude or phase speed that is with $\mathcal{O}(p^k)$ indicates that it is able to be of order up to p^k.

Acknowledgement

The authors received partial support from an Air Force grant FA 9550-12-1-0191 and an NSF grant DMS 1216465. We also would like to thank Paul D. Williams (University of Reading, UK) for his helpful comments.

REFERENCES

[1] J. Amezcua, E. Kalnay, and P. D. Williams, *The effects of the RAW filter on the climatology and forecast skill of the SPEEDY model*, Mon. Wea. Rev., 139 (2011), pp. 608–619.

[2] J. Amezcua and P. D. Williams, *The composite-tendency Robert–Asselin–Williams (RAW) filter in semi-implicit integrations*, Q. J. R. Meteorol. Soc, (2015).

[3] R. Asselin, *Frequency filter for time integrations*, Mon. Wea. Rev., 100 (1972), pp. 487–490.

[4] R. Bleck, *Short-range prediction in isentropic coordinates with filtered and unfiltered numerical models*, Mon. Wea. Rev., 102 (1974), pp. 813–829.

[5] E. Cordero and A. Staniforth, *A problem with the Robert–Asselin time filter for three-time-level semi-implicit semi-lagrangian discretizations*, Mon. Wea. Rev., 132 (2004), pp. 600–610.

[6] R. Daley, C. Girard, J. Henderson, and I. Simmonds, *Short-term forecasting with a multilevel spectral primitive equation model part I - model formulation*, Atmosphere, 14 (1976), pp. 98–116.

[7] M. Déqué and D. Cariolle, *Some destabilizing properties of the Asselin time filter*, Mon. Wea. Rev., 114 (1986), pp. 880–884.

[8] D. R. Durran, *The third-order Adams-Bashforth method: An attractive alternative to leapfrog time differencing*, Mon. Wea. Rev., 119 (1991), pp. 702–720.

[9] ——, *Numerical methods for fluid dynamics*, vol. 32 of Texts in Applied Mathematics, Springer, New York, second ed., 2010. With applications to geophysics.

[10] C. T. Gordon and W. F. Stern, *A description of the GFDL global spectral model*, Mon. Wea. Rev., 110 (1982), pp. 625–644.

[11] E. Hairer and G. Wanner, *Solving ordinary differential equations. II*, vol. 14 of Springer Series in Computational Mathematics, Springer-Verlag, Berlin, 2010. Stiff and differential-algebraic problems, Second revised edition.

[12] G. J. Haltiner and J. M. McCollough, *Experiments in the initialization of a global primitive equation model*, J. Appl. Meteor., 14 (1975), pp. 281–288.

[13] N. Hurl, W. Layton, Y. Li, and C. Trenchea, *Stability analysis of the Crank–Nicolson-Leapfrog method with the Robert–Asselin–Williams time filter*, BIT, 54 (2014), pp. 1009–1021.

[14] Y. Kurihara, *On the use of implicit and iterative methods for the time integration of the wave equation*, Mon. Wea. Rev., 93 (1965), pp. 33–46.

[15] Y. Li and C. Trenchea, *A higher-order Robert–Asselin type time filter*, J. Comput. Phys., 259 (2014), pp. 23–32.

[16] ——, *A higher order Robert-Asselin type time filter in the semi-explicit integrations*, tech. rep., University of Pittsburgh, 2015.

[17] Y. Li and P. D. Williams, *Analysis of the RAW filter in composite-tendency leapfrog integrations*, tech. rep., University of Pittsburgh, 2014.

[18] D. Lilly, *On the computational stability of numerical solutions of time-dependent non-linear geophysical fluid dynamics problems*, Mon. Wea. Rev., 93 (1965), pp. 11–25.

[19] L. Magazenkov, *Trudy Glavnoi Geofizicheskoi Observatorii*, (Transactions of the Main Geophysical Observatory), 410 (1980), pp. 120–129.

[20] N. Oger, O. Pannekoucke, A. Doerenbecher, and P. Arbogast, *Assessing the influence of the model trajectory in the adaptive observation Kalman filter sensitivity method*, Q. J. R. Meteorol. Soc, 138 (2012), pp. 813–825.

[21] R. L. Pfeffer, I. M. Navon, and X. Zou, *A comparison of the impact of two time-differencing schemes on the NASA-GLAS climate model*, Mon. Wea. Rev., 120 (1992), pp. 1381–1393.

[22] A. Quarteroni, R. Sacco, and F. Saleri, *Numerical mathematics*, vol. 37 of Texts in Applied Mathematics, Springer-Verlag, Berlin, second ed., 2007.

[23] D. Ren and L. M. Leslie, *Three positive feedback mechanisms for ice-sheet melting in a warming climate*, J. Glaciol., 57 (2011), pp. 1057–1066.

[24] A. Robert and M. Lépine, *An anomaly in the behaviour of the time filter used with the leapfrog scheme in atmospheric models*, Atmosphere-Ocean, 35 (1997), pp. S3–S15.

[25] A. J. Robert, *The integration of a low order spectral form of the primitive meteorological equations*, J. Met. Soc. Japan, 44, Ser. 2 (1966), pp. 237–245.

[26] R. E. Schlesinger, L. W. Uccellini, and D. R. Johnson, *The effects of the Asselin time filter on numerical solutions to the linearized shallow-water wave equations*, Mon. Wea. Rev., 111 (1983), pp. 455–467.

[27] M. Watanabe and Coauthors, *Improved climate simulation by MIROC5: Mean states, variability, and climate sensitivity*, J. Climate, 23 (2010), pp. 6312–6335.

[28] P. D. Williams, *A proposed modification to the Robert–Asselin time filter*, Mon. Wea. Rev., 137 (2009), pp. 2538–2546.

[29] P. D. Williams, *The RAW filter: An improvement to the Robert–Asselin filter in semi-implicit integrations*, Mon. Wea. Rev., 139 (2011), pp. 1996–2007.

[30] P. D. Williams, *Achieving seventh-order amplitude accuracy in leapfrog integrations*, Mon. Wea. Rev., 141 (2013), pp. 3037–3051.

[31] C.-C. Young, Y.-C. Liang, Y.-H. Tseng, and C.-H. Chow, *Characteristics of the RAW-filtered leapfrog time-stepping scheme in the ocean general circulation model*, Mon. Wea. Rev., 142 (2013), pp. 434–447.

[32] C.-C. Young, Y.-H. Tseng, M.-L. Shen, Y.-C. Liang, M.-H. Chen, and C.-H. Chien, *Software development of the TaIwan Multi-scale Community Ocean Model (TIMCOM)* , Environ. Modell. Software, 38 (2012), pp. 214 – 219.

[33] S. Zalesak, *Fully multidimensional flux-corrected transport algorithms for fluids*, J. Comput. Phys., 31 (1979), pp. 335–362.

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis (Eds)

MATHEMATICAL MODELING OF REACTION MECHANISM OF FORMATION OF PHOTOCHEMICAL SMOG BY APPLYING THE SEMI-IMPLICIT METHOD

LEÓN R. FRIDA*, AGUILAR M. ARMANDO†, GARCÍA R. OMAR†, PINEDA B. MIGUEL† AND AXOTLA G. JUAN†

* Departamento de Ciencias Químicas, Facultad de Estudios Superiores Cuautitlán UNAM
Km. 2.5 Cuautitlán Teoloyucan, 54714 Estado de México
e-mail: fridam@unam. mx

† Departamento de Matematicas, Facultad de Estudios Superiores Cuautitlán UNAM
Km. 2.5 Cuautitlán Teoloyucan, 54714 Estado de México
e-mail: caronte4509@yahoo.com.mx, mnazarethp@gmail.com, jc_axotla@fesc.cuautitlan2.unam.mx

Key words: Rigid System, Jacobean, Formaldehyde, Photochemical Smog.

Abstract. To simulate the different mechanisms we considered a reactor of constant temperature and volume, where the only reactions that are carried out are those with reported kinetic constants [1]. For example, to simulate the formaldehyde kinetic you make a serial of seven chemical reactions where intervene nine chemical species [2].The change in concentration with respect to the time of one specie is mathematically represented by means of an ordinary differential equation. In the studied cases, the mechanism of reaction can be represented as a system of nonlinear ordinary differential equations. In the simulation of the mechanism of reaction, the medullar part is the solution of all the ordinary differential equation that describe the temporary evolution of the concentration of each the species. The differential equation that comes from the kinetic present what it is called rigidity, principally due to the simultaneous presence of radical with called rigidity, principally due to the simultaneous presence of radicals with a really short life time as the presence of hydroperoxide HO_2• and species that remain almost constant as the oxygen. The main problem to carry out the precise integration of the differential equation system that represents the mechanism of reaction of the atmospheric chemistry, is the wide when variation of the kinetic constant as it appears in the reactions 2 and 7 which brings as a consequence the instability when applying an explicit numerical method because for any change in so different scales. When this happens, it is said that the differential equation system is rigid. In order to solve the problem that the rigidity of a differential equation system represented we should use special numerical method that ensures precision and stability in its integration. To achieve this whit a classical explicit method it is required a lot of computing time, besides the possible instability. When using the semi implicit method, we developed a computer package using language C++ to solve the system of nonlinear ordinary differential equation. Solving the matrix system with the method mentioned above, it is found the numerical value of the concentration of the five chemical species for every time step, given the initial concentration. The computer program used to solve the system of

differential equation was developed in UNAM.

1 INTRODUCTION

All The core part in mathematical simulation of reaction mechanism of volatile organic compounds contained in highly polluted atmospheres, it is the solution of the set of ordinary differential equations describing the time evolution of the concentration of the chemical species. [1] The differential equations that result from chemical kinetics feature called rigidity, mainly due to the simultaneous presence of radical free with life time very short and of species that remain nearly constant as it is the content of oxygen, carbon dioxide and nitrogen.

Proposed mechanisms of photochemical smog formation reaction by different organic compounds, it was to make the computer simulation of the chemical kinetics from different initial conditions, in order to see the change of the concentrations of each species at the time.

To perform the mathematical simulation [2] of the different mechanisms is considered a reactor volume and temperature constant, where only carry out reactions indicated for each mechanism.

The change of concentration with focus in the time of a sort it is represented mathematically by means of an ordinary differential equation. In the cases studied, the reaction mechanism can be represented as a system of nonlinear ordinary differential equations, in the following steps:

$$\frac{dC_1}{dt} = f_1(C_1 \ldots\ldots\ldots C_n)$$

$$\frac{dC_2}{dt} = f_2(C_1 \ldots\ldots\ldots C_n)$$

•

•

•

$$\frac{dCn}{dt} = f_n(C_1 \ldots\ldots\ldots C_n)$$

Where C_1, $C_2 \ldots C_n$ represent the concentrations of each of the n chemical species.

This type of system is known as autonomous, due to the independent variable time does not appear explicitly in the functions $f_1, f_2, \ldots, f_n$.

The main problem to realize the integration of systems of differential equations that represent of atmospheric chemistry reaction mechanisms, it is the wide variation of the kinetic constants, which brings as consequence the instability by applying an explicit numerical method as for a change in the independent variable, the dependent variables change at very different scales. When this happens, it is said that the system of differential equations is a rigid system.

The matrix $\boldsymbol{J}$, which contains the partial derivatives of functions $\boldsymbol{f_i}$, with respect to each of the chemical species, concentration plays an important role in the application of numerical methods to rigid systems. The determinant of the matrix $\boldsymbol{J}$ is known as Jacobean. The rigid system is characterized when the matrix $\boldsymbol{J}$ has widely disparate characteristic values.

More stringent, a rigid system occurs when:

• The real parts of the eigenvalues of the matrix ***J*** is less than zero.

• The quotient that results from dividing the real part of the maximum characteristic value, between the real parts of the smallest characteristic value is much greater than one.

A special numerical method that ensures accuracy and stability in their integration should be used to solve the problem that represents the rigidity of a system of differential equations. Long time is required to achieve this with an explicit classical method, computer, apart from the possible instability.

This work used the semi-implicit method, which allows to solve the above mentioned problem.

If we consider as an example the following differential equations with their respective initial conditions:

$$y' = f_1(x, y, z);\ y(x_0) = y_0$$
$$z' = f_2(x, y, z);\ z(x_0) = z_0$$

It is denoted as u the numerical value of y as the numerical value of z v, then using the semi-implicit method, differential equations are represented in the following way:

$$\frac{u_{n+1} - u_n}{h} = \frac{1}{2}[f_1(x_{n+1}, u_{n+1}, v_{n+1}) + f_1(x_n, u_n, v_n)]$$
$$\frac{v_{n+1} - v_n}{h} = \frac{1}{2}[f_2(x_{n+1}, u_{n+1}, v_{n+1}) + f_2(x_n, u_n, v_n)]$$

Where u_0 and v_0 is the initial condition y_0 and z_0.

Applying the formula of Taylor to functions on the right side, are the following linear equations with respect to u_{n+1} and v_{n+1}:

$$u_{n+1} - u_n = \frac{h}{2}[f_1(B) + (u_{n+1} - u_n)\frac{\partial f_1(B)}{\partial y} + (v_{n+1} - v_n)\frac{\partial f_1(B)}{\partial z} + \ldots + f_1(I)]$$
$$v_{n+1} - v_n = \frac{h}{2}[f_2(B) + (u_{n+1} - u_n)\frac{\partial f_2(B)}{\partial y} + (v_{n+1} - v_n)\frac{\partial f_2(B)}{\partial z} + \ldots + f_2(I)]$$

where B is (x_{n+1}, u_n, v_n) e $I\ (x_n, u_n, v_n)$.

If defined:

$$u^* = u_{n+1} - u_n$$

Then:

$$u^* = \frac{h}{2}f_1(B) + \frac{h}{2}u^*\frac{\partial f_1(B)}{\partial y} + \frac{h}{2}v^*\frac{\partial f_1(B)}{\partial z} + \frac{h}{2}f_1(I)$$
$$v^* = \frac{h}{2}f_2(B) + \frac{h}{2}u^*\frac{\partial f_2(B)}{\partial y} + \frac{h}{2}v^*\frac{\partial f_2(B)}{\partial z} + \frac{h}{2}f_2(I)$$

Fixing these equations in matrix form, it is:

$$\begin{pmatrix} u^* \\ v^* \end{pmatrix} = \begin{pmatrix} \frac{h}{2}f_1(B) \\ \frac{h}{2}f_2(B) \end{pmatrix} + \begin{pmatrix} \frac{h}{2}f_1(I) \\ \frac{h}{2}f_2(I) \end{pmatrix} + \begin{pmatrix} \frac{h}{2}\frac{\partial f_1(B)}{\partial y} & \frac{h}{2}\frac{\partial f_1(B)}{\partial z} \\ \frac{h}{2}\frac{\partial f_2(B)}{\partial y} & \frac{h}{2}\frac{\partial f_2(B)}{\partial z} \end{pmatrix}\begin{pmatrix} u^* \\ v^* \end{pmatrix}$$

Then:

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} u^* \\ v^* \end{pmatrix} - \begin{pmatrix} \frac{h}{2}\frac{\partial f_1(B)}{\partial y} & \frac{h}{2}\frac{\partial f_1(B)}{\partial z} \\ \frac{h}{2}\frac{\partial f_2(B)}{\partial y} & \frac{h}{2}\frac{\partial f_2(B)}{\partial z} \end{pmatrix}\begin{pmatrix} u^* \\ v^* \end{pmatrix} = \begin{pmatrix} \frac{h}{2}f_1(B) + \frac{h}{2}f_1(I) \\ \frac{h}{2}f_2(B) + \frac{h}{2}f_2(I) \end{pmatrix}$$

Therefore:

$$\begin{pmatrix} 1 - \frac{h}{2}\frac{\partial f_1(B)}{\partial y} & -\frac{h}{2}\frac{\partial f_1(B)}{\partial z} \\ -\frac{h}{2}\frac{\partial f_2(B)}{\partial y} & 1 - \frac{h}{2}\frac{\partial f_2(B)}{\partial z} \end{pmatrix}\begin{pmatrix} u^* \\ v^* \end{pmatrix} = \begin{pmatrix} \frac{h}{2}f_1(B) + \frac{h}{2}f_1(I) \\ \frac{h}{2}f_2(B) + \frac{h}{2}f_2(I) \end{pmatrix}$$

Solving this system of two simultaneous linear equations are the values of u^* and v^*, and therefore the values of $v_{n+1}.1$ and v_{n+1}.

To use the semi-implicit method a computer package in c++ was developed to solve systems of nonlinear ordinary differential equations.

2 REACTION MECHANISM

To illustrate the method used semi-implicit and simulate the chemical kinetics of the reaction mechanism proposed mathematical modeling of formaldehyde (HCHO) considered is shown as you pollute the main secondary highly polluted atmospheres.

In this case indicates the mechanism of reaction of formaldehyde [3,4] where it can react with in the presence of light or the hydroxyl radical release hydro peroxide, carbon monoxide and water respectively.

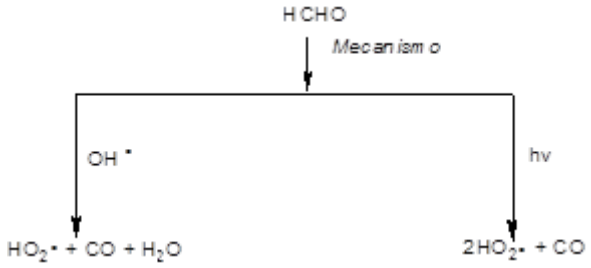


Figure 1. Mechanism of reaction of formaldehyde.

In the proposal chemical kinetics are considered chain reactions, as initial reactions of the cycle [5] basic photochemical smog, indicated in the following three first reactions and their respective kinetic constants.

$NO_2 + h\gamma \longrightarrow NO + O$	$k_1=0.533$
$O + O_2 + M \longrightarrow O_3 + M$	$k_2=2.183X10^{-5}$
$NO + O_3 \longrightarrow NO_2 + O_2$	$k_3=26.59$

Propagation reactions correspond to reaction mechanism shown above in Figure 1.

$HCHO + h\gamma \xrightarrow{2O_2} 2HO_2^{\bullet} + CO$ $\quad k_4=2.35x10^{-3}$

$HCHO + HO^{\bullet} \xrightarrow{O_2} HO_2^{\bullet} + CO + H_2O$ $\quad k_5=1.6x10^{4}$

And the corresponding termination reactions of the hydroxyl and hydro peroxide to form nitrogen dioxide and nitric acid respectively.

$HO_2^{\bullet} + NO \longrightarrow {}^{\bullet}OH + NO_2$ $\quad k_6=1.214x10^{4}$

${}^{\bullet}OH + NO_2 \longrightarrow HNO_3$ $\quad k_7=I.613x10^{4}$

The different chemical species formed are the following nine: C_1=NO, C_2=NO_2, C_3=O_3, C_4=O, C_5=HCHO, C_6=${}^{\bullet}$OH, C_7=$HO_2^{\bullet}$, C_8=HNO_3 y C_9=CO.

3 MATHEMATICAL SIMULATION

The change of concentration of chemical species focus in the time it is given by the following functions or differential equations:

```
f_r[1]:=k1*c[2]-k3*c[1]*c[3]-k6*c[1]*c[7];
f_r[2]:=-k1*c[2]+k3*c[3]*c[1]+k6*c[1]*c[7]-k7*c[2]*c[6];
f_r[3]:=k2*c[4]*C_O2*C_M-k3*c[3]*c[1];
f_r[4]:=k1*c[2]-k2*c[4]*C_O2*C_M;
f_r[5]:=-k4*c[5]-k5*c[5]*c[6];
f_r[6]:=-k5*c[5]*c[6]+k6*c[1]*c[7]-k7*c[2]*c[6];
f_r[7]:=2*k4*c[5]+k5*c[5]*c[6]-k6*c[1]*c[7];
f_r[8]:=k7*c[2]*c[6];
f_r[9]:=k4*c[5]+k5*c[5]*c[6];
```

While the partial derivatives with the concentration of each of the nine species in each of the seven reactions and chemical, it is given by the set of the following equations

```
Df_Dy[1,1]:=-k3*c[3]-k6*c[7];
Df_Dy[1,2]:=k1;
Df_Dy[1,3]:=-k3*c[1];
Df_Dy[1,7]:=-k6*c[1];

Df_Dy[2,1]:=k3*c[3]+k6*c[7];
Df_Dy[2,2]:=-k1-k7*c[6];
Df_Dy[2,3]:=k3*c[1];
Df_Dy[2,6]:=-k7*c[2];
Df_Dy[2,7]:=k6*c[1];

Df_Dy[3,1]:=-k3*c[3];
Df_Dy[3,3]:=-k3*c[1];
Df_Dy[3,4]:=k2*C_O2*C_M;
```

```
Df_Dy[4,2]:=k1;
Df_Dy[4,4]:=-k2*C_O2*C_M;

Df_Dy[5,5]:=-k4-k5*c[6];
Df_Dy[5,6]:=-k5*c[5];

Df_Dy[6,1]:=k6*c[7];
Df_Dy[6,2]:=-k7*c[6];
Df_Dy[6,5]:=-k5*c[6];
Df_Dy[6,6]:=-k5*c[5]-k7*c[2];
Df_Dy[6,7]:=k6*c[1];

Df_Dy[7,1]:=-k6*c[7];
Df_Dy[7,5]:=2*k4+k5*c[6];
Df_Dy[7,6]:=k5*c[5];
Df_Dy[7,7]:=-k6*c[1];

Df_Dy[8,2]:=k7*c[6];
Df_Dy[8,6]:=k7*c[2];

Df_Dy[9,5]:=k4+k5*c[6];
Df_Dy[9,6]:=k5*c[5];
```

Where is the matrix resulting Jacobean J as follows:

$-K_3C_3-K_6C_7$	K_1	$-K_3C_1$	0	0	0	$-K_6C_1$	0	0
$K_3C_3+K_6C_7$	$-K_1-K_7C_6$	K_3C_1	0	0	$-K_7C_2$	K_6C_1	0	0
$-K_3C_3$	0	$-K_3C_1$	$K_2CO_2C_M$	0	0	0	0	0
0	K1	0	$-K_2CO_2C_M$	0	0	0	0	0
0	0	0	0	$-K_4-K_5C_6$	$-K_5C_5$	0	0	0
K_6C_7	$-K_7C_6$	0	0	$-K_5C_6$	$-K_5C_5-K_7C_2$	K_6C_1	0	0
$-K_6C_7$	0	0	0	$2K_4+K_5C_6$	K_5C_5	$-K_6C_1$	0	0
0	K_7C_6	0	0	0	K_7C_2	0	0	0
0	0	0	0	$K_4+K_5C_6$	K_5C_5	0	0	0

Solving the matrix system by the above mentioned method, you will find the numerical value of the concentrations of nine chemical species for each step in time, given the initial concentrations, specifically set the more stable compounds.

The computer program used to solve the system of differential equations, developed in the Center of research theory of the Faculty of studies Cuautitlán of UNAM, includes:

• The routine implementation of the Semi-implicit method for rigid systems of coupled nonlinear ordinary differential equations.

• The routine of Gauss elimination method for solving systems of linear algebraic equations

simultaneous.

• Graphing of results written in c++.

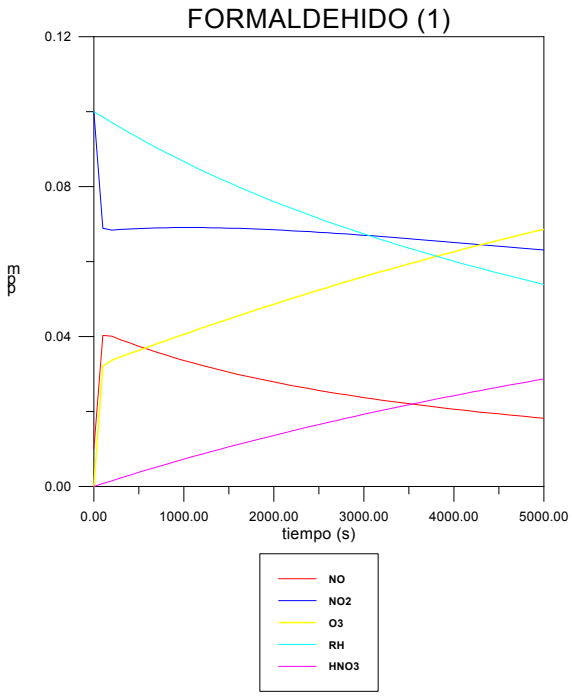
The first page must contain the Title, Author(s), Affiliation(s), Key words and the Summary. The Introduction must begin immediately below, following the format of this template.

4. ANALYSIS OF RESULTS

In the first run of the program (case 1), were assigned as initial conditions NOx concentrations closer to the atmosphere highly contaminated, i.e. 0.1 ppm NO_2 which can be found in greater proportion due to the oxidizing conditions of the environment, and 0.01 ppm of NO. The concentration for each hydrocarbon depicted in graphs as RH, was 0.1ppm, taking into account that the emissions allowed by Mexican legislation in total hydrocarbons (NOM 041 ECOL 93) emission standards [6] for the cars of prior to 1979 are 700 ppm, and from 1994 to date, is 200 ppm, and that on average will emit 400 compounds in different proportions.

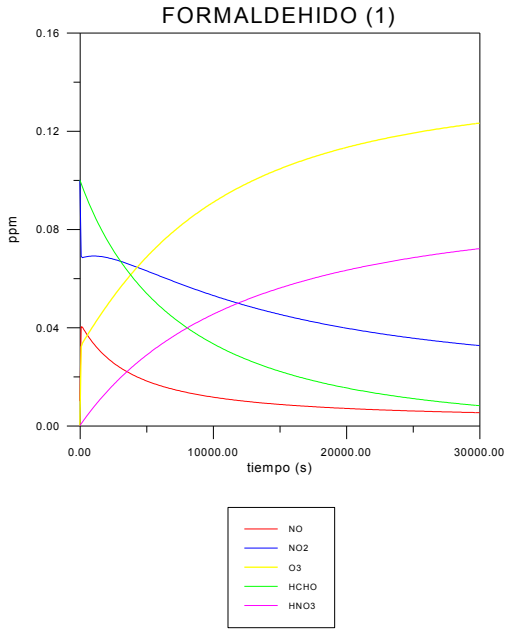
The time selected for the majority of cases was 5000 seconds, since during this time are observed changes in concentration of ozone and the emergence of other secondary pollutants, except in some cases where it is carried out only in the opening 100 seconds so it is 400 seconds, or in older times chart, as it is the case of acetone, which doubled the time.

Analyze the curves corresponding to the first run, you can see similar behavior in which the NO concentration increases, the NO_2 decreases and starts the formation of ozone reaching a maximum concentration of 0.072 ppm, while nitric acid reaches a concentration 0.0325 ppm, consumption of formaldehyde, the increase of the concentration of NO maximum to 0.0419 ppm and subsequent consumption, reduction of NO_2 to 0.0674 average ppm and subsequent production, formaldehyde is consumed rapidly, reducing its concentration to 0.059 ppm.



Graphic 1. Mathematical simulation of formaldehyde to 5000 seconds

The same initial conditions were considered in the second run (case 2) only change the response time with different times to observe the response of the program, managing to find that 30 000 seconds it is reached stability or equilibrium of reaction, reaching consume practically formaldehyde and not, to decrease to less than half the NO_2, ozone reaching concentrations de0.12 ppm and nitric acid of 0.0725 ppm.



Graphic 1. Mathematical simulation of formaldehyde to 30000 seconds

5 CONCLUSIONS

The proposed computational program responds to the solution of differential equations, derived partial and Jacobean of rigid systems, product of the mathematical simulation of chemical kinetics of volatile organic compounds in highly polluted atmospheres, this program can be applied to a myriad of possibilities [7] where you can change. The type of compound, the reaction mechanism, the concentration of the chemical species or reaction time.

In the event proposed for this presentation that was formaldehyde, chemically explains the high reactivity and justifies the high concentration of pollutants side such as ozone which is measured in different cities to evaluate air pollution and nitric acid which can produce acid rain.

REFERENCES

[1] McRae, G.J., Goodin, W.R., Seinfeld, J.H. *Numerical solution of the atmospheric diffusion equation for chemical reacting flows*. Journal of Computational Physics. 45. 1-42, (1982).

[2] León R. F. Ma., *Análisis Comparativo de los Mecanismos de Reacción y Cinética Química de Compuestos Orgánicos Volátiles para la Formación de Smog Fotoquímico en la Ciudad de México*, Universidad Politécnica de Madrid, Tesis, (1998).

[3] Atkinson, R., Baulch, D.L., Cox, R.A., Crutzen, P.J., Hampson Jr., R.F., Kerr, J.A. (Chairman), Troe, J., *Evaluated Kinetic and Photochemical Data for Atmospheric Chemistry: Supplement III*, IUPAC Subcommittee on Gas Kinetic Data Evaluation for Atmospheric Chemistry, Vol. 18, No. 2, pp. 881-1097, (1989).

[4] Baulch, D.L., Cox, R.A., Crutzen, P.J., Hampson Jr., R.F., Kerr, J.A. (Chairman), Troe, J., Watson, R.T., *Evaluated Kinetic and Photochemical Data for Atmospheric Chemistry: Supplement II, CODATA* Task Group on Gas Phase Chemical Kinetics, Vol. 13, No. 4, pp. 1259-1380, (1984).

[5] Seinfeld, John H. ; Pandis, Spyros N, *Atmospheric Chemistry and Physics - From Air Pollution to Climate Change* (2nd Edition) John Wiley & Sons

[6] Norma Oficial Mexicana NOM-041-ECOL/93 (1993). *Contaminantes de vehículos de gasolina.* Diario Oficial de la Federación (Ed.), (2008).

[7] León Rodríguez Frida María, Aguilar Márquez Armando, Axotla García Juan Carlos, Pineda Becerril Miguel de Nazareth. *Effect Ketones in the Butane Reactivity in the Formation of de Ozone in the MZMC (Metropolitan Zone in Mexico City*), International Symposium on Environmental Science and Technology, Donghua University, Shanghai, China, (2009).

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis(Eds)

STRUCTURAL OPTIMIZATION OF A WING BODY WITH UNCERTAIN AERODYNAMIC LOADS

DIMITRIOS I. PAPADIMITRIOU AND COSTAS PAPADIMITRIOU

Department of Mechanical Engineering, University of Thessaly
Pedion Areos, 38334, Volos, Greece
e-mail: dpapadim@uth.gr, costasp@uth.gr

Key words: Uncertainty Propagation, Robust Topology Optimization, Adjoint Method, Wing Body

Abstract. An algorithm for the structural optimization of wing body structures is presented in this paper. The structural analysis for the computation of displacements and stresses within the wing body is based on a finite volume solver. The topology optimization, based on the original methodology which includes the SIMP algorithm and the smoothing of computed fields using the solution of the Helmholtz equation is implemented. This is applied to solve the minimum compliance problem in a wing-like structure in order to compute the optimal distribution of material within the body that maximizes its stiffness constrained by a user defined volume. The study takes into account the uncertainties in the loading of the wing model and the optimization is robust since the minimization of the mean value and standard deviation of the compliance is sought with respect to these uncertainties using a sparse grid-based technique.

1 INTRODUCTION

Topology optimization [1] has become a very common means of structural optimization, based on the evolution of the optimal structure topology by adding or removing material from an initial structure. Similar to any optimization problem, the consideration of the uncertainties is very useful in topology optimization, as well. The uncertainties are usually related to the loading forces and the parameters of the structure. The robust optimization problem can be solved using several different algorithms such as stochastic or deterministic ones concerning both the estimation of the statistics of the objective function and their optimization. [2, 3, 4, 5, 3, 6, 7].

The robust topology optimization considering uncertainties in stiffness and loading parameters has been presented in [8, 9], whereas a small number of attempts related to reliability-based topology optimization can be found in the literature [10, 11].

A robust topology optimization for wing-like body structures is presented in this paper. The elasticity equations are solved using a finite-volume solver. The robust topology optimization of the structure includes uncertainties in the loading parameters and sparse grids [12, 13, 14] are used to evaluate the multi-dimensional integrals over the uncertain parameter space. The minimization of the mean value and standard deviation is accomplished using sensitivity analysis on these sparse grid points. The presented algorithm will be demonstrated to the robust topology optimization of a wing-like body structure.

2 TOPOLOGY OPTIMIZATION

2.1 Problem Definition

The topology optimization problem is defined by the minimization of the compliance function constrained by the total volume of the structure. The constrained optimization problem is defined as follows

$$\begin{aligned} min. \quad & F = \int_{\Omega} \bar{b}_i u_i d\Omega + \int_{S_1} \bar{\sigma}_i u_i dS \\ s.t. \quad & C = \int \rho V \leq V_f \\ and \quad & \frac{\partial \sigma_{ij}}{\partial x_j} + \bar{b}_i = 0 \quad , x \in \Omega \\ & \sigma_{ij} n_j = \bar{\sigma}_i \quad , x \in S_1 \\ & u_i = 0 \quad , x \in S_2 \end{aligned} \tag{1}$$

where ρ are the local densities, which are considered as the design variables and V the corresponding volumes. V_f is a given fraction of the total volume, $\bar{b}_i$ are the body forces, σ_{ij} are the stresses and ϵ_{ij} are the strains, which are expressed in terms of the displacements u_i through the equations

$$\sigma_{ij} = \mu \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) + \lambda \delta_{ij} \frac{\partial u_k}{\partial x_k} \tag{2}$$

and

$$\epsilon_{ij} = \frac{1}{2} \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) \tag{3}$$

respectively. Also, λ and μ are the Lamé parameters.

The compliance function F given by eq. 1 is totally equivalent to the strain energy function $G = \frac{1}{2} \int_{\Omega} \sigma_{ij} \epsilon_{ij} d\Omega$ which can be proved by the Gauss divergence theorem.

2.2 Self-Adjoint Topology Optimization

The discretized minimum compliance problem is formulated as, [1]

$$\begin{aligned} min. \quad & F = \mathbf{f}(\rho)\mathbf{u}(\rho) \\ s.t. \quad & C = \Sigma \rho_i V_i \leq V_f \\ and \quad & \mathbf{K}(\rho)\mathbf{u}(\rho) = \mathbf{f} \end{aligned} \tag{4}$$

where $\mathbf{u}$ and $\mathbf{f}$ are the displacement and force vectors and $\mathbf{K}$ is the stiffness matrix.

In order to find the derivatives of the objective function F with respect to the design variables ρ_i, the augmented objective function is formed by introducing the adjoint variable vector $\mathbf{\Psi}$, as follows

$$F_{aug} = F + \mathbf{\Psi}\left(\mathbf{f} - \mathbf{K}\mathbf{u}\right) \tag{5}$$

The sensitivities $\tilde{F}_{aug}$ of F_{aug} with respect to ρ_i yield

$$\begin{aligned} \tilde{F}_{aug} &= \mathbf{f}\tilde{\mathbf{u}} + \mathbf{\Psi}\left(-\tilde{\mathbf{K}}\mathbf{u} - \mathbf{K}\tilde{\mathbf{u}}\right) \\ &= \left(\mathbf{f} - \mathbf{K}\mathbf{\Psi}\right)\tilde{\mathbf{u}} - \mathbf{\Psi}\tilde{\mathbf{K}}\mathbf{u} \end{aligned} \tag{6}$$

Due also to symmetry, the adjoint equations, derived by eliminating the term depending on $\tilde{\mathbf{u}}$ yield

$$\mathbf{K}\mathbf{\Psi} = \mathbf{f} \tag{7}$$

proving the discrete minimum compliance problem is self-adjoint. The sensitivities $\tilde{F}_{aug}$ are finally given by

$$\tilde{F}_{aug} = -\mathbf{\Psi}\tilde{\mathbf{K}}\mathbf{u} = -\mathbf{u}\tilde{\mathbf{K}}\mathbf{u} \tag{8}$$

depending on the sensitivities of the stiffness matrix. The stiffness matrix is expressed in terms of the ρ parameters through the SIMP algorithm [1] by the expression

$$\mathbf{K}_i = \rho_i^p \bar{\mathbf{K}} \tag{9}$$

where $\bar{\mathbf{K}}$ are the stiffness matrices of the selected material and p is a user-defined parameter. The derivatives of $\mathbf{K_i}$ with respect to ρ_i are given by

$$\tilde{\mathbf{K}}_i = p\rho_i^{p-1}\bar{\mathbf{K}} \tag{10}$$

The computation of the objective function sensitivities does not require, thus, the solution of additional equations.

The optimization algorithm applied to minimize the compliance subject to the volumetric constraint is based on the bisection approach [15, 16]. Also, the smoothing of the design parameter and sensitivity derivative fields is achieved through the solution of the Helmholtz equations, as described in [17].

2.3 Robust Continuous Topology Optimization

The robust topology optimization is based on the minimization of the mean value and standard deviation of the compliance objective function with respect to uncertainties in the loading and model parameters. Thus the objective function becomes $F_{rob} = w\mu_F + (1-w)\sigma_F$ for user defined weight w.

The robust topology optimization is based on the computation of the mean value and standard deviation of the objective function, using the Gauss-Hermite quadrature on sparse grids, [13]. The PDF of the uncertain parameters is considered to be Gaussian and thus the Gauss-Hermite quadrature is required. the mean value of F or else the first statistical moment is computed by the expression

$$\mu_F = w_i F_i \tag{11}$$

where summation is implied for the index i, The expression for the second statistical moment is as follows

$$\mu_{2,F} = w_i F_i^2 \tag{12}$$

and the standard deviation is given by the expression

$$\sigma_F = \sqrt{\mu_{2,F} - \mu_F^2} \tag{13}$$

The first order sensitivities of the mean value and standard deviation with respect to the design parameters are given by

$$\frac{d\mu_F}{db_k} = w_i \frac{dF_i}{db_k} \tag{14}$$

and

$$\frac{d\sigma_F}{db_k} = \frac{\frac{d\mu_{2,F}}{db_k} - 2\mu_F \frac{d\mu_F}{db_k}}{\sigma_F} \tag{15}$$

where

$$\frac{d\mu_{2,F}}{db_k} = 2w_i F_i \frac{dF_i}{db_k} \tag{16}$$

3 APPLICATION

The robust topology optimization methodology is applied to the optimization of a 3D wing-like body structure. The distribution of the magnitude of the displacement vector is shown in Fig. 1.

Two distributed forces are acting on the body one at the top and one at the bottom surface, both facing towards the body domain. The constraint volume was set equal to half the volume of the structure. The two forces were modelled by independent Gaussian distributions with mean equal to their nominal values and standard deviation equal to 20% of their mean values. The optimal distribution of the density design parameters on

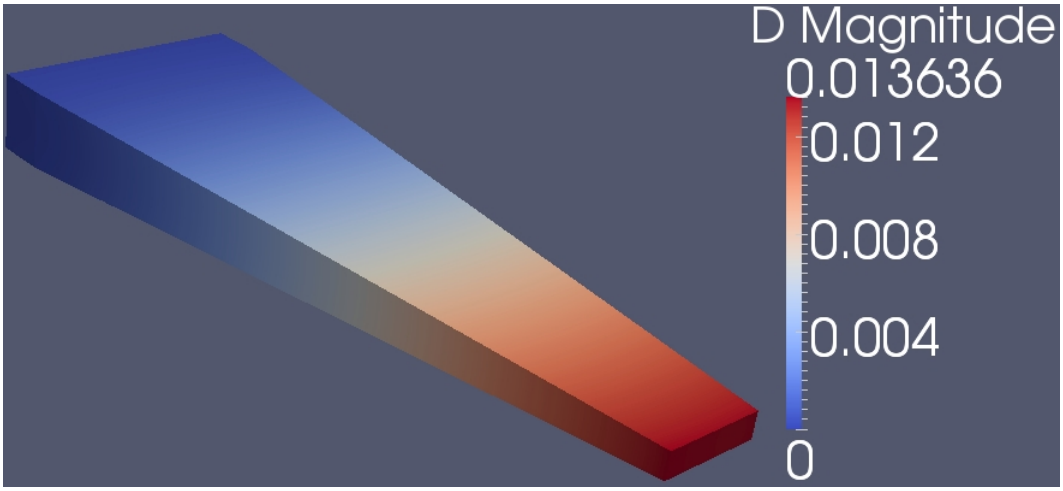


Figure 1: Wing-like body structure case. Distribution of the magnitude of the displacement vector.

the structure obtained by a single-point deterministic topology optimization is illustrated in Fig. 2.

The robust topology optimization of the wing-like body structure results to a distribution of the density variables depicted in Fig. 3 where the mean value of the compliance objective function was minimized with uncertainties regarded for the two values of the distributed forces.

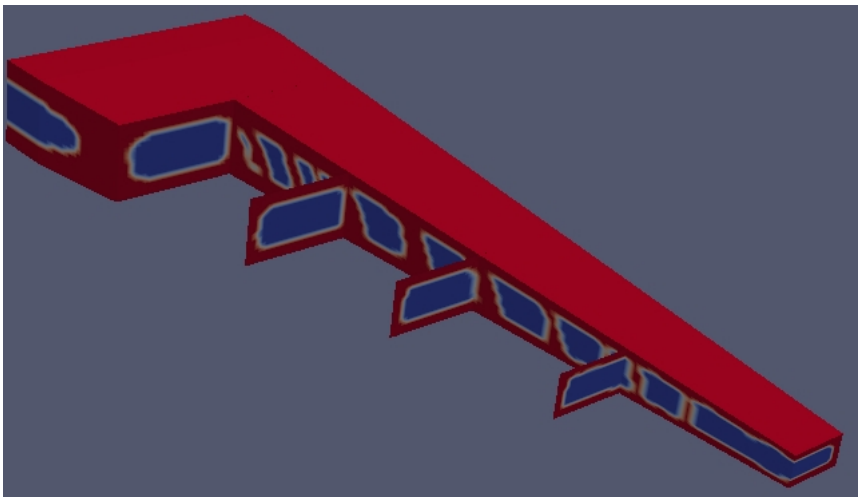


Figure 2: Wing-like body structure case. Plane cut view of the distribution of the optimal values of the density design parameters obtained by single-point topology optimization. The red colors correspond to the area which is solid, while the blue colors correspond to the area from which the solid has been removed.

The comparison of the single-point and the robust topology optimization of the wing-like structure can be seen in Figs. 4 and 5. In these figures, the convergence history of

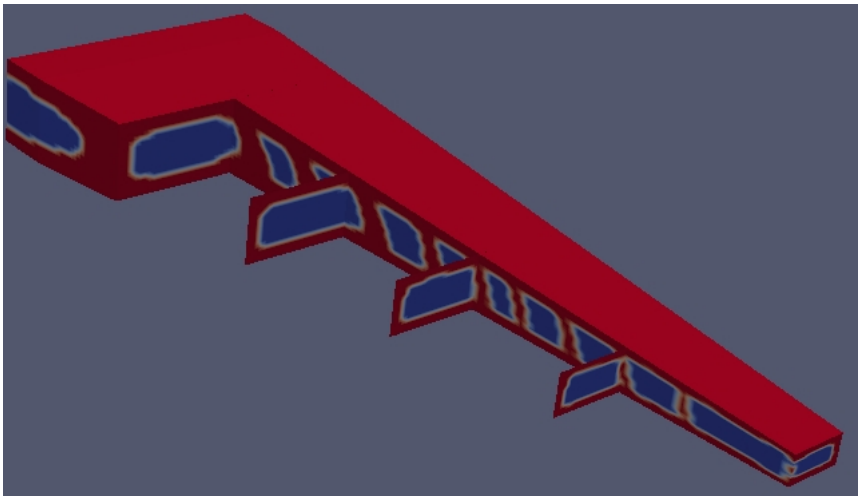


Figure 3: Wing-like body structure case. Plane cut view of the distribution of the optimal values of the density design parameters obtained by robust topology optimization for minimum mean value of the compliance function using a first-level sparse grid.

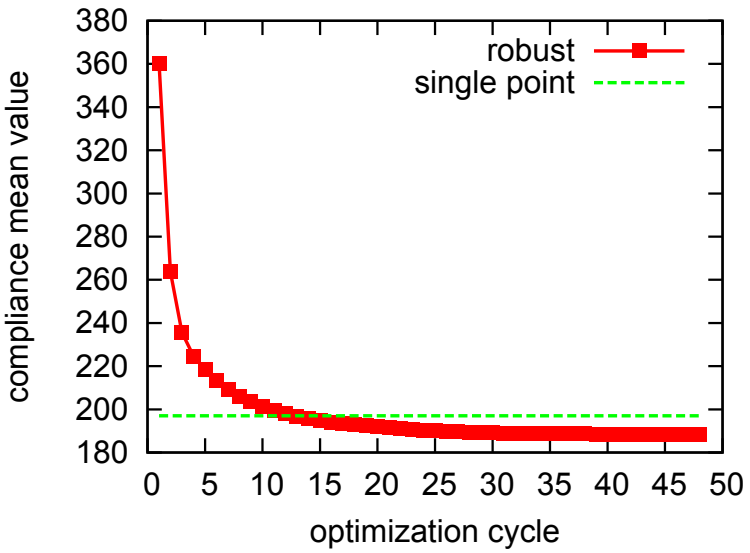


Figure 4: Wing-like body structure case. Convergence history of μ_F obtained by robust topology optimization (minimization of μ_F) compared with μ_F obtained by single point optimization (minimization of $F(\bar{c})$).

the mean value and standard deviation of the compliance function are shown. The two moments obtained by robust optimization are lower than the values that correspond to the single-point optimization derived design. This fact proves the importance and requirement of paying the additional cost of minimizing statistical moments of the objective function rather than the function for the nominal values of the uncertain parameters, in the case

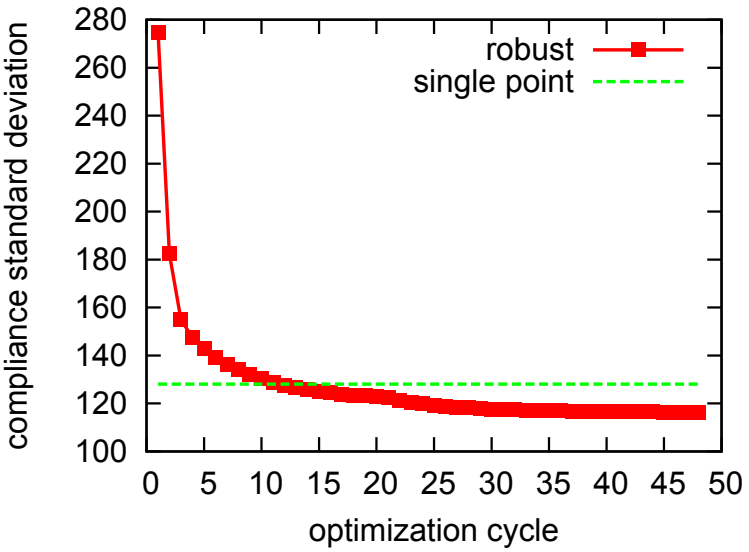


Figure 5: Wing-like body structure case. Convergence history of σ_F obtained by robust topology optimization (minimization of σ_F) compared with σ_F obtained by single point optimization (minimization of $F(\bar{c})$).

that the uncertainties are due.

4 CONCLUSIONS

A robust topology optimization algorithm based on sparse grid quadrature was applied to the structural optimization of a wing-like structure body with uncertainties concerning its loading parameters. The topology obtained through robust optimization was quite different than that obtained through single point deterministic topology optimization and proved to be more insensitive to variations in the uncertain quantities resulting to lower values of both the mean value and the standard deviation of the compliance objective function.

ACKNOWLEDGEMENTS

The research project is implemented within the framework of the Action "Supporting Postdoctoral Researchers" of the Operational Program "Education and Lifelong Learning" (Actions Beneficiary: General Secretariat for Research and Technology), and is co-financed by the European Social Fund (ESF) and the Greek State.

REFERENCES

[1] Bendsoe, M.P., Sigmund, O., Topology Optimization: Theory, Methods, and Application, Second ed. Springer, New York, 2003.

[2] Taguchi, G., Chowdhury, S., and Wu, Y., Taguchi's quality engineering handbook,

John Wiley & Sons, New York, 2004.

[3] Martinelli, M., and Duvigneau, R., On the use of second-order derivatives and metamodel-based Monte-Carlo for uncertainty estimation in aerodynamics, *Computers & Fluids* (2010) **39**(6):953–964.

[4] Kumar, A., Nair, P. B., Keane, A. J., Shahpar, S., Robust design using bayesian monte carlo, *International Journal for Numerical Methods in Engineering* (2008) **73**(11):1497–1517.

[5] Putko, M., Newman, P., III, A. T., and Green, L., Approach for uncertainty propagation and robust design in CFD using sensitivity derivatives, AIAA Paper, Vol. 25, No. 28, 2001.

[6] Papoutsis-Kiachagias, E., Papadimitriou, D., and Giannakoglou, K., Discrete and continuous adjoint methods in aerodynamic robust design problems, CFD and Optimization, An ECCOMAS Thematic Conference, 2011.

[7] Papadimitriou, D., and Giannakoglou, K., Third-order sensitivity analysis for robust aerodynamic design using continuous adjoint, *International Journal for Numerical Methods in Fluids* (2013) **71**(5):652–670.

[8] Asadpoure, A., Tootkaboni, M., Guest, J.K., Robust topology optimization of structures with uncertainties in stiffness - Application to truss structures, *Computers and Structures*)(2011) **89**:1131–1141.

[9] Guest, J.K., Igusa, T., Structural optimization under uncertain loads and nodal locations, *Computer Methods in Applied Mechanics and Engineering* (2008) **198**:116–124.

[10] Kharmanda, G., Olhoff, N., Mohamed, A., Lemaire , M., Reliability-based topology optimization, *Structural and Multidisciplinary Optimization* (2004) **26**(5):295–307.

[11] Jung, H.S. Cho, S., Reliability-based topology optimization of geometrically nonlinear structures with loading and material uncertainties, *Finite Elements in Analysis and Design* (2004) **41**(3):311–331.

[12] Zenger, C., Sparse grids, In Proceedings of the Research Workshop of the Israel Science Foundation on Multiscale Phenomenon, Modelling and Computation (p. 86) (1997).

[13] Gerstner T., Griebel M., Numerical integration using sparse grids, *Numerical Algorithms* (1998) **18**:209–232.

[14] Bungartz, H. J., Griebel, M., Sparse grids, *Acta numerica* (2004) **13**:147–269.

[15] Sigmund, O., A 99 line topology optimization code written in Matlab, *Structural and Multidisciplinary Optimization* (2001) **21**(2):120–127.

[16] Andreassen, E., Clausen, A., Schevenels, M., Lazarov, B. S., Sigmund, O., Efficient topology optimization in MATLAB using 88 lines of code, *Structural and Multidisciplinary Optimization* (2011) **43**(1):1–16.

[17] Lazarov, B. S., Sigmund, O., Filters in topology optimization based on Helmholtztype differential equations., *International Journal for Numerical Methods in Engineering* (2011) **86**(6), 765–781.

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis (Eds)

AN INTERCOMPONENT HEAT TRANSFER IN A GRAVITATIONAL FLOW MOVEMENTS OF PARTICLES IN AN INCLINED CHUTE

I. N. LOGACHEV*, K. I. LOGACHEV* O. A. AVERKOVA*

* Belgorod State Technological University named after V.G. Shukhov
(BSTU named after V.G. Shukhov), Kostykova str., 46, 308012 Belgorod, Russia
e-mail: kilogachev@mail.ru, web page: http://www.bstu.ru

Key words: Granular Materials, ejection of air, local exhaust ventilation, Applications.

Abstract. To ensure the complete localization of dust emissions, arising from the formation the overpressure in shelters, when heated conveyor overload materials is usually performed through suction of air from two shelters - from the upper (the upper shelter of the drive drum conveyor) and the lower (the lower shelter is place of loading the conveyor.) The main problem when designing of aspiration systems such congestions is to determine the required flow of air, sucked out of the shelters. We have developed a scientifically based method [1,2] which allows us not only to determine the required flow of air, removed from the shelter, but also to choose the rational allocation schemes of aspirating pipes, which is especially important in the cascade arrangement of equipment (feeder -rumble - crusher - conveyor), which covers are linked. The solution of this problem have been achieved in the use of the correct model of dual-velocity continuum "particulate matter - the air", which allowed at a modern level to estimate the dynamic interaction of the flow of air particles under overload of bulk material, as a free running particle jets, and as it moves through a closed overload troughs. The existence of combinations of these regimes is characterized by almost all industrial overload nodes. Intercomponent heat transfer modifies the picture of air movement in these nodes, which we have been examined in the gravitational movement of the particle flux in an inclined chute. The developed procedures of calculation of performance aspiration systems have proliferated in Russia and abroad in the mining and metallurgical industries.

1 INTRODUCTION

The inter-component heat and mass exchange plays a double role. On the one hand, an additional force, thermal pressure caused by buoyant forces, occurs in the chute. On the other hand, the mass exchange results in an additional source or outlet of the gaseous component.

2 INTER-COMPONENT HEAT EXCHANGE IN AN INCLINED CHUTE

Heat exchange, as well as the force interaction between the components, is defined by the flow pattern of the particles and the nature of their movement in the chute. An experimental study of heat exchange was carried out by means of a unit used to examine the inducing properties of an unheated particle flow (Fig. 1).

The value of the heat flow from the particles to air was determined using the enthalpy method:

$$Q = c_2 G_2 (t_к - t_н), \text{ W}, \quad (1)$$

where c_2 is the air heat capacity, J/kg·°; G_2 is the air mass flow rate, kg/s, $t_н$, $t_к$ are air inlet and outlet temperature, at the chute inlet and at its outlet, respectively, °C. The chute walls were heat sealed to prevent the heat exchange with surrounding air.

Research was made on crushed granite (mono fraction of 1.25-2.5 mm) and iron ore (poly fraction with d_{av} – 2.5 mm, the grain composition of which is shown in Table 3.2.) Being heated up to 200-300°C, the material was transferred through a heat-sealed chute with a section of 0.15 x 0.15 m at Θ = 45°, 60°, 75°. As shown by experimental studies, the rate of heat exchange varies with the relative velocity of the particles (Fig. 2a) and with their volume concentration (Fig. 2b), which is consistent with a generalization of heat exchange in dispersed through flows, as performed by Z.R. Gorbis [3]. The established behavior of the inter-component heat exchange for the case of an accelerated fall of particles under consideration was also confirmed by later experiments performed by A.S. Semenov [4] who studied the heat exchange between falling 10.5 mm steel balls and the air in a vertical chute with a section of 0.14 x 0.14 m.

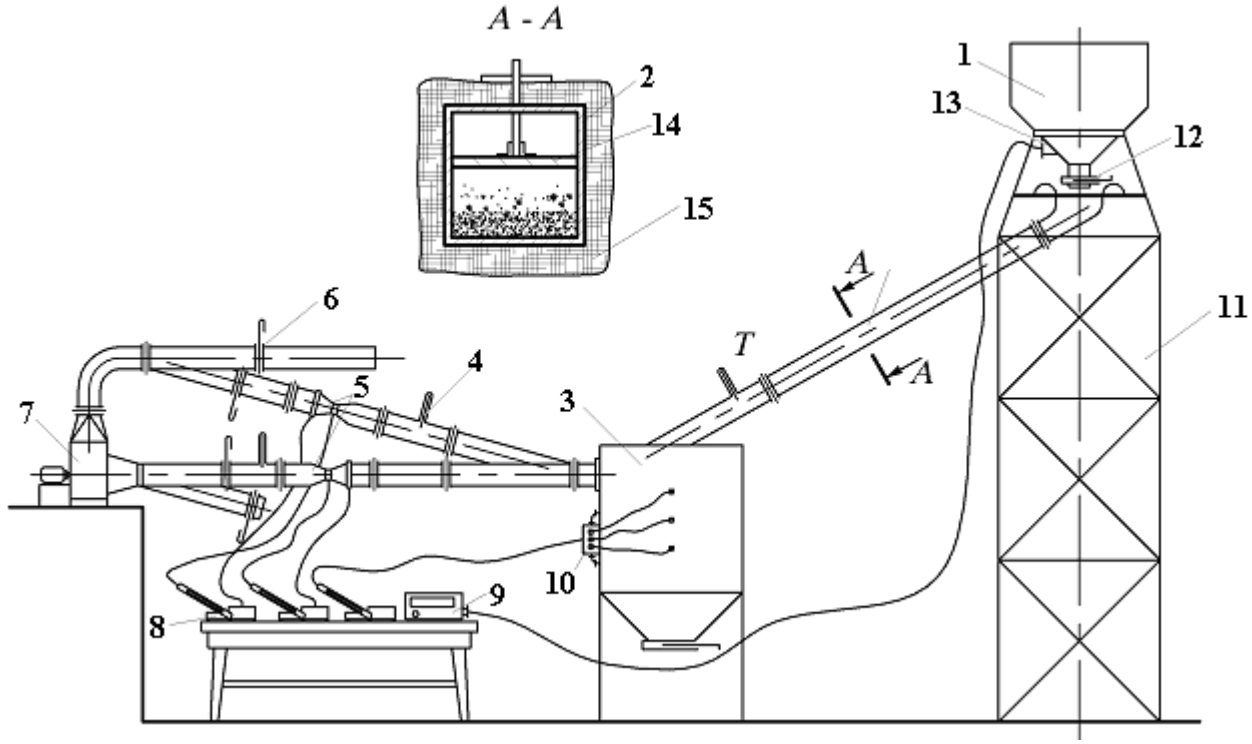


Figure 1: Diagram of the experimental arrangement for the study of injective properties of bulk materials: 1 – upper bin; 2 – chute; 3 – lower bin; 4 – thermometer; 5 – Venturi tube; 6 – damper; 7 – fan; 8 – micropressure gauge; 9 – galvanometer; 10 – blending chamber; 11 – metal frame; 12 – diaphragm; 13 – thermocouple; 14 – chute upper wall; 15 – heat insulation layer

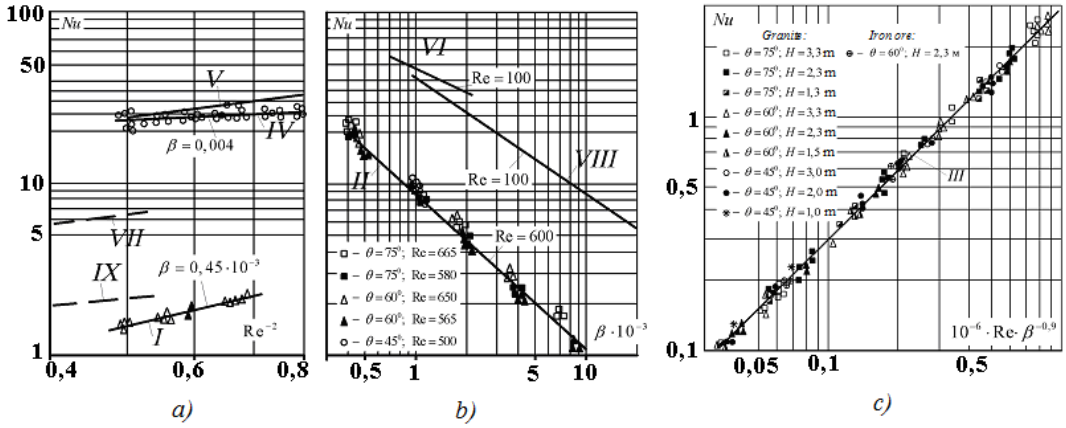


Figure 2: Influence of the number Re and volume concentration on the inter-component heat exchange in the fall of crushed granite particles in an inclined chute (I, II, III) and steel balls in a vertical chute (IV – according to A.S. Semenov), in the flow of free (V, VI – according to Z.R. Gorbis) and stagnant gas suspension (VII, VIII – according to Z.R. Gorbis; IX – according to Yu.N. Morozov)

However, quantitatively, the heat exchange in inclined chutes is significantly different from flows of free gas suspension and heat exchange in a vertical chute. Here, almost every particle participates in heat exchange, and its rate is much higher than in case of particles moving in an inclined chute, where most of them moves near the bottom in constraint conditions. Thus, in our case, we can speak of a conditional (apparent) heat exchange coefficient.

Here, the heat exchange process is analogous to the heat exchange in a mechanically stagnated gas suspension, where there can be seen "dead" zones – areas of a weak interaction with the air – in the flow of particles on the braking elements of mines. This can explain the near values of the Nusselt number (curves I, IX [5, 3]), as well as the coincident behavior of the heat exchange with the increase of the volume concentration (slope angle of lines II and VIII).

As a result of the statistical processing of the experimental data in the range $0.0002 < \beta < 0.01$; $400 < \mathrm{Re} < 700$, the following correlation was obtained [6]:

$$Nu = 2{,}95 \cdot 10^{-6}\ \mathrm{Re} \cdot \beta^{-0{,}9}\,, \tag{2}$$

which allows defining the inter-component heat exchange in an inclined chute. Here, the Nusselt and Reynolds numbers are expressed in terms of the average particle diameter and the average relative velocity along the chute length:

$$Nu = \frac{\alpha d}{\lambda_e}\,; \quad \mathrm{Re} = \frac{(v_{1av} - v_2)d}{\nu}\,; \quad v_{1av} = \frac{v_{1н} + v_{1k}}{2}\,,$$

where α is the heat exchange coefficient, W/m^2·°; λ_e is the air thermal conductivity, W/m·°.

3 THERMAL HEAD

As a result of the heat exchange, the air density in the chute is different from the density of the surrounding air, and its unit volume is influenced by Archimedes' buoyant force. Dynamic equation (8) for a prismatic chute is as follows (we suppose $\bar{v}_2 \approx const$, $\bar{\beta}_2 \approx 1$)

$$dp = -(\rho_0 - \rho_2)g_x dx - \lambda \frac{dx}{D} \cdot \frac{|v_2| v_2}{2} \rho_2 + dP_э . \tag{3}$$

Calculate the value

$$P_T = \int_0^l (\rho_0 - \rho_2) g_x dx , \tag{4}$$

usually called the thermal head. Express this value in terms of the chute height and the averaged density of the air

$$P_T = (\rho_0 - \bar{\rho}_2) gH , \tag{5}$$

where

$$\bar{\rho}_2 = \frac{1}{l} \int_0^l \rho_2 dx . \tag{6}$$

Open the symbol of averaging. To this end, express the air density in terms of temperature. Use the thermal expansion coefficient β_T defined by the equation

$$\beta_T = \frac{1}{\rho_2} \left(\frac{\partial \rho_2}{\partial T} \right)_P , \tag{7}$$

to obtain

$$\rho_2 = \rho_0 \exp[\beta_T (T_0 - T_2)], \tag{8}$$

where T_0, ρ_0 are the temperature (°K) and the density of the surrounding air (kg/m^3); T_2, ρ_2 are the temperature (°K) and density (kg/m^3) of the air in the chute.

To determine the temperature T_2, use heat-transfer equation (92) and the expression for the inter-component heat exchange (95) of Appendix 1. Assuming that the process is stable and ignoring pulse moments, this equation for a one-dimensional problem is as follows

$$d(c_2 \rho_2 T_2 v_2 S_{ж}) = \frac{\beta_1}{V_ч} S_ч \alpha (T_1 - T_2) S_{ж} dx . \tag{9}$$

In addition to the above assumptions, we assume that the temperature of the material is constant along the chute. This assumption is based on the fact that the material is not almost cooled in case of relatively low transfer heights due to a short stay in the chute (about 1 sec). Field measurements (Table 1) showed that the relative cooling does not exceed the accuracy of measurements and varies in the range of 1-3%.

In addition, average the volume concentration of the material, assuming that

$$\beta_1 \approx \beta = \frac{G_1}{\rho_1 S_{ж} v_{1av}} . \tag{10}$$

Integrate equation (9) taking into account accepted simplifications under the condition that $T_2 = T_o$ at the beginning of the chute (at $x = 0$) to obtain

$$T_2 = T_1 - (T_1 - T_0) \exp\left(-\frac{x}{l} W_\alpha \right), \tag{11}$$

where

$$W_\alpha = \beta \frac{S_ч}{V_ч} \alpha S_ж l \Big/ (c_2 \rho_2 v_2 S_ж). \tag{12}$$

Then, the expression for the air density in the chute is as follows

$$\rho_2 = \rho_0 \exp\left[-\beta_T (T_1 - T_2) \left(1 - e^{-\frac{x}{l} W_\alpha} \right) \right], \tag{13}$$

and the averaged density value is

$$\bar{\rho}_2 = \rho_0 \exp[\beta_T (T_1 - T_0)] \cdot \left\{ Ei[\beta_T (T_1 - T_0)] - Ei\left[\beta_T (T_1 - T_0) e^{-W_\alpha} \right] \right\} \Big/ W_\alpha , \tag{14}$$

where $Ei(f)$ is an exponential integral function with the argument f.

Insert this result for $\bar{\rho}_2$ into equation (5) to obtain

$$P_T = \left(\rho_0 - \frac{\rho_0 + \rho_{2k}}{2} \Pi \right) gH , \tag{15}$$

where Π is a correction coefficient equal to

$$\Pi = 2e^{-A} \left[Ei(A) - Ei(Ae^{-W_\alpha}) \right] \Big/ \left[W_\alpha \cdot (1 + \rho_{2k} / \rho_0) \right], \tag{16}$$

$$A = (\ln \rho_{2k} / \rho_0) \Big/ (1 - e^{-W_\alpha}) . \tag{17}$$

Table 1:Change in the temperature of the material and the vapor-air mixture in the chute in case of transfers of heated wet materials

Transfer group name	Flow of material G_1, kg/s	Drop height H, m	Chute cross-sectional area $S_ж$, m^2	Temperature of material, °C		Temperature of vapor-air mixture, °C		
				$t_{1н}$	t_{1k}	$t_{2н}$	t_{2k}	t_0
Transfer of burnt ore material from the drum cooler to the conveyor belt	10	5.5	0.2	78	77	25	50	7
Transfer of burnt ore material from a conveyor to a conveyor	200	6.0	0.4	65	63	20	45	12
Transfer of burnt ore material from a conveyor through an intermediate bin to another conveyor	200	12.0	0.4	62	60	20	40	10
Transfer of iron ore pellets from the drum cooler to a conveyor	8	3.0	0.2	73	70	30	35	6
Transfer of iron ore pellets from a conveyor to another conveyor	30	3.5	0.8	70	68	30	50	10

In the area $w_\alpha < 1$; $0.6 < \rho_{2k}/\rho_0 < 1$, the coefficient Π is almost equal to 1, and the value of the averaged air density in the chute is equal to the arithmetic mean value [7]. In the general case, the thermal head equals

$$P_T = \left(\rho_0 - \frac{\rho_{2н} + \rho_{2k}}{2} \right) gH \,. \tag{18}$$

Here, ρ_{2k} is the air density at the end of the chute at T_{2k} calculated taking into account the correlation obtained for the inter-component heat exchange (2) according to

$$T_{2k} = T_1 - (T_1 - T_{2н}) e^{-W_\alpha} \,. \tag{19}$$

4 AIR VELOCITY IN THE CHUTE

Integrate dynamic equation (3) to obtain

$$P_к - P_н = -P_T - \lambda \frac{l}{D} \frac{|v_2| v_2}{2} \rho_2 + P_э \tag{20}$$

or, expressing the pressure at the beginning and end of the chute in terms of the coefficients of local resistances

$$P_к = P_0 + \zeta_k \frac{|v_2| v_2}{2} \rho_2 \,, \; P_н = P_0 - \zeta_н \frac{|v_2| v_2}{2} \rho_2 \,, \tag{21}$$

equation (20) is as follows

$$\sum \zeta \frac{|v_2| v_2}{2} \rho_2 = P_Э - P_T \,. \tag{22}$$

Which shows that the difference between the induction head and thermal one determines the air flow and the direction of air flow in the chute, when transferring heated material. Three cases are possible in this regard.

Case 1: $P_Э > P_T$. The air moves downwards (forward flow). The value of the thermal head acts as an additional resistance. The volume of induced air is defined by an obvious equality:

$$Q_Э = \sqrt{(P_Э - P_T) / R_{ж}} \,. \tag{23}$$

Case 2: $P_Э < P_T$. Air moves toward the falling material (counter flow) under the prevailing thermal head. The induction head only slows the movement:

$$Q_{ж} = \sqrt{(P_T - P_Э) / R_{ж}} \,. \tag{24}$$

However, it should be borne in mind that the sum of the coefficients of local resistances will not generally be equal to the similar amount in case of the forward flow.

Case 3: $P_Э = P_T$. There is no direction of air movement in the chute. Only local aerodynamically unstable air circulations can occur in this case. Consider in detail the condition of the aerodynamic instability. Designate the temperature in the chute as T_{2cp} (note that in the limiting case $T_{2cp} \to T_1$). The air density is according to (8)

$$\rho_2 = \rho_0 \exp\left(\beta_T (T_0 - T_{2cp}) \right) \tag{25}$$

or, considering that in most practical cases $(T_2 - T_0)\,\beta_T << 1$,

$$\rho_2 = \rho_0\left[1-(T_{2cp}-T_0)\beta_T\right]. \tag{26}$$

The value of the thermal head in this case is

$$P_T = gH\rho_0(T_{2av}-T_0)\beta_T. \tag{27}$$

The value of the induction head

$$P_Э = \kappa_m\psi^*\frac{\rho_2}{\rho_1}G_1(v_{1k}^3 - v_{1H}^3)/(6a_T S_{ж}). \tag{28}$$

Then, the condition of equality of these heads takes the form of the following criterial equation

$$(1-n^3)\cdot \mathrm{Re}_k^2/(6Eu_0^*) = Gr, \tag{29}$$

where *Gr* is Grashof number, which characterizes the ascensional forces and equals

$$Gr = \beta_T\frac{gH^3}{\nu^2}(T_{2av}-T_0), \tag{30}$$

Re_k is Reynolds number, which characterizes the kinetic capacity of the particle flow at the end of the chute and equals

$$\mathrm{Re}_k = v_{1k}H/\nu, \tag{31}$$

Eu_0^* – is the modified Euler criterion which characterizes the aerodynamic drag strength of the particles and equals

$$\mathrm{Eu}_0^* = S_{ж}\frac{C_y^2}{2}\rho_0\Big/(G_1 v_{1k}). \tag{32}$$

The balance of the forces described by criterial equation (29) has been confirmed during an experiment involving the transfer of heated crushed granite (Fig. 3).

Thus, the amount of air being moved along the chute (with the local suction units operating) is, in the general case, equal to

$$Q_ж = \frac{P_Э - P_T + P_2 - P_1}{\sqrt{|P_Э - P_T + P_2 - P_1|/R_ж}}, \tag{2}$$

where P_1, P_2 are rarefactions kept up with the local suction units in the upper hood and the lower one (adjacent to the upper and lower ends of the chute, respectively), Pa.

Or in a dimensionless form

$$\varphi_k|\varphi_k| = Bu\cdot\left[|1-\varphi_k|^3 - |n-\varphi_k|^3\right]\Big/3 - Eu_T, \tag{34}$$

where

$$Eu_T = (P_T - P_2 + P_1)\Big/\left(\sum\zeta\frac{v_{1k}^2}{2}\rho_2\right). \tag{35}$$

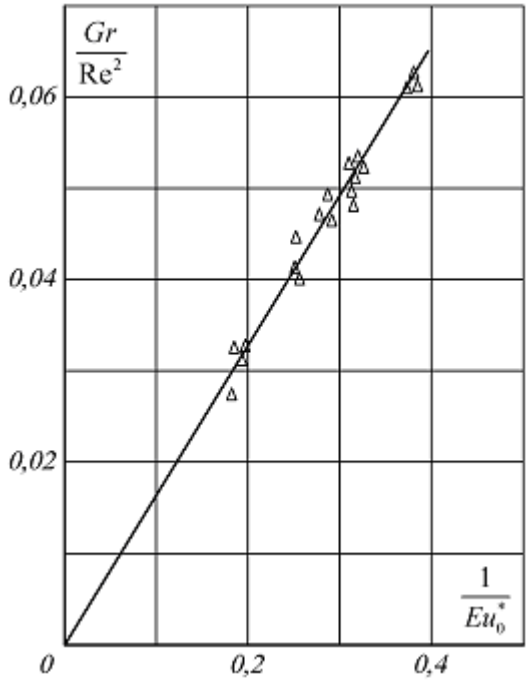


Figure: 3. The balance of the forces of induction and thermal pressures in the chute when transferring heated granite (d_{av}= 1.88 mm, $n \approx 0$)

The minus symbol before the value φ_k (or $Q_{ж}$) denotes a case of a counter flow, the case of balance occurs with

$$(1-n^3)Bu = 3Eu_T\,. \tag{36}$$

4 INFLUENCE OF MASS EXCHANGE ON THE VOLUMES OF INDUCED AIR

Consider the movement of the heated wet material accompanied with moisture evaporation from the surface of falling particles. Equations of mass exchange for a one-dimensional problem will be as follows

$$\frac{d}{dx}\beta_1\rho_1 v_1 S_{ж} = -J\cdot S_{ж};\ \frac{d}{dx}\beta_2\rho_2 v_2 S_{ж} = J\cdot S_{ж}\,, \tag{37}$$

where J is the volumetric evaporation rate, kg/(s·m^3). Respectively, for an impulse

$$\frac{d}{dx}\beta_1\rho_1 v_1 v_1 S_{ж} = S_{ж}\beta_1 a_T\rho_1 - S_{ж}\beta_1\kappa_m\psi^*\frac{|v_1-v_2|(v_1-v_2)}{2}\rho_2 - Jv_1 S_{ж}\,, \tag{38}$$

$$\frac{d}{dx}\beta_2\rho_2 v_2 v_2 S_{ж} = S_{ж}\beta_2 g_x(\rho_2-\rho_0) + S_{ж}\beta_1\kappa_m\psi^*\frac{|v_1-v_2|(v_1-v_2)}{2}\rho_2 - \frac{d}{dx}\beta_2 P_2 S_{ж} - \lambda\frac{S_{ж}\beta_2}{D}\frac{v_2^2}{2}\rho_2 + Jv_1 S_{ж}\,. \tag{39}$$

Assuming still that the volume concentration of the material is small ($\beta_1 << 1$; $\beta_2 \approx 1$), the last correlation with (36) can be written as ($S_{ж}$ - const):

$$\frac{dp}{dx} = g_x(\rho_2 - \rho_0) + \beta_1 \kappa_m \psi^* \frac{|v_1 - v_2|(v_1 - v_2)}{2} \rho_2 - \frac{\lambda}{D} \frac{v_2^2}{2} \rho_2 + J(v_1 - v_2). \tag{40}$$

Express the mass transport equation of the gaseous component in terms of the moisture content (m) and the flow rate of dry air ($G_в$)

$$J = \frac{G_в}{S_ж} \frac{dm}{dx}, \tag{41}$$

and write obvious equations of flow

$$\rho_2 v_2 S_ж = G_в(1+m) \tag{42}$$

and the averaged values of the velocities of the components

$$\overline{v}_1 = \frac{1}{l}\int_0^l v_1 dx = v_{1k} \cdot \frac{2}{3} \cdot \frac{1-n^3}{1-n^2}, \tag{43}$$

$$\overline{v}_2 = \frac{1}{l}\int_0^l v_2 dx \approx G_в(1+m_{ср}) / (\overline{\rho}_2 S_ж). \tag{44}$$

Assuming that the densities and velocities of the components on the right side of equation (40) are averaged, after integration on the condition that

$$P(0) = -P_1 - \zeta_н \frac{v_{2н}^2}{2} \rho_{2н}; \; P(l) = -P_2 + \zeta_k \frac{v_{2k}^2}{2} \rho_{2k} \tag{45}$$

the following equation is obtained

$$\sum \zeta^* \frac{\overline{v}_2^2}{2} \rho_2 = -P_T + P_Э + P_2 - P_1 + P_J, \tag{46}$$

where P_J is the pressure force that occurs due to moisture evaporation from the falling particles (for brevity, we call this value the interphase pressure), which is equal to

$$P_J = \int_0^l J(v_1 - v_2)dx \approx G_1(m_k - m_н)(\overline{v}_1 - \overline{v}_2) / S_ж, \tag{47}$$

$\Sigma\zeta^*$ is the sum of the coefficients of local resistance which is equal to

$$\sum \zeta^* = \zeta_н \left(\frac{1+m_н}{1+m_{ср}}\right)^2 \frac{\overline{\rho}_2}{\rho_{2н}} + \zeta_k \left(\frac{1+m_к}{1+m_{ср}}\right)^2 \frac{\overline{\rho}_2}{\rho_{2k}} + \lambda \frac{l}{D}. \tag{48}$$

In the dimensionless form, the equation can be as:

$$\overline{\varphi}^2 = Bu\left[|1-\overline{\varphi}|^3 - |n-\overline{\varphi}|^3\right] \Big/ 3 - Eu_T + Eu_J \left(\frac{2}{3}\frac{1-n^3}{1-n^2} - \overline{\varphi}\right), \tag{49}$$

where $Bu = \frac{\psi^* \kappa_m G_1 v_{1k}}{a_T S_ж \rho_1 \sum \zeta}$, Eu_T are numbers defined by correlation (35), and at $\Sigma\zeta \rightarrow \Sigma\zeta'$ and $\rho_2 \rightarrow \overline{\rho}_2$

$$Eu_J = G_в(m_k - m_н)\cdot v_{1k} \Big/ \left(S_ж \sum \zeta^* \frac{v_{1k}^2}{2} \bar{\rho}_2 \right). \tag{50}$$

From this, the mass flow rate of non-condensing (dry) part of the air can be obtained

$$G_в = \bar{\varphi}\cdot v_{1k} S_ж \frac{\bar{\rho}_2}{1+m_{av}}. \tag{51}$$

Then, the amount of vapor-air mixture transferring from the chute to the lower cavity (hood) can be obtained based on the obvious equality

$$G_{2k} = G_в(1+m_k) = \bar{\varphi} v_{1k} S_ж \bar{\rho}_2 \frac{1+m_k}{1+m_{av}}. \tag{52}$$

Thus, the amount of induced air during transfers of wet materials is increased not only due to the water vapors resulted from evaporation, but also due to additional forces of the interphase pressure.

12 CONCLUSIONS

It was established that due to non-uniformity of particles distribution the intensity of intercomponent heat exchange in a chute is by an order lower than the intensity of heat exchange in a vertical channel with a uniform distribution of particles. Counteracting forces resulting from the heat exchange of buoyancy forces contribute to a decrease in suction properties of a bulk material stream. The direction and value of the air velocity in a chute at transfer of a heated material is determined by the relation between heat and induction pressures. When transferring wet heated materials, the amount of air induced is increased due to water vapors resulting from condensation and auxiliary interphase pressure forces.

The reported study was partly supported by RFBR, research projects No. 14-41-08005r_ofi_m and 14-08-31069-mol_a

REFERENCES

[1] Logachev, I.N. and Logachev, K.I. Industrial Air *Quality And Ventilation: Controlling Dust Emissions*. CRC Press, (2014).
[2] Neykov, O.D. and Logachev, I.N. *Aspiration in the production of powder materials*. Moscow: Metallurgy (1973).
[3] Gorbis, Z.R. *Heat exchange and hydromechanics of dispersed through flows. 2nd revised and enlarged edition*. M: Energiya, (1970).
[4] Semenov, A. *Study of ventilation in chuting of hot materials. Abstract Thesis ... Cand.Sc. (Engineering)*. Sverdlovsk, (1975).
[5] Morozov, Yu.I. *Study of heat exchangers with a falling layer: Thesis ... Cand.Sc. (Engineering)*. Kiev, (1967).
[6] Logachev, I.N. Intercomponent heat exchange in a stream of a bulk material in closed

chute transfers. *Proceedings of the 4th Republic Interuniversity Conference for Evaporation, Combustion and Gas Dynamics of Dispersion Systems*. Odessa: OGU, (1969).

[7] Logachev, I.N. Thermal head in hot material transfers. *Air ventilation and treatment*. Moscow: Nedra (1970) **4**: 120-124.

Legend

a_t – acceleration of a stream of particles in a chute, m/s;
c – airborne speed of particles, m/s;
c_y – conventional airborne speed, m/s;
c_1 – heat capacity of material particles, J/(kg·K);
c_2 – air heat capacity (with ρ = const), J/(kg·K);
D – hydraulic diameter of a chute (channel), m;
d, $d_э$ – particle diameter (sphere diameter equivalent to a particle in terms of volume), m;
E – specific energy, J/kg;
G – mass flow (G_1 – particles, G_2 – air, $G_в$ – dry air), kg/s;
g – gravity factor (g_x – chute x-direction gravity factor), m/s^2;
H – drop height of particles, m;
$h = x = x/l_\infty$ – dimensionless drop height of particles;
I – intensity of interphase transformations, kg/(s·m^2);
k – particle drag coefficient ($k_г$, k_f, k_s – geometric, $k_д$ – dynamic);
k_m – particle frontal area/volume ratio, 1/м;
$L_э$, $Q_{э,к}$ – induced air flow in a chute, m^3/s;
l – chute length, m;
M – mass force (M_1 – particles, M_2 – air), N/kg
n – relation of the initial particle speed in a chute to the particle speed in the chute channel;
P – pressure ($P_э$ – chute injection pressure, $P_т$ – chute thermal pressure, P_a, P_0 – outside chute,
P_j – chute interphase pressure), Pa;
$P_ч$ – particle weight, N;
$Q_{жс}$ – chute air flow, m^3/s;
Q_{21} – air-to-particles heat exchange rate, W/m^3;
q – heat flow, W/m^2;
R – aerodynamic drag of bombarding particles, N;
$P_п$ – aerodynamic force of stream particle, N;
R, R_0 – aerodynamic force of single particle, N;
$R_{жс}$ – chute hydraulic characteristic, kg/m^3;
S – area of particles flow section, m^2;
S, $S_{жс}$ – cross sectional area of a chute (channel), m^2;
s – surface ($s_м$, s_4 – particles, $s_ш$ – sphere), m^2;
T – temperature, °K;
T_{2av} – mean air temperature in a chute, °K;

T_0 – average air temperature outside a chute, °K;
t, τ – time (τ_∞ – relaxation time), s;
V – volume ($V_ч$ – particle volume), m^3;
υ, v, ϑ – velocity (υ, υ_1 – particles; $\upsilon_{1к}$, $\upsilon_к$ – particles at the chute outlet; υ_{10}, $\upsilon_{1н}$ – particles at the chute inlet; υ_2, u – air), m/s;
$u_{вх}$ – exhaust pipe entry section air velocity, m/s;
$w = \upsilon - u$ – relative particle velocity, m/s;
w – material humidity, %;
x – path of particles over a chute, m;
α – interelement exchange ratio (α_m – mass, kg/(s·m^2·K); α_T, α – heat, W/(m^2·K);
β – volume concentration (β_1 – particles, β_2 – air), m^3/m^3;
$\beta_т$ – air thermal expansion coefficient, 1/°K;
ζ – local drag factor (LDF);
λ – hydraulic resistance coefficient;
$\lambda_г$ – air thermal conductivity, W/(m·K);
v – air kinematic viscosity coefficient, m^2/s;
$\vec{\Pi}$ – surface force vector, N/m^2;
Π_c – material particle constraint ratio, w/o unit of measurement;
$\Pi_д$ – dynamical interference activity factor, w/o unit of measurement;
ρ – density (ρ_1, ρ_m – particle material; ρ_2, ρ – particle stream air; ρ_0 – air outside a chute; ($\rho_{2н}$, $\rho_{2к}$ – air at the chute inlet and outlet), kg/m^3;
τ – time, s;
ψ – particle resistance coefficient (ψ_0 – particles in the area of self-similarity, $\psi_{0ш}$ – sphere in the area of self-similarity, ψ_c – airborne particles, ψ * – stream particles), w/o unit of measurement.

Criteria:
$Re = wd\rho/\eta$ – Reynolds number;
$Bu = \psi^* k_m G_1 \upsilon_{1k} / \left(\sum \zeta a_T S_{ж} \rho_1\right)$ – Butakov number;

$Eu = S_{ж} \frac{c_y}{2} \rho_0 / (G_1 v_{1k})$, $Eu^* = \Delta p / \left(0,5 \sum \zeta \upsilon_{1k}^2 \rho_2\right)$ – Euler number;

$Gr = \beta_т \frac{gH^3}{v^2} (T_{2av} - T_0)$ – Grashof number;

$Nu = \alpha d / \lambda_г$ – Nusselt number.

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis(Eds)

COUPLED THERMO- AERODYNAMICAL PROBLEMS IN DESIGN OF PROTECTION CLOTH

IRINA CHERUNOVA*, MAHESH DHONE† AND NIKOLAI KORNEV†

*Don State Technical University
Shevchenko str. 147, 346500 Shakhty, Russia
e-mail: i_sch@mail.ru, web page: http://www.sssu.ru/

†University of Rostock
A. Einstein Str.2, 18059 Rostock, Germany
e-mail: nikolai.kornev@uni-rostock.de - Web page: http://www.uni-rostock.de/

Key words: Human Thermodynamics, Conjugate Heat Transfer, Computational Fluid Dynamics, Cloth Design

Abstract. The paper presents results of calculation of thermodynamic interaction between the human body and the ambient air at very low temperatures. The numan body is clothed in a warm coverall. Temperature transport is calculated numerically by the solution of heat conduction equation. Simplified thermodynamic model of the human body by the surface heat flux obtained from empirical data is applied. Results of investigations are used for design of real protection cloth.

1 INTRODUCTION

Aerodynamic and thermodynamic interaction between body and environment under low temperature and wind conditions is a typical problem with thermomechanical coupling of fluids (air or water) and structures (human body and clothes). The body generates the heat within certain organs which is then transferred by the thermal diffusion through the body substance possessing very non uniform properties. A large fraction of the heat generated by internal organs is transported to the body periphery by a complicated net of blood vessels. The heat penetrates through the cloth and is transferred to surrounding medium by natural and forced convections. Wind causes big areas of the overpressure on the cloth surface which results in deformations and local change of the cloth thickness. In its turn, change of the local thickness leads to an alteration of heat conduction properties of the cloth. This means that the heat exchange between body and air is changed not only by intensification of convective heat transfer but also due to change of thermodynamic properties of cloth caused by wind induced deformations.

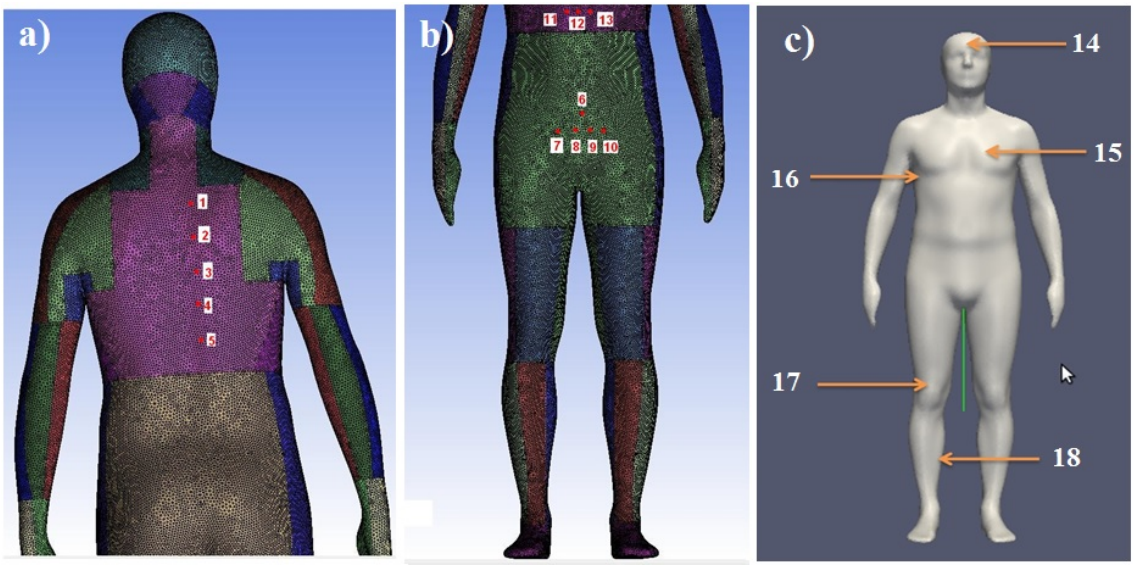


Figure 1: Points on the skin where the temperature is measured and calculated

The second effect which has still not been discussed thoroughly in the literature has been considered in our previous papers [1], [2] and [3]. Under strong wind conditions the heat transfer from the human body can sufficiently be increased due to change of the thermal conductivity caused by cloth deformation under wind induced pressures. For instance, at 10 m/s the heat increase could be of ten percent [2].

In this paper we restricted ourselves to the problem of interaction of body with surrounding medium through the heat diffusion and radiation. Since the wind influence has already been considered in our paper [2] it is not discussed here. Results of the present analysis are used for design of protection cloth with heating. To keep the human body temperature on the acceptable level, the local heating elements are embedded into the cloth textile. The power of this heating is calculated numerically. The results of the paper are utilized for the design of real protection cloth for the work under low temperatures and wind conditions in oil and gas industry.

2 MATHEMATICAL MODEL

2.1 Governing equation

The heat transport in the system "human body-cloth-ambient air" is described by the unsteady heat conduction equation

$$\rho C_p \frac{\partial T}{\partial t} = \lambda\left(\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} + \frac{\partial^2 T}{\partial z^2}\right) + f \tag{1}$$

where the specific heat capacity C_p, the heat conduction coefficient λ and the density ρ are changed depending on the material where the heat is conducted. f is the heat release

source. The steady state solution is obtained from time marching unsteady simulation. The equation (1) is solved with the Neumann boundary condition

$$\dot{q} = \dot{q}_{specified}, \text{ on the human body} \tag{2}$$

enforced on the body surface and the Dirichlet boundary condition

$$T = T_{air}, \text{ away from the human body} \tag{3}$$

enforced at the outer boundary of the computational domain.

2.2 Numerical implementation

The equation (1) is solved numerically using the OpenFoam toolkit [4]. The finite volume method is implemented on an unstructured grid with approximately 4.3 Mio of cells in the computational domain $15m \times 5.5m \times 5.5m$. Elongation of the computational domain in the longitudinal direction is because the same grid is applied for thermo-aerodynamic calculations which will be presented later. The grid has different resolutions in different areas. The finest unstructured grid is located in the block of $3m \times 1.8m \times 1.8m$ closest to the body. Away from the body the grid resolution is gradually reduced having relatively high resolution in the block $7.5m \times 3.6m \times 3.6m$. The real human body geometry of height of $1.8m$ designed at the Hohenstein Institute [5] on the basis of the detailed anthropological study of different categories of peoples is located in the symmetry plane at 5m from the inlet of the computational domain. For reliable resolution of the temperature boundary layers six thin prism layers are added on the inner and the outer surfaces of the cloth. A thin prism layer of the thickness of 1mm adjacent to the inner air layer between the cloth and the body is served as the heating layer with heat release source f.

2.3 Cloth construction

A cloth of thickness 30 mm comprises a savior, a flamestat cotton (upper sheet), an insulation thinsulate, and a taffeta as a lining. This cloth is used in oil industry for work at very low temperatures under oil contamination conditions. The air layer of the thickness of $10mm$ is located between the cloth and the body. The cloth textile has the following properties: the heat capacity is $C_p = 1800J/kgK$, the density is $8.57kg/m^3$, the thermal conductivity is $\lambda = 0.04W/mK$.

2.4 Simplified thermodynamic model of the human body by the surface heat flux obtained from empirical data

The heat release from the human body can be calculated using relatively simple models based on empirical information. The group of models proposed by Stolwijk and Fiala [6], [7] are based on the partial differential equations describing the heat transport in cross sections of body which are coupled through the blood transport along the body. Since

such models are not quite reliable we use a simplified model specifying the heat flux on the body surface $\dot{q}$. The heat flux $\dot{q}$ is calculated from the condition that the resulting temperatures from simulations are equal to these measured on the real human body clothed in the protection coverall. Temperature measurements were performed at -22 deg of the ambient air at 18 points shown in Fig. 1 directly on the skin of the human body. Results of calculations performed without wind speed are presented in the table 1. The body surface is subdivided into 34 elements (see Fig. 2). The heat flux on each element was varied as long as the temperature on the skin from simulations (second column in table 1) is approximately equal to the measured ones (first column in table 1). As follows from the fourth column in the table 1 the heat flux given in Fig. 2 provides the temperature distribution close to the measured one with the discrepancy less than five percent with the only exception at point 5. The total heat flux necessary to maintain prescribed temperatures is around $31W$. This flux doesn't take the radiation and sweating into account. Consideration of these additional effects results in the total heat flux about 70 Watt.

The data presented in Fig. 2 is some kind of thermodynamic model of the human body. To prove it we used an additional test. At the temperature of $+29^\circ$ of Celsius of ambient air the human body is in equilibrium with air and has the temperature of the skin around +29 deg. To escape generation of new grids we use the copper cloth instead of textile one. The copper possesses very high conductivity $\lambda = 401W/mK$ and acts as an ideal conductors of the heat from body as though there is no cloth. The temperature distribution obtained for this case is presented in the fifth column of the table 1 which is close to the equilibrium temperature 29°. The discrepancy is around five percent. Therefore the simplified model thermodynamic model of the human body by the surface heat flux obtained from empirical data can be considered as the reliable one.

3 RESULTS

The aim of the calculation is the determination of the heat release f which allows one to maintain the temperature on the body skin in the comfortable range at the temperature of the ambient air of -40°. The comfortable body skin temperatures are taken from the measurements performed at -22° with the real human clothed in the protection coverall (see column 2 in the table 1). In numerical simulations the area and the power were varied as long as the simulated temperature at 18 points shown in Fig.1 is close to the comfortable one. The heating elements were selected under the following conditions:

- Heating elements can not be located close to the sensible inner organs because it can have negative impact on health,
- Heating elements can not be located in zones of high sweating,
- The necessary area of the heating elements and heating power should be minimal,
- Both overheating and under heating of the body should be avoided.

Distribution of heating elements with necessary area and heat flux is given in Fig. 3. The total heat flux necessary to maintain the comfortable temperatures is estimated

Table 1: Temperature distribution at 18 points at −22 deg with cloth, without heating: measurement (second column), simulation (third column) and difference between measurement and simulation (fourth column) and at +29 deg without cloth: simulation (fifth column) and difference between measurement and simulation (last column).

Point	Experiment deg	Simulation deg	Error %	Simulation deg	Error %
Air temperature		−22 deg		+29 deg	
1	31.50	30.72	2.5	30.73	5.5
2	32.80	32.01	2.4	30.74	5.3
3	32.40	32.55	0.5	30.77	5.5
4	33.90	34.21	0 .9	30.82	5.4
5	32.20	34.41	6.9	30.88	5.8
6	31.40	32.05	2.1	30.71	5.4
7	32.80	33.06	0.8	30.73	5.3
8	32.40	32.38	0.1	30.72	5.3
9	33.80	32.32	4.4	30.72	5.1
10	32.10	32.71	1.9	30.72	5.4
11	32.90	32.70	0.6	30.71	5.2
12	32.60	33.43	2.5	30.71	5.2
13	34.00	32.90	3.2	30.71	5.0
14	29.67	29.37	1.0	29.28	0.9
15	31.20	31.23	0.1	29.75	2.4
16	34.82	34.85	0.1	30.20	3.4
17	23.49	23.52	0.1	29.73	3.1
18	26.75	26.63	0.4	29.65	2.4

as 12 Watt. Additionally about 30 Watt of the heat flux are necessary to cover the heat loss due to radiation. The final temperature distribution presented in the table 2 gives the averaged discrepancy of 0.44° and the biggest local discrepancy of 1.06° at the point 9. Figure 4 shows temperature distributions on the body skin at the ambient air temperatures of −22° and −40°. Obviously, the numerical simulations have a serious disadvantage in areas of contacts of different body parts, i.e. legs and hands. These artifacts of a pure numerical nature can be eliminated by increase of resolution in contact areas. Within this paper these regions are simply excluded from analysis.

4 CONCLUSIONS

The paper presents an application of numerical simulations for design of a special cloth for work under extremely low temperatures. The heating elements inside of the cloth are

Table 2: Temperature distribution at 18 points at −40 deg with cloth and heating: comfort temperature (second column), simulation (third column) and difference between the comfort temperature and simulation (fourth column).

Point	Comfortable T	Simulation	Error
1	31.5	30.72	0.78
2	32.8	33.70	0.90
3	32.4	33.01	0.61
4	33.9	33.27	0.63
5	32.2	31.65	0.55
6	31.4	31.34	0.06
7	32.8	32.67	0.13
8	32.4	32.25	0.15
9	33.8	32.74	1.06
10	32.1	31.98	0.12
11	32.9	33.11	0.21
12	32.6	32.54	0.06
13	34.0	34.46	0.46
15	31.20	30.54	0.66
16	34.82	34.55	0.27
17	23.49	23.27	0.22
18	26.75	26.20	0.55

applied to maintain the comfortable temperatures on the body skin. The distribution of the temperature on the body, in the cloth, in the air layer between the body and the cloth and in the ambient air is calculated from the temperature transport equation using the OpenFoam toolkit. Sizes and the power of heating elements are selected from numerical simulations. The results of numerical simulations seem to be reliable with exception in areas of contacts of different body parts. Results with account of full thermo- aerodynamic interaction will be presented in next works of the authors. At present the numerical results are used to construct the protecting coverall of the new generation for employees of the gas and oil industry.

Acknowledgments. This work was carried out with the financial support within the framework of the joint project 11032p/20253 of the German Federal Ministry of Education and Research (BMBF) and the Russian Foundation for Assistance to Small Innovative Enterprises (FASIE).

REFERENCES

[1] Cherunova I., Jacobi G. and Kornev N. Study of human body aerodynamics for protection cloth design. *Proceedings of the 6th International Conference on Vortex Flows and Vortex Models*, November 17-20, 2014, Nagoya, Japan.

[2] Cherunova I., Kornev N., Jacobi G., Treshchun I., Gross A., Turnow J., Schreier S. and Paschen M. Application of computational heat transfer and fluid mechanics for protection cloth design. *Journal of Engineering Physics and Thermophysics* (2014), **87(4)**:829–837. Vol. 87, No. 4, pp. 829-837.

[3] Cherunova I., Kornev N., Jacobi G., Treshchun I., Gross A., Turnow J., Schreier S. and Paschen M. Application of computational fluid mechanics for protection cloth design. *Proceedings of the 11th World Congress on Computational Mechanics (WCCM XI)*, 22.-25.07.2014, Barcelona, Spain.

[4] Weller, H.G., Tabor, G., Jasak, H. and Fureby C. A tensorial approach to CFD using object oriented techniques, *Computers in Physics* (1998), **12(6)**: 620–631.

[5] www.hohenstein.de

[6] Fiala, D., Lomas K.J., Stohrer M.. Computer prediction of human thermoregulatory and temperature responses to a wide range of environmental conditions. *Int. J. Biometeorol* (2001), **45**:143159.

[7] Cherunova, N. Kornev, and I. Brink. Mathematical model of the ice protection of a human body at high temperatures of surrounding medium, *Eng. Research* (2012), **76(3)**:97–103.

Body part	Area	Heat flux	Total heat flux
abdomen front	0.163	16.92	2.76
abdomen inferior	0.1175	15.83	1.86
abdomen posterior	0.161	15.83	2.55
face front	0.0132	26.57	0.35
face superior	0.0152	26.57	0.4
feet front	0.11	15.83	1.74
feet posterior	0.0358	15.83	0.57
hand inferior	0.0452	15.83	0.72
hand superior	0.0438	15.83	0.69
head-front	0.0061	26.57	0.16
head-posterior	0.0626	26.57	1.66
lower-arms-front	0.163	15.83	2.58
lower-arms-inferior	0.163	15.83	2.58
lower-arms-posterior	0.163	15.83	2.58
lower-arms superior	0.163	15.83	2.58
lower legs-front	0.185	17.37	3.21
lower legs inferior	0.185	17.37	3.21
lower legs posterior	0.185	17.37	3.21
lower legs superior	0.185	17.37	3.21
neck front	0.0303	15.83	0.48
neck posterior	0.0303	15.83	0.48
neck superior	0.0303	15.83	0.48
shoulder	0.0337	15.83	0.53
thorax front	0.115	18.1	2.08
thorax inferior	0.032	2.54	0.08
thorax posterior	0.104	15.83	1.65
upper arms front	0.149	15.83	2.36
upper arms posterior	0.149	15.83	2.36
upper arms superior	0.149	15.83	2.36
upper arms inferior	0.025	2.54	0.06
upper legs front	0.26	15.31	3.98
upper legs-inferior	0.26	15.31	3.98
upper legs-posterior	0.26	15.31	3.98
upper legs superior	0.26	15.31	3.98

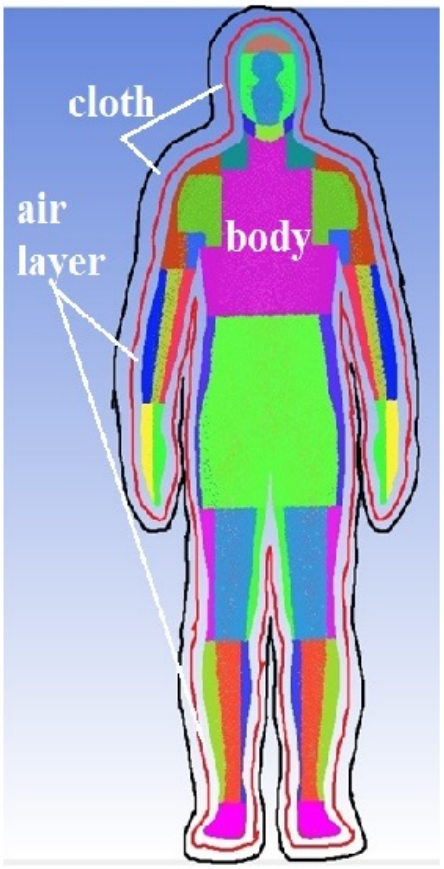


Figure 2: Distribution of the heat flux in W/m^2 (third column) and the total heat flux in Watt on different body parts which is necessary to get the temperature distribution obtained from measurement (see 2nd column in the table 1).

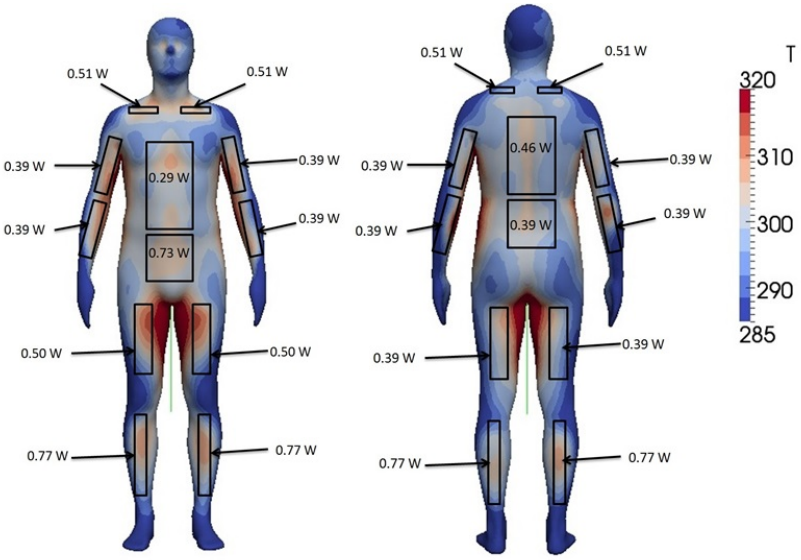


Figure 3: Distribution of heating elements and their power necessary to maintain the comfortable temperature at the temperature of the ambient air of $-40°$.

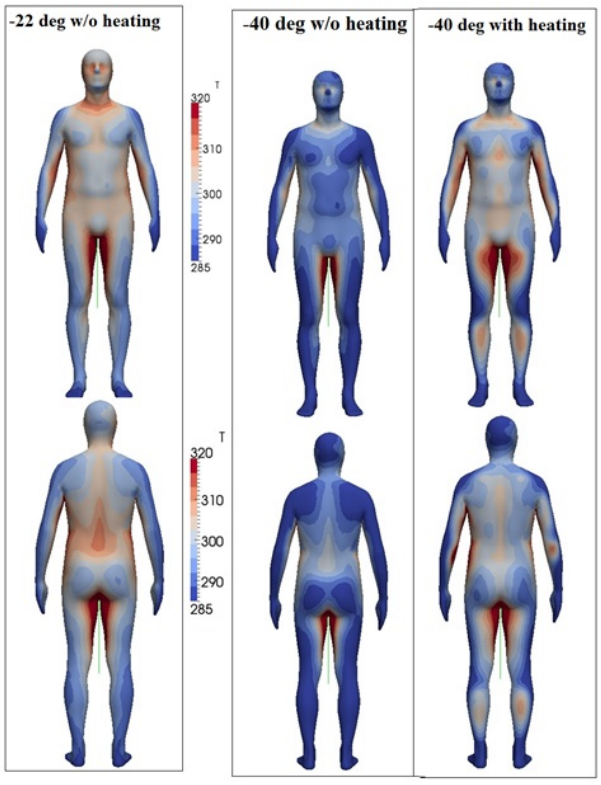


Figure 4: Temperature distribution on the body in cloth: left: $-22°$ without heating, middle: $-40°$ without heating, right: $-40°$ with heating.

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis (Eds)

RESPONSE OF RC COLUMNS WITH TRANSIENT CREEP IN A NATURAL FIRE ENVIRONMENT

A. SADAOUI[*], K. LATTARI[*] AND A. KHENNANE[†]

[*] Department of civil Engineering, Faculté de Génie de la Construction
University of Mouloud Mammeri, Tizi-ouzou, Hanaoua II, Algeria
e-mail: a_sadaoui@yahoo.com

[*] Department of civil Engineering, Faculté de Génie de la Construction
University of Mouloud Mammeri, Tizi-ouzou, Hanaoua II, Algeria
e-mail: lattari_kamel@hotmail.fr

[†] University of New South Wales (UNSW), Australian Defence Force Academy
PQ Box 7916, Canberra, Australia
email: a.khennane@adfa.edu.au

Key words: RC columns, Transient creep, Fire scenarios, High temperatures, Finite element analysis, Geometrical nonlinearity.

Abstract. The aim of this study is to investigate the effect of transient creep on the structural response of RC columns subjected to natural or parametric fires using a finite element model developed by the authors. The model, capable of analysing the response of RC columns from pre-fire stages to collapse in fire environment, is specially developed for the analysis of RC structures under severe thermo-mechanical loads, which accounts for transient creep explicitly as an additional component of the total strain of the concrete or implicitly through the use of the materials' properties recommended by EC2. Through the obtained results, it is shown that the transient creep phenomenon significantly influences the fire response of RC columns. It was also found that the conventional method based on standard fire exposure may not be conservative if the resulting fire has a decay phase similar to the severe fire scenario used in this work.

1 INTRODUCTION

The performance of a reinforced concrete (RC) structure in a fire environment depends particularly on the behaviour of the columns subjected to fire. These structural elements form the main mean of load transfer to the foundations because a failure of column is more detrimental to a building than that of a beam or a slab, and more likely to result in a progressive collapse [1,2,3]. Columns are subject to a combination of thermo-mechanical actions that arise from restrained thermal elongations, degradation of the mechanical properties of the constituents, and transient creep in the concrete. The transient creep, also usually referred to as transitional thermal creep or load induced creep, is unique to concrete and develops only during first time heating of the concrete under compressive stresses [4,5,6].

Temperature increase in fire conditions is well known to decrease the load bearing capacity of concrete and to increase its deformability. The structural and chemical changes taking place in the material, the internal stresses caused by the temperature gradients, as well as the high pore pressures caused by the evaporation of pore water, combine to develop internal microcracks or damage within the concrete [7]. For instance, as the temperature increases, the cement paste begins to decompose as a consequence of dehydration. Physical-chemical changes also take place in the aggregates. For this reason the decrease of compressive strength of concrete at elevated temperatures depends also on the type of aggregates used. Hence the mechanical properties, such as the elastic modulus and compressive strength, decrease with increasing temperatures. The concrete becomes more compliant. Its thermal dilatation coefficient increases non-linearly with temperature [8]. Concrete spalling may also appear as a result of thermal gradients and increasing pore pressure [9].

The majority of studies carried out on the behaviour of RC columns are based on the standard heating conditions, which do not take into account post-flashover fire temperature-time curve. There is no reliable experimental data, mathematical models, or design specifications for predicting the behavior of these elements under realistic fire scenarios. Thus, the current approach of determining fire resistance of RC columns by testing under standard fire conditions may not be realistic, since a number of factors such as compartmentation characteristics (opening size, fire load density, thermal properties) are not accounted for. Furthermore, the mechanical response depends on both load and temperature histories. Such dependence manifests itself in the absence of transient creep if a concrete specimen has already been subject to a high temperature cycle prior to testing. It is believed that any structural analysis of heated concrete that ignores this phenomenon will yield erroneous results, particularly for columns exposed to fire. Consequently, the design of columns is of paramount importance to the structural integrity of the whole structure, and should be based on more direct analytical approach that determines the structural fire response, which takes into account the fire compartmentation characteristics, the ventilation conditions, and, among all, the thermo-mechanical properties of the structural materials as well as transient creep phenomenon effects.

The aim of this study is to investigate the effect of transient creep on the structural response of RC columns subjected to natural or parametric fires using a finite element model developed by the authors [3,4]. The model, capable of analysing the response of RC columns from pre-fire stages to collapse in fire environment, is specially developed for the analysis of RC structures under severe thermo-mechanical loads, which accounts for transient creep explicitly as an additional component of the total strain of the concrete or implicitly through the use of the materials' properties recommended by EC2 [10].

2 BASIS OF THE NUMERICAL MODEL

The thermo-mechanical response of RC columns is obtained in an uncoupled way with incrementing time in steps. Because of the dependence of the material properties on temperature, thermal and mechanical responses are obtained incrementally. But initially, the

element is analyzed at ambient temperature under service loads only, just prior to the beginning of a fire.

2.1 Main steps of fire analysis

Within a time step, the fire resistance analysis of a restrained RC column subject to natural fire scenarios is performed through four main steps as briefly illustrated on figure 1.

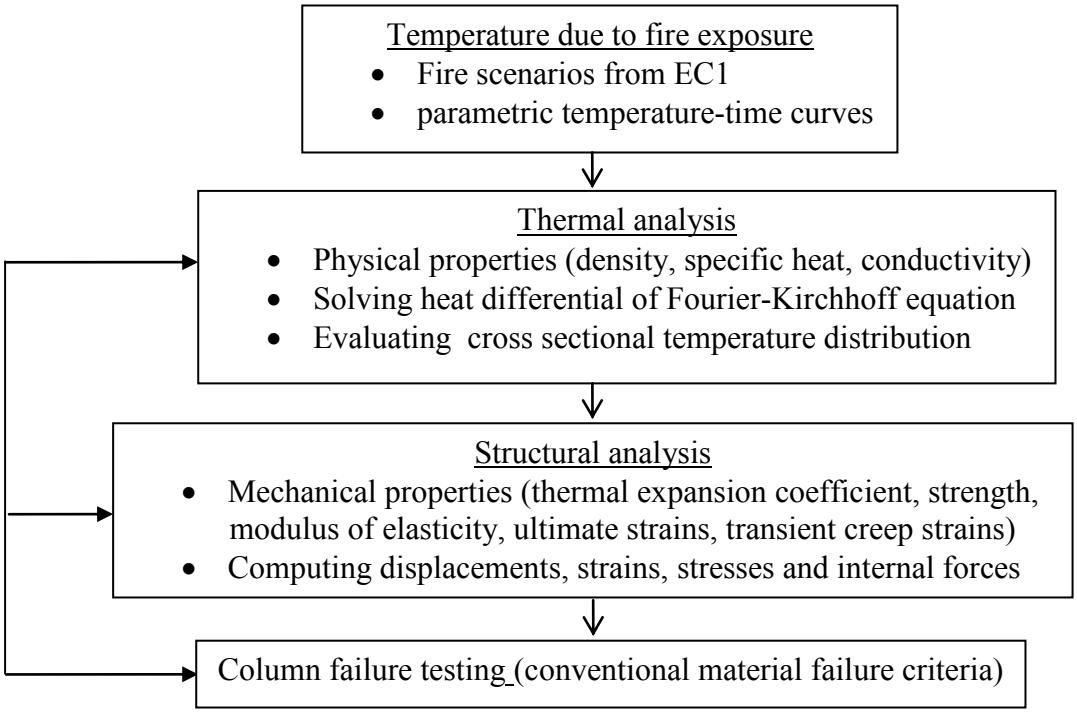


Figure 1: Main steps of fire analysis of RC columns

The first step consists in determining the temperature distribution due to fire exposure with help of the natural fire scenario defined in EC1 [11] as parametric temperature time curves. The second step is the thermal analysis to determine the temperature distribution along the cross-section of the column in order to evaluate the changes occurring in the material mechanical properties. This is followed by a static structural analysis to obtain the equilibrium configuration of the structural element at any time of the fire exposure. The computer model generates various critical output parameters, such as temperatures, stresses, strains, displacements and fire induced transient creep forces at various given fire exposure times. These parameters are then used to check against pre-determined failure criteria, in which the failure of a RC column is said to occur when the column element is unable to resist the specified applied service load (conventional material failure criterion). The failure occurs either by loss of the stability of the column or by exceeding the load-bearing capacity of the cross-section, which may be expressed by the failure criterion material (crushing of concrete and/or yielding of the steel).

2.2 Thermal analysis

The computed temperature-time curves describing the fire scenario are used to calculate the temperatures within the column cross-section using a finite element formulation. Cross sectional area of each column finite element is subdivided into a number of layers or sub-layers forming a two dimensional mesh. The temperature rise is obtained by establishing a heat balance of the cross section and detailed equations for the calculations are derived [12]. The temperature is assumed to be uniform along the length of the column and thus the calculations are performed for a unit length of the cross section. Steel reinforcement is not specifically considered in the thermal analysis because it does not significantly influence the temperature distribution. The initial moisture content is taken equal to 6%, which, according to EC2, corresponds to a peak value of the specific heat equal to 3700 J / kg. ° K.

2.3 Structural analysis

The cross-sectional temperatures generated from thermal analysis are used as input to the structural analysis step. A two-dimensional Euler-Bernoulli beam-column element based on the finite element displacement method approach and a Lagrangian co-rotational formulation are used in order to describe the deformation behavior of the column at any time of the fire exposure. Gaussian and Simpson rules are also employed to integrate stiffness matrices and internal forces along the longitudinal and transverse directions. Furthermore, the equilibrium equations which are derived from the incremental formulation of the principle of virtual displacements, make use of a tangential operator derived by integrating the stress-strain rates of concrete and steel at elevated temperatures. The formulation includes geometrical nonlinear effects resulting from large displacements, and material nonlinearities due the degradation of the elastic and inelastic properties with temperature. Detailed equations for formulations have been already described to some extent in [13]. Also, the concrete model accounts for cracking and crushing, as well as transitional thermal creep, which is considered in the present work explicitly as an additional component of the total strain, or implicitly through the use of the materials' properties of the concrete as recommended by the EC2.

2.4 Displacement and strain decompositions

Given, the equilibrium configurations of the element at the instants t_0 , t and t + Δt such as represented schematically on figure 2. Taking into account the geometrical configuration of a beam element for which the longitudinal dimension is very important compared to the two other dimensions, and the Navier Bernoulli hypothesis of straight sections, it follows that the axial strain of any point at time t is given by:

$$ {}^{t}\varepsilon_x = {}^{t}u_0' - y\,{}^{t}v_0'' + \frac{1}{2}\left({}^{t}v_0'\right)^2 \tag{1} $$

The prime denotes derivative with respect to ${}^{0}x$. In an expression analogous to (1), the strain at t+Δt is obtained as :

$$ {}^{t+\Delta t}\varepsilon_x = {}^{t+\Delta t}u_0' - y\,{}^{t+\Delta t}v_0'' + \frac{1}{2}\left({}^{t+\Delta t}v_0'\right)^2 \tag{2} $$

The preceding expressions give the strains at times t and t+Δt of any point of the section in the case of important displacements and small strains. In order to introduce the incremental formulation of the principle of virtual displacements, let's write the following incremental decompositions in stress and displacements:

$$^{t+\Delta t}\sigma_x = {}^{t}\sigma_x + \Delta\sigma_x \tag{3}$$

$$^{t+\Delta t}u_0 = {}^{t}u_0 + \Delta u_0 \tag{4}$$

$$^{t+\Delta t}v_0 = {}^{t}v_0 + \Delta v_0 \tag{5}$$

Substituting (4) and (5) in (2) leads:

$$^{t+\Delta t}\varepsilon_x = {}^{t}\varepsilon_x + \Delta\varepsilon_x \tag{6}$$

$$\Delta\varepsilon_x = \Delta u_0' - y\Delta v_0'' + {}^{t}v_0'\Delta v_0' + \frac{1}{2}\left(\Delta v_0'\right)^2 \tag{7}$$

which in turn can be decomposed into a linear and non linear parts :

$$\Delta\varepsilon_x = \Delta e_x + \Delta\eta_x \tag{8}$$

$$\Delta e_x = \Delta u_0' - y\Delta v_0'' + {}^{t}v_0'\Delta v_0' \tag{9}$$

$$\Delta\eta_x = \frac{1}{2}\left(\Delta v_0'\right)^2 \tag{10}$$

2.5 Materials thermo-mechanical law

In a time increment, Δt, the total strain increment, $\Delta\varepsilon_x$, of a generic material fibre of a beam-column finite element is assumed to be the sum of increments of elastic, $\Delta\varepsilon_e$, plastic, $\Delta\varepsilon_p$, thermal, $\Delta\varepsilon_{th}$, transient creep, $\Delta\varepsilon_{tr}$; the latter being non-zero in concrete and assumed that develops under first time of heating and under a state of compression only [6]. The sum elastic and plastic parts of strain increment is termed the mechanical strain increment $\Delta\varepsilon_\sigma = \Delta\varepsilon_e + \Delta\varepsilon_p$ which can be written as:

$$\Delta\varepsilon_\sigma = \Delta\varepsilon_x - \Delta\varepsilon_{th} - \Delta\varepsilon_{tr} \tag{11}$$

We assume that the relationship between the mechanical strain and the longitudinal normal stress, σ, is given by the constitutive law $\sigma=f(\varepsilon_\sigma,\Theta)$, where f is a functional pertinent to the chosen material. In the present work, we use the temperature dependant constitutive laws of concrete and reinforcing steel as suggested in EC2. An isotropic strain-hardening model is assumed in loading-unloading cycles. The adopted incremental thermo-mechanical law can be written in the form:

$$\Delta\sigma_x = E_T\left(\Delta\varepsilon_x - \Delta\varepsilon_{TH} - \Delta\varepsilon_{TR}\right) \tag{12}$$

where $E_T = E_T(^t\varepsilon_x)$ is the tangent modulus corresponding to a strain $^t\varepsilon_x$ of the material σ-ε curve at time fire exposure t. $\Delta\varepsilon_{th}$, and $\Delta\varepsilon_{tr}$ are respectively the incremental strains of free thermal elongation and transient creep computed during the time step Δt at any given iteration Newton-Raphson process as described in the following section.

2.6 Free thermal elongation and transient creep strains

The free thermal elongation strain in concrete and steel, $\Delta\varepsilon_{th,c}$ and $\Delta\varepsilon_{th,s}$, is assumed to be a function of the current temperature, Θ, and is given by the generic relation $\Delta\varepsilon_{th}=\Delta\varepsilon_{th}(\Theta)$. The approximation of the function for concrete and steel as defined in EC2 is adopted here. The concrete transient creep strain, $\Delta\varepsilon_{tr,c}$, is assumed to be a function of the current stress σ_c, time and temperature. We adopt the model proposed by Anderberg and Thelanderson [14], accounted for with a simple formula as:

$$\Delta\varepsilon_{tr,c} = \frac{k\sigma_c}{f_{c20}}\Delta\varepsilon_{th,c} \tag{13}$$

where f_{c20} is the compressive strength at ambient temperature (=20°C), $\Delta\varepsilon_{th,c}$ is the free thermal elongation strain increment of the concrete and, k a coefficient varying between 1.8 and 2.35.Noting that the transient creep strain is proportional to σ_c and thus a sign that is opposite to the sign of $\Delta\varepsilon_{th,c}$ where concrete is stressed in compression.

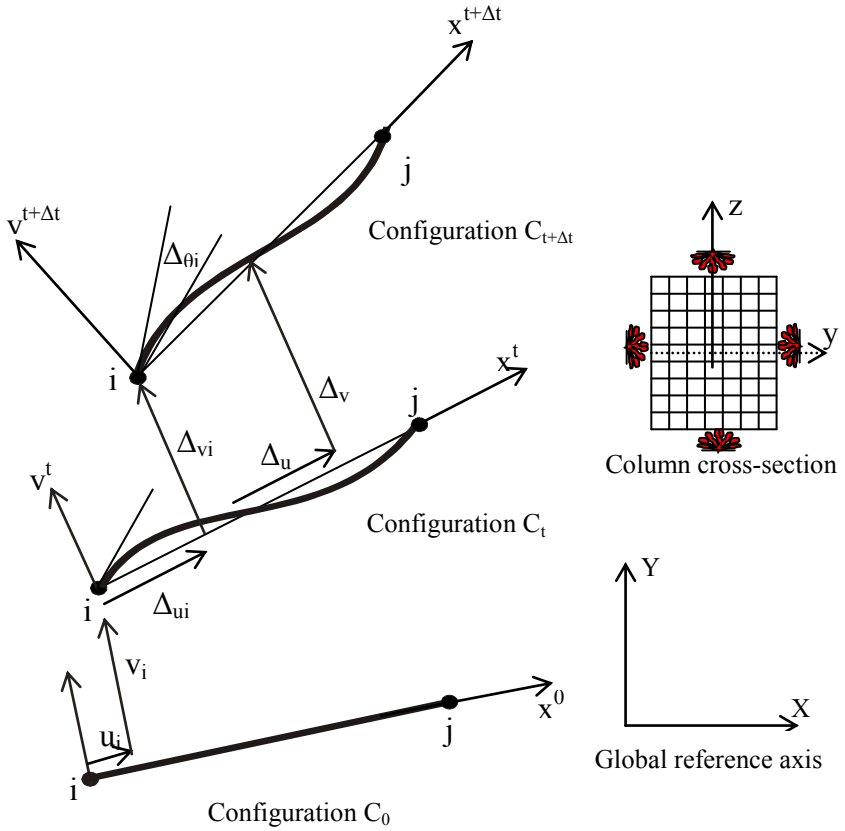


Figure 2: Equilibrium configurations of a structural finite element ij in a Lagrangian co-rotational formulation at any time fire exposure (t)

3 FIRE SCENARIOS

Three fire scenarios are investigated; namely ISO834 and two design fire scenarios: Fire I and Fire II. There is no decay phase in the time-temperature curves for the standard fire scenario. However, in realistic fires represented by the two design fires, there is always a decay phase, since the amount of fuel or ventilation runs out leading to the burn out phase of the fire. The parametric fire time-temperature curve proposed in EC1 are selected to represent the design fire scenarios used in the analysis, which are influenced by compartment properties such as the fuel load, ventilation opening and wall linings. To develop the two design fire scenarios, a fire is assumed to occur in a room with dimensions of 6mx4mx3m as illustrated in figure 3 [13]. Two values of the fuel load and the opening dimensions are also assumed. More details about the properties of the room for the two fires are represented in Table 1. The values were assumed in such a way that Fire I represents a severe design fire whose temperatures exceeds 1200°C, such as in a library or storage room where large amount of combustible materials and sufficient ventilation are available. Fire II represents a moderate design fire whose temperatures reach 600°C. The obtained time-temperature curves are shown in Figure 4. The results from the analysis indicate that the fire scenario has a significant influence on the temperature distribution across the column section. As expected, the temperature at various depths of concrete increases with the fire exposure time for the standard fire scenario. The results also show that the increase in concrete is slightly larger for the ISO 834.

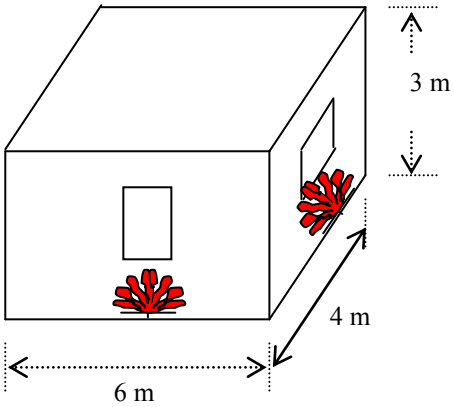


Figure 3: Compartment in fire

Table 1: Compartment characteristics

Design Fire	Lining material	$Ws^{0.5}/m^2°K$ (Thermal capacity lining material)	Opening (m)	MJ/m^2 floor area (Fire load)
Fire I	Gypsum	488	2.25x1.5	1200
Fire II	Concrete	1500	2.85x1	400

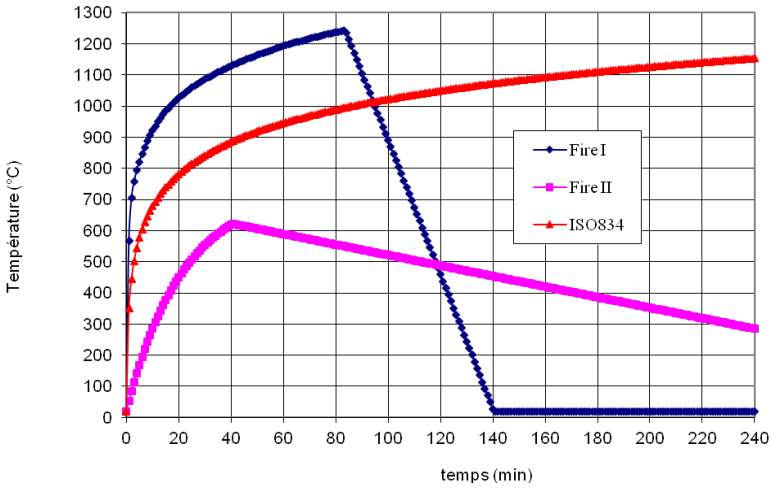


Figure 4: Various fire scenarios

4 RESULTS AND DISCUSSIONS

Geometrical and loading details of the column are shown on figure 5 having an initial sinusoidal imperfection of $\omega_o = \ell/1000$. It is sufficiently slender with an initial deflection as to induce second order effects. Additionally, a compressive load of P_0=740 kN is applied with an eccentricity e_0 of 15 mm. The column is pinned at both ends with the loaded end free to elongate axially. It is heated on all the four faces while keeping the applied load constant. The column is assumed to be made of normal concrete with a compressive strength f'_c = 30 MPa, tensile strength f_t= 3.5 MPa, and reinforced with steel rebars having a yield strength f_y = 400 MPa. The concrete cover relative to the stirrups is taken as equal to c= 30 mm for all the columns. In the analysis, the column is also assumed to be exposed to fire from all sides, and the materials' properties suggested by the EC2 are adopted. The fire resistance is evaluated based on the conventional material failure criterion previously defined particularly expressed by crushing of concrete and/or yielding of the steel. Fire induced spalling is supposed not occur since the column is made of normal concrete (NC), which exhibits a relatively high permeability. This provides an easy mechanism for the water vapor to escape from the concrete. Thus, with increasing temperatures the pore pressure continuously dissipates in the column, and there is no significant vapor pressure build-up. This is agreement with reported test results, which clearly show negligible spalling in NC members during fire exposure [15].

To investigate the effect of transient creep on its behavior, the column is analysed under two different cases. In the first case, transient creep is considered implicitly through the use of the variations of the material properties with temperature as recommended by the Eurocode [10]. In the second case, transient creep is considered explicitly as an additional strain component. The obtained results are shown in the form of the progression with time fire exposure of the axial displacement, u, of the loaded end and the lateral deflection, v, of the middle of the column.

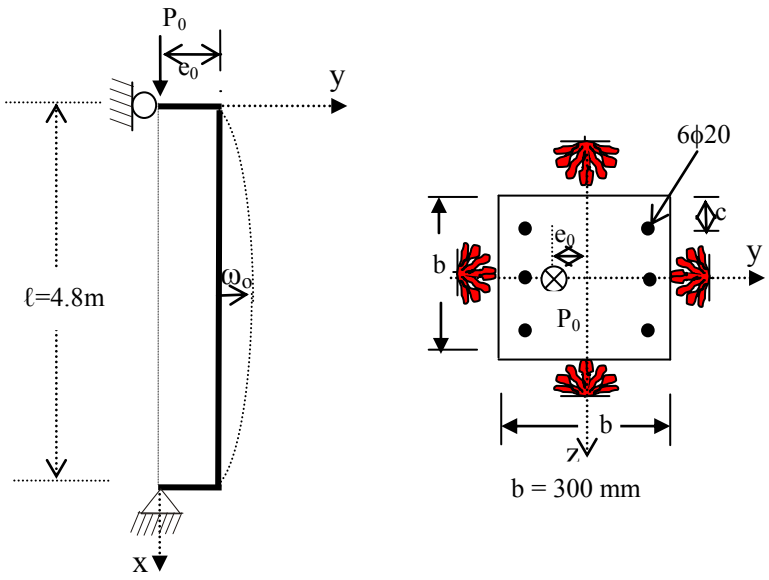


Figure 5: Geometrical details of the column

4.1 Effect of fire scenario

Figures 6 and 7 show the results with fire exposure for the three fire scenarios when transient creep is considered implicitly through the use of Eurocode's properties. It can be seen that the lowest fire resistance is obtained for the column exposed to the severe fire design Fire I. It is about 38 min in Fire I, 62 min in ISO834 and 180 min in Fire II. This is due to the rapid increase in temperature for Fire I as can be seen from figure 4. A reduction (or recovery) in the displacements is also noticeable particularly for the moderate fire design FireII as a result of the decay phase. Hence, the result reveals that in many applications, the fire resistance values computed based on standard fire exposure may not be conservative if the resulting fire has a decay phase similar to the one of design Fire I used in this study.

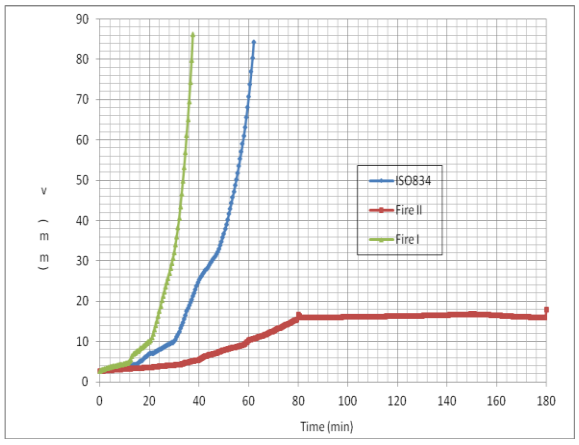


Figure 6: Effect of fire scenario on the deflection

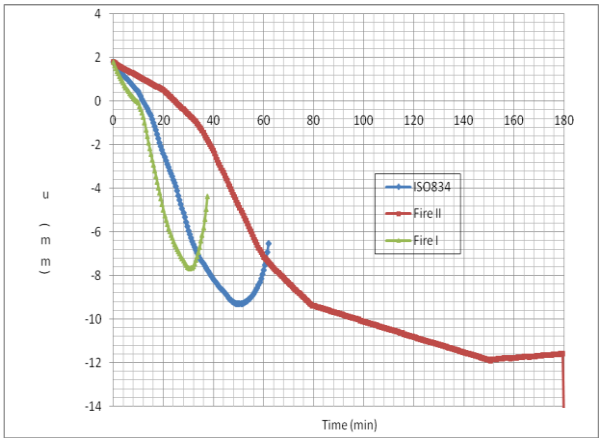


Figure 7: Effect of fire scenario on the axial displacement

4.2 Effect of transient creep

Figure 8 shows the model's predictions under the severe fire design (Fire I) when transient creep is considered explicitly with: k = 0, k = 1.8 and k = 2.35. Temperature dependent material properties used are those derived originally by Anderberg and Thelandersson which also reported in [5,4] since Eurocode's material data with temperature presents some ambiguity on whether these properties include transient creep or not. This can be explained by the fact the transient creep strains counteracts the thermal expansion and thus creating additional compressive stresses. Noting that the column is under a state of bending due to the influence of the eccentric load applied initially prior to the beginning of the fire. The increase in compressive stresses results in more plasticization of the compressed zone and more cracking in the tensile zone, and hence in a reduction of the overall stiffness of the beam. The sudden increase in the lateral displacement v indicates the onset of buckling which precipitates the failure of the column.

The results are also compared with the implicit model when the properties recommended by Eurocode are used. As shown in figure 9, it can be seen that the results agree better when transient creep is taken into account. The value of k = 1.8, which is the least recommended, was found to give the best agreement

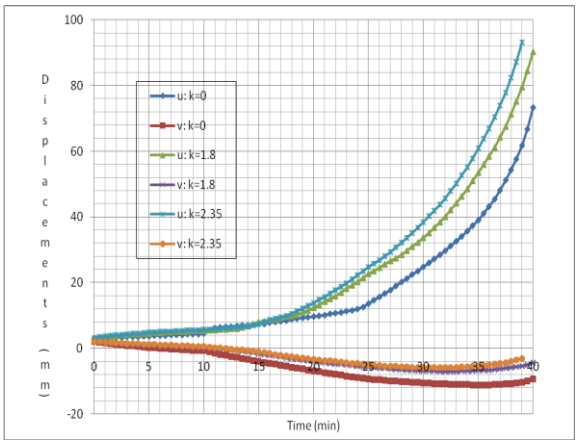


Figure 8: Effect of transient creep on the column response

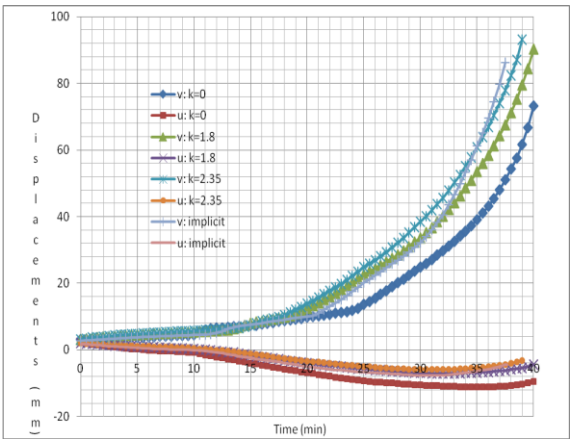


Figure 9: Model predictions using both the Eurocodes's properties and transient creep

5 CONCLUSIONS

- The transient creep on the fire response of RC columns is investigated using two different approaches. In the first approach transient creep is considered explicitly as an additional component of the total strain, while in the second approach, the phenomenon is taken into account implicitly through the use of concrete data properties with temperature as recommended by the Eurocode.
- The transient creep modelled as an additional strain component is a more robust

approach since it provides an insight into the response of the column. The value of the transient creep strains coefficient k = 1.8 is found to give the best agreement comparatively to the implicit model.

- The transient creep effects counteracts the thermal expansion and thus creating additional compressive stresses, which clearly explains the rapid failure of the column at elevated temperatures where transient creep strains are important.
- The conventional methods of evaluating fire resistance, computed based on standard fire exposure, may not be conservative if the resulting fire has a decay phase similar to the severe fire scenarios as the one used in this study.

REFERENCES

[1] Martins, M.B. & Rodrigues, J.P.C. Fire resistance of RC columns with elastically restrained thermal elongation, *Engineering structure,* 32 (10), 3330-3337, (2010).

[2] Lie, T.T., Lin, T.D., Allen, D.E. and Abrahams, M.S. Fire resistance of RC concrete columns, *Ottawa: National research council of Canada,* Division of Building research, (1984), NRCC, 23065.

[3] Sadaoui A., Illouli S. and Khennane A. Effect of restraints on the response of RC columns in a parametric fire, FraMCoS-8, *8th International Conference on Fracture Mechanics of Concrete and Concrete Structures*, 10-14 mars (2013), Toledo, Spain.

[4] Sadaoui A., Khennane A. Effect of transient creep on the behaviour of reinforced concrete columns in fire, *Engineering Structures*, (2009), Vol. 31, issue 9, pp. 2203-2208.

[5] Van Foeken,R.J. Numerical analysis of reinforced concrete structures at high temperatures including cracking, plasticity and creep, *Num. Meth. Thermal Problems.*, (1985), V4, p. 1127-83.

[6] Khoury GA, Graiger B.N., Sullivan P.J.E. Transient thermal strain of concrete: literature review, conditions within specimen and behaviour of individual constituents, *Magazine of Concrete Research*, 1985, Vol.37, p. 131-144.

[7] Bratina S., Saje M. and Plantine I. The effects of different strain contributions on the response of RC beams in fire, *Engeneering Structures*, (2009), 29, pp. 418-430.

[8] Harmathy, T.Z. Thermal properties of concrete at elevated temperatures. ASTM *Journal of materials*, (1970), Vol.5, p. 47-74.

[9] Rostasy F.S., Ehm C., Heinrichsmeyer K. Changes of pore structure of cement mortars due to temperature, *Cement and Concrete Research*, (1980), Vol. 10, n°2, p. 157-164.

[10] European standard EN 1992-1-2, Eurocode 2: Design of concrete structures- part 1-2: General rules- Structural fire design. Brussels: CEN, (2004).

[11] European standard EN 1991-1-2, Eurocode 1: Basis of design and actions on structures, part 1-2: actions on structures exposed to fire. Brussels: CEN- CEN, (2002).

[12] Terro M. J. Numerical modeling of the behavior of concrete structures, *ACI Structural Journal*, V.95, (1998), N°2, pp. 183-193.

[13] Sadaoui, A., Kaci, S. and Khennane, A. Behavior of RC frames in a fire environment including transitional thermal creep, *Aust J Struct Eng*, (2007), AJSE, 7(3), 167-184.

[14] Anderberg, G.Y. and Thelanderson, S. Stress and deformation of concrete at high temperatures. 2 Experimental investigation and material behaviour: Bulletin 54, Lund Institute of Technolgy, Lund, (1976).

[15] Lie T.T. and Lin T.D. Influence of restraint on fire performance of reinforced concrete columns, *Fire Safety Science*, Proceedings of the 1st International Symposium, Berkeley, USA, , pp. 291–300, (1986).

VI International Conference on Computational Methods for Coupled Problems in Science and Engineering
COUPLED PROBLEMS 2015
B. Schrefler, E. Oñate and M. Papadrakakis (Eds)

THERMOELASTIC PHENOMENA IN THE TRANSVERSALLY GRADED LAMINATES

J. JĘDRYSIAK[*] AND E. PAZERA[†]

[*]Department of Structural Mechanics
Łódź University of Technology
al. Politechniki 6, 90-924 Łódź, Poland
e-mail: jarek@p.lodz.pl

[†]Department of Structural Mechanics
Łódź University of Technology
al. Politechniki 6, 90-924 Łódź, Poland
email: ewelina.pazera@p.lodz.pl

Key words: Thermoelasticity, Transversally Graded Laminates, Functionally Graded Materials, Tolerance Modelling, Asymptotic Modelling.

Abstract. The object of considerations is a laminate, made of two components non-periodically distributed as microlaminas along one direction. In this note, two models describing the thermoelasticity problems are proposed: the tolerance model, taking into account the effect of the microstructure size on the overall behaviour of these laminates, and the asymptotic model, neglecting this effect. To obtain the governing equations of the tolerance and the asymptotic model, the tolerance averaging technique is applied.

1 INTRODUCTION

In this paper we deal with the problem of thermoelasticity in a two-phase laminate. The cells are composed of two sublayers of different material constituents. The thickness of the cells (cf. Fig. l) is denoted by l and is constant (so-called the uniform distribution of the cells). It is assumed that macroscopic properties of this laminate change continuously along one direction (perpendicular to the laminas). Hence, this object can be treated as made of a functional graded material (FGM), cf. Suresh and Mortensen [1]. Such laminates can be called *transversally graded laminates* (TG-type laminates), cf. Jędrysiak [2]. A fragment of the laminate of this kind is shown in Fig. 1.

Although usually in such composites basic cell cannot be simply defined and thermomechanical phenomena can be considered only with reference to micromechanical models with idealised geometries. The functionally graded laminates are not homogeneous in a macroscale, nonetheless in the analysis of various issues concerning these laminates, assumptions of idealisation can be similar to those applied to analyse composites with a periodic structure and their overall behaviour can be described by adopted and modified methods, which are used for macroscopically homogeneous composites. Some of the averaged methods adapted to analyse and determine TG-type laminates are discussed in Suresh and Mortensen [9] and Reiter et al. [7]. Between these methods, there are techniques

based on the asymptotic homogenizations, cf. Jikov et al. [8] and methods using models with microlocal parameters, cf. Matysiak [10]. There are the alternative methods for functionally graded materials, such as a Green's function approach, cf. Kim and Noda [11], the Laplace transformation, cf. Ootao and Tanigawa [12], higher-order plate theory is proposed by Aboudi et al. [13] and reformulated by Bansal and Pindera [14], stochastic thermal stresses in objects with functionally graded materials properties are analysed by Chiba [15], a numerical homogenization technique is used by Schmauder and Weber [16], the boundary element method by Goldberg and Hopkins [17], meshless computational method is proposed by Sladek J. et al. [18], the finite difference method by Sadowski et al. [19]. Unfortunately, the most of the know approaches, do not take into account the effect of microstructure size on the overall behaviour of these laminates.

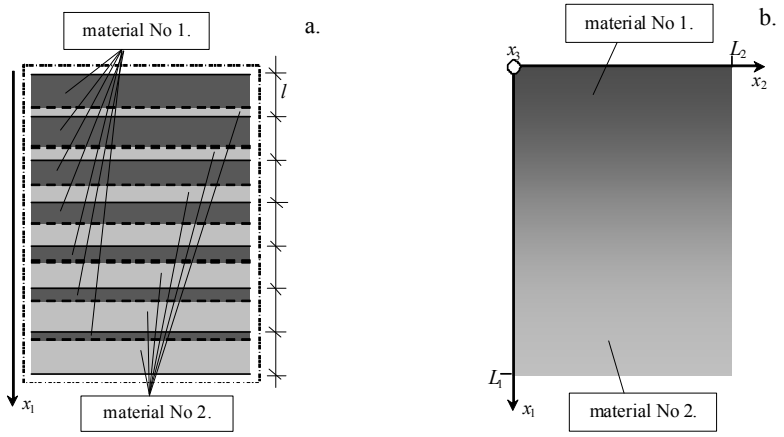


Figure 1: The cross-section of considered laminate: a. microstructure, b. macrostructure.

In order to obtain the averaged equations, taking into account this effect, *the tolerance averaging technique* is applied, proposed by Woźniak and Wierzbicki [15] to describe various thermo-mechanical phenomena in periodic composites. This way of modelling and the results of applying this technique for functionally graded media can be found in Jędrysiak [16], Woźniak [17], Woźniak et al. [18], Ostrowski and Michalak [19], Ostrowski [20]. TG-type laminates was also investigated, e.g. a heat conduction by Jędrysiak and Radzikowska [21]. The tolerance averaging technique replace equations with functional, highly-oscillating, tolerance-periodic and non-continuous coefficients by a system of differential equations with slowly-varying coefficients.

The main aim of this note is to obtain equations, describing thermoelasticity problems for TG-type laminates. The equations of two models are received. One of them is *the tolerance model*, which takes into account the effect of the microstructure size on the overall behaviour of the laminates with functionally graded material properties. And the second - *the asymptotic model* - which neglects this effect. Both of these models are based on concepts of the tolerance modelling. Applications of these models are shown in some examples. Applications of these models are shown in some examples. The problem of thermoelasticity in micro-periodic structures was also considered by use of other mathematical models, approaches and shown in many papers. We shall mention here Refined Averaged Theory presented in paper of Baczyński [22] and Effective Modulus Theory in papers by Ignaczak [23].

2 MODELLING FOUNDATIONS

Let subscripts $i,j,..$, run over 1,2,3 (related to the coordinate system $Ox_1x_2x_3$ and subscripts $\alpha,\beta,..$, run over 2,3 (related to the coordinate system Ox_2x_3). Denote: $\mathbf{x}\equiv(x_1,x_2)$, $x\equiv x_1$, the time coordinate by t, derivatives respected to x_i, x_α, x by respectively ∂_i, ∂_α, ∂ respectively, the laminate length along the x-axis by H, length dimensions along x_α-axis by L_α. Using denotations $\Omega\equiv(0,H)$, $\Xi\equiv(0,L_2)\times(0,L_3)$, the laminate under consideration occupies in the physical space the region $\Omega\times\Xi$. This laminate is made of two materials distributed in m laminas, the thickness of the lamina (l) is constant ($H=ml$) and can be called *the microstructure parameter*. It is assumed that condition $l<<H$ is satisfied. Denote by $\Delta\equiv(-l/2, l/2)$ the basic cell in the interval Ω. Every lamina n is made of two homogeneous sublayers with thickness l'_n, l''_n, depended on the argument x. The properties of sublayers are described by specific heats c' and c'', heat conduction tensors with components k'_{ij}, k''_{ij}, mass densities ρ', ρ'', elasticity tensors with components c'_{ijkl}, c''_{ijkl}, thermoelasticity tensors b'_{ij}, b''_{ij}; where i,j,k,l=1,2,3. Denote material volume fractions in the n^{th} lamina by $\nu'_n\equiv l'_n/l$, $\nu''_n\equiv l''_n/l$. We assume that the following conditions are satisfied: $|\nu'_{n+1}-\nu'_n|<<1$ and $\nu'_n+\nu''_n=1$. It can be shown that sequences $\{\nu'_n\}$, $\{\nu''_n\}$, n=1,…,m, can be approximated by continuous functions $\nu'(\cdot)$, $\nu''(\cdot)$. These functions can be called *the fraction ratios of materials* and describe gradation of material properties along the x-axis. The non-homogeneity ratio is denoted and defined by $\nu(\cdot)\equiv\sqrt{\nu'(\cdot)\nu''(\cdot)}$. The unknown displacements along the x_i-axis are denoted by u_i (i=1,2,3) and the temperature field by Θ.

The thermoelasticity problems for laminates with transversally graded material properties can be described by the know equations:

$$\begin{aligned}&\partial_j(c_{ijkl}\partial_l u_k)-\rho\ddot{u}_i=\partial_j(b_{ij}\Theta),\\&\partial_j(k_{ij}\partial_i\Theta)-c\dot{\Theta}=T_0 b_{ij}\partial_j\dot{u}_i,\end{aligned}\tag{1}$$

where T_0 is a constant with a temperature dimension.

Unfortunately, all coefficients of the above equations (1), k_{ij}, c, c_{ijkl}, b_{ij}, ρ, are highly-oscillating, tolerance-periodic and non-continuous functions in x. By using the tolerance modelling we can replace these equations by differential equations with smooth, slowly-varying functional coefficients, cf. Woźniak et al. [17].

3 INTRODUCTORY CONCEPTS

For this purpose, some basic concepts of the tolerance modelling, e.g. *an averaging operator, a tolerance-periodic function, a slowly-varying function, a highly-oscillating function*, are introduced below. The averaging operator for an arbitrary integrable function f, is defined by:

$$\langle f\rangle(x)=\tfrac{1}{l}\int_{x-l/2}^{x+l/2}f(y)dy,\quad x\in[l/2,H-l/2].\tag{2}$$

Denote k^{th} gradient of function $f=f(x)$ by $\partial^k f$, $x\in\Omega$, k=0,1, $\partial^0 f\equiv f$ and function defined in $\Omega\times R^m$ by $\tilde{f}^{(k)}(\cdot,\cdot)$. Function $f\in H^\alpha(\Omega)$ is the tolerance-periodic function (with respect to cell

Δ and tolerance parameter δ), $f \in TP_{\delta}^{\alpha}(\Omega,\Delta)$, if for k=0,1, the following conditions hold:

a. $(\forall x \in \Omega)(\exists \widetilde{f}^{(k)}(x,\cdot) \in H^0(\Omega)[\| \partial^k f|_{\Omega_x}(\cdot) - \widetilde{f}^{(k)}(x,\cdot)\|_{H^0(\Omega_x)} \leq \delta]$,

b. $\int_{\Delta(\cdot)} \widetilde{f}^{(k)}(\cdot,y)dy \in C^0(\overline{\Omega})$,

where function $\widetilde{f}^{(k)}$ is the periodic approximation of $\partial^k f$ in $\Delta(x)=x+\Delta$, $x \in \Omega$, k=0,1.

Function $F \in H^{\alpha}(\Omega)$ is *the slowly-varying function* (with respect to cell Δ and tolerance parameter δ), $F \in SV_{\delta}^{\alpha}(\Omega,\Delta)$ if:

a. $F \in TP_{\delta}^{\alpha}(\Omega,\Delta)$,

b. $(\forall x \in \Omega)[\widetilde{F}^{(k)}(x,\cdot)\big|_{\Delta(x)} = \partial^k F(x), \quad k=0,1]$.

Thus, periodic approximation $\widetilde{F}^{(k)}$ of $\partial^k F$ in $\Delta(x)$ is a constant function for every $x \in \Omega$.

Function $\phi \in \mathrm{H}^{\alpha}(\Omega)$ is *the highly oscillating function* (with respect to cell Δ and tolerance parameter δ), $\phi \in HO_{\delta}^{\alpha}(\Omega,\Delta)$ if:

a. $\phi \in TP_{\delta}^{\alpha}(\Omega,\Delta)$,

b. $(\forall x \in \Omega)[\widetilde{\phi}^{(k)}(x,\cdot)\big|_{\Delta(x)} = \partial^k \widetilde{\phi}(x), \quad k=0,1]$.

Function $f \equiv \phi F \in TP_{\delta}^{\alpha}(\Omega,\Delta)$, for every $F \in SV_{\delta}^{\alpha}(\Omega,\Delta)$, satisfies the condition:

c. $\widetilde{f}^{(1)}(x,\cdot)\big|_{\Delta(x)} = F(x)\partial^{1}\widetilde{\phi}(x)\big|_{\Delta(x)}$, if α=0, we denote $\widetilde{f} \equiv \widetilde{f}^{(0)}$.

Let us denote by $h(\cdot)$ and $g(\cdot)$ highly oscillating, continuous functions, $h,g \in HO_{\delta}^{1}(\Omega,\Delta)$, defined on $\overline{\Omega}$, which have piecewise continuous and bounded gradients $\partial^1 h$, $\partial^1 g$. Functions $h(\cdot)$, $g(\cdot)$ are *the fluctuation shape functions* of the 1st kind, if they satisfy conditions:

a. $\partial^k h \in O(l^{\alpha-k}), \partial^k g \in O(l^{\alpha-k})$, for k=0,1, α=1, $\partial^0 h \equiv h, \partial^0 g \equiv g$,

b. $<h>(x) \approx 0, \ <g>(x) \approx 0$, for every $x \in \Omega$,

and depend on l as a parameter. Set of all fluctuation shape functions of the 1st kind is denoted by $FS_{\delta}^{1}(\Omega,\Delta)$.

4 TOLERANCE MODELLING

4.1 Fundamental assumptions

The tolerance modelling is based on two basic assumptions, cf. [18, 17, 2]. The first is *the decomposition of the micro-macro*, where it is assumed that the fundamental unknowns (in the present issue – the displacement field and the temperature field) can be taken as a sum of averaged part (averaged field) and oscillating part. Furthermore it is assumed that the oscillating part can be expressed as a product of a known oscillating function, called the fluctuation shape function, and the new unknown, *the fluctuation amplitude*. It is assumed that the new basic unknowns, the averaged fields and the amplitudes of fluctuations, are slowly-varying functions of that co-ordinate which parameterizes the perpendicular direction to the laminas. For the displacement and for the temperature, respectively:

$$u_i(x,\mathbf{x},t) = w_i(x,\mathbf{x},t) + h(x)v_i(x,\mathbf{x},t) \tag{3}$$

$$\Theta(x,\mathbf{x},t) = \theta(x,\mathbf{x},t) + g(x)\psi(x,\mathbf{x},t) \tag{4}$$

and $w_i(\cdot,\mathbf{x},t),v_i(\cdot,\mathbf{x},t),\theta(\cdot,\mathbf{x},t),\psi(\cdot,\mathbf{x},t)\in SV_\delta^1(\Omega,\Delta)$ for every t; $h(\cdot),g(\cdot)\in FS_\delta^1(\Omega,\Delta)$. Functions $w_i(\cdot,\mathbf{x},t)$ and $\theta(\cdot,\mathbf{x},t)$, are the basic unknowns, called *the macrodisplacements* and *the macrotemperature*, respectively; $v_i(\cdot,\mathbf{x},t)$ and $\psi(\cdot,\mathbf{x},t)$ are additional basic unknowns, called *the fluctuation amplitudes of displacements* and *the fluctuation amplitudes of temperature*, respectively; $h(\cdot)$ and $g(\cdot)$ are the known fluctuation shape functions. The fluctuation shape functions $h(x),g(x),x\in\overline{\Omega},$ can be given in the following form:

$$h(x)=g(x)=\begin{cases}[-l\sqrt{3}\dfrac{\nu(\bar{x})}{\nu'(\bar{x})}(2x/l+\nu''(\bar{x}))] & \text{for } x\in[-l/2,-l/2+l\,\nu'(\bar{x})],\\ [l\sqrt{3}\dfrac{\nu(\bar{x})}{\nu''(\bar{x})}(2x/l-\nu''(\bar{x}))] & \text{for } x\in[-l/2+l\,\nu'(\bar{x}),l/2],\end{cases} \tag{5}$$

where $\bar{x}$ is a centre of $\Delta=(-l/2,l/2)$. Because the non-homogeneity ratio $\nu(\cdot)$ is a slowly-varying function the mean values of h and g in every lamina are equal to zero. The shape of this function, $g(x)$ takes into account discontinuities of gradient of temperature field on intersections. For the displacement field it is the most common fluctuations shape function, $h(x)$ considered in such problems.

The second assumption is so-called *the tolerance averaging approximation* in which it is assumed that some components $O(\delta)$ are negligible small, e.g. in formulas:

$$\begin{aligned}&<f>(x)=<\overline{f}>(x)+O(\delta),\\ &<fF>(x)=<f>(x)F(x)+O(\delta),\\ &<f\partial(hF)>(x)=<f\partial h>(x)F(x)+O(\delta),\end{aligned}$$

where $x\in\Omega,\ 0<\delta<<1,\ f\in TP_\delta^1(\Omega,\Delta),F\in SV_\delta^1(\Omega,\Delta),h\in FS_\delta^1(\Omega,\Delta).$

4.2 Modelling procedure

The modelling procedure is outlined in Woźniak [17]. The starting point is the formulation of the extended action functional:

$$\begin{aligned}&\Lambda(u_i(\cdot),\Theta(\cdot),p_i(\cdot),f(\cdot))=\\ &\quad=\int_\Omega\int_\Xi\int_{t_0}^{t_1}\Lambda(z,\partial_j u_i(z,\xi,t),u_i(z,\xi,t),\partial_i\Theta(z,\xi,t),\Theta(z,\xi,t),p_i(z,\xi,t),f(z,\xi,t))dtd\xi dz,\end{aligned} \tag{6}$$

where $\Lambda(\cdot,\partial_j u_i,u_i,\partial_i\Theta,\Theta,p_i,f)\in HO_\delta^0(\Omega,\Delta),\ p_i(\cdot,\xi,t),f(\cdot,\xi,t)\in HO_\delta^0(\Omega,\Delta)$.
Lagrangian Λ is given by the formula:

$$\Lambda=\tfrac{1}{2}(\rho\dot{u}_i\dot{u}_j\delta_{ij}-c_{ijkl}\partial_j u_i\partial_l u_k+k_{ij}\partial_i\Theta\,\partial_j\Theta)+p_iu_i+f\Theta. \tag{7}$$

The constitutive equations determine p_i and f:

$$p_i=\partial_j(b_{ij}\Theta),\qquad f=c\dot{\Theta}+T_0b_{ij}\partial_j\dot{u}_i. \tag{8}$$

For Lagrangian (7) we can write the Euler-Lagrange equations:

$$\frac{\partial}{\partial t}\frac{\partial\Lambda}{\partial\dot{u}_i}+\partial_j\frac{\partial\Lambda}{\partial\,\partial_j u_i}-\frac{\partial\Lambda}{\partial u_i}=0,\qquad \partial_j\frac{\partial\Lambda}{\partial\,\partial_j\Theta}-\frac{\partial\Lambda}{\partial\Theta}=0. \tag{9}$$

The fundamental equations of thermoelasticity problems (1) are obtained by using the extended principle of stationary action from equations (9) combined with equations (7) and (8). Then, we apply the tolerance modelling to action functional (6). Substituting micro-macro decompositions (3) and (4) to formula (6) and using averaging operator (2) to this action functional, we arrive at the tolerance averaged functional $\mathrm{A}(u_i(\cdot),\Theta(\cdot),p_i(\cdot),f(\cdot))$:

$$\mathrm{A}_{hg}(w_i,v_i,\theta,\psi,p_i,f)=\int_\Omega\int_\Xi\int_{t_0}^{t_1}<\Lambda_{hg}>(x,\partial_j w_i,w_i,\partial_\alpha v_i,v_i,\partial_i\theta,\theta,\partial_\alpha\psi,\psi,p_i,f)dtd\xi dx, \tag{10}$$

where the averaged form $<\Lambda_{hg}>$ of lagrangian (7) can be written as:

$$\begin{aligned}<\Lambda_{hg}>&=\tfrac{1}{2}<k_{ij}>\partial_i\theta\partial_j\theta\ +\tfrac{1}{2}<k_{i1}\partial g>\partial_i\theta\,\psi+\tfrac{1}{2}<k_{j1}\partial g>\partial_j\theta\,\psi+\tfrac{1}{2}\underline{<k_{\alpha\beta}gg>}\partial_\alpha\psi\partial_\beta\psi+\\&+\tfrac{1}{2}<k_{11}\partial g\partial g>\psi\psi+\tfrac{1}{2}<\rho>\dot w_i\dot w_j\delta_{ij}+\tfrac{1}{2}\underline{<\rho hh>}\dot v_i\dot v_j\delta_{ij}-\tfrac{1}{2}<c_{ijkl}>\partial_j w_i\partial_l w_k-\tfrac{1}{2}<c_{ijkl}\partial h>\partial_j w_i v_k-\\&-\tfrac{1}{2}<c_{i1kl}\partial h>v_i\partial_l w_k-\tfrac{1}{2}<c_{i1kl}\partial h\partial h>v_i v_k-\tfrac{1}{2}\underline{<c_{i\alpha k\beta}hh>}\partial_\alpha v_i\partial_\beta v_k+\\&+<p_i>w_i+\underline{<p_ih>}v_i+<f>\theta+\underline{<fg>}\psi,\end{aligned} \tag{11}$$

and $<p_i>$, $<p_ih>$, $<f>$, $<fg>$ are defined by the constitutive equations:

$$\begin{aligned}&<f>=<c>\dot\theta+<T_0b_{ij}>\partial_j\dot w_i+<T_0b_{i1}\partial h>\dot v_i,\\&<fg>=<cgg>\dot\psi+<T_0b_{i\beta}hg>\partial_\beta\dot v_i,\\&<p_i>=\partial b_{ij}\theta+<b_{ij}>\partial_j\theta,\\&<p_ih>=-<b_{i1}\partial g>\theta\ +\underline{<b_{i\beta}gh>}\partial_\beta\psi.\end{aligned} \tag{12}$$

Underlined terms in formulas (11) and (12) depend on the microstructure parameter l. Using the extended principle of stationary action to Λ_{hg} the following system of equations with slowly-varying coefficient is obtained:

$$\begin{aligned}&\frac{\partial}{\partial t}\frac{\partial<\Lambda_{hg}>}{\partial\dot w_i}+\partial_j\frac{\partial<\Lambda_{hg}>}{\partial\,\partial w_i}-\frac{\partial<\Lambda_{hg}>}{\partial w_i}=0,\\&\frac{\partial}{\partial t}\frac{\partial<\Lambda_{hg}>}{\partial\dot v_i}+\partial_\alpha\frac{\partial<\Lambda_{hg}>}{\partial\,\partial_\alpha v_i}-\frac{\partial<\Lambda_{hg}>}{\partial v_i}=0,\\&\partial_j\frac{\partial<\Lambda_{hg}>}{\partial\,\partial_j\theta}-\frac{\partial<\Lambda_{hg}>}{\partial\theta}=0,\qquad\partial_\alpha\frac{\partial<\Lambda_{hg}>}{\partial\,\partial_\alpha\psi}-\frac{\partial<\Lambda_{hg}>}{\partial\psi}=0,\end{aligned} \tag{13}$$

5 GOVERNING EQUATIONS

5.1 The tolerance model equations

The following form of averaged equations for thermoelasticity problems is obtained by combining equations (13) with formulas (11) and (12):

$$\begin{aligned}&\partial_j(<c_{ijkl}>\partial_l w_k+<c_{ijk1}\partial h>v_k)-<\rho>\ddot w_i=\partial<b_{ij}>\theta+<b_{ij}>\partial_j\theta,\\&<c_{i1kl}\partial h>\partial_l w_k\underline{-\partial_\alpha(<c_{i\alpha k\beta}hh>\partial_\beta v_k)}+<c_{i1k1}\partial h\partial h>v_k+\underline{<\rho hh>\ddot v_i}=<b_{i1}\partial h>\theta-\underline{<b_{i\beta}gh>\partial_\beta\psi},\\&\partial_j(<k_{ij}>\partial_i\theta\ +<k_{1j}\partial g>\psi)=<c>\dot\theta+<T_0b_{ij}>\partial_j\dot w_i+<T_0b_{i1}\partial h>\dot v_i,\\&\underline{\partial_\alpha(<k_{\alpha\beta}gg>\partial_\beta\psi)}-<k_{i1}\partial g>\partial_i\theta-<k_{11}\partial g\partial g>\psi=\underline{<cgg>\dot\psi}+\underline{<T_0b_{i\beta}hg>\partial_\beta\dot v_i},\end{aligned} \tag{14}$$

where the underlined terms depend on the microstructure parameter l.

In the above equations (14) the coefficients are slowly-varying functions in x, in contrast to equations (1), where the coefficients are non-continuous, highly oscillating and tolerance-periodic functions. Equations (14) together with micro-macro decompositions (3) and (4) constitute *the tolerance model of thermoelasticity problems for transversally graded laminates*. It is necessary to formulate boundary conditions for the macrodisplacements w_i and the macrotemperature θ on the edges $x=0$, H and $x_\alpha=0$, L_α, but for the fluctuation amplitudes of displacements v_i and the fluctuation amplitude of temperature ψ − only on the edges $x_\alpha=0$, L_α.

5.2 The asymptotic model equations

Equations of *the asymptotic model* can be obtained directly from equations of the tolerance model, neglecting underlined terms (involving the microstructure parameter l):

$$\begin{aligned}
&\partial_j(<c_{ijkl}>\partial_l w_k+<c_{ijk1}\partial h>v_k)-<\rho>\ddot{w}_i=\partial<b_{ij}>\theta+<b_{ij}>\partial_j\theta,\\
&<c_{i1kl}\partial h>\partial_l w_k+<c_{i1k1}\partial h\partial h>v_k=<b_{i1}\partial h>\theta,\\
&\partial_j(<k_{ij}>\partial_i\theta+<k_{1j}\partial g>\psi)=<c>\dot{\theta}+<T_0 b_{ij}>\partial_j\dot{w}_i+<T_0 b_{i1}\partial h>\dot{v}_i,\\
&<k_{i1}\partial g>\partial_i\theta+<k_{11}\partial g\partial g>\psi=0,
\end{aligned} \tag{15}$$

The above equations describe the thermoelastic effects in laminates on the macro-level.

6 EXAMPLE

The considered laminate with transversally graded material properties has a height equals H=0.5 m, thickness of the lamina l=0.01 m. It is assumed that the both materials are isotropic and homogeneous. Hence, let us denote by indexes 1, 2 properties of the first and the second material: Young's moduli E_1, E_2, Poisson's ratios ν_1, ν_2, mass densities ρ_1, ρ_2, expansion coefficients α_1, α_2, specific heats c_1, c_2, heat conductions k_1, k_2, Lame's constants λ_1, λ_2, μ_1, μ_2, thermoelasticity constants b_1, b_2. So for the first material (steel) there are: $E_1=210\cdot10^9$[Pa], $\nu_1=0.3$[-], $\rho_1=7800$[kg/m^3], $\alpha_1=10.8\cdot10^{-6}$[1/K], $c_1=460$[J/(kgK)], $k_1=58$[W/(m^{2}K)], $\lambda_1=1.21\cdot10^{11}$[Pa], $\mu_1=8.08\cdot10^{10}$[Pa], $b_1=5.67\cdot10^6$[Pa/K] and for the second material (aluminum): $E_2=69\cdot10^9$[Pa], $\nu_2=0.33$[-], $\rho_2=2720$[kg/m^3], $\alpha_2=23.1\cdot10^{-6}$[1/K], $c_2=920$[J/(kgK)], $k_2=229$[W/(m^{2}K)], $\lambda_2=5.0354\cdot10^{10}$[Pa], $\mu_2=2.5940\cdot10^{10}$[Pa], $b_2=4.6879\cdot10^6$[Pa/K]. The fluctuation shape functions $h(x)$, $g(x)$ are given by formulas (5).

The six different distribution functions of material properties are considered:

$$\begin{aligned}
&\nu'_{n0}=0.5,\ \nu'_{n1}=x_1/H,\ \nu'_{n2}=(x_1)^2/H^2,\ \nu'_{n3}=(x_1)^3/H^3,\\
&\nu'_{n4}=(1-e^{\frac{2x_1}{H}})/(1-e^2),\ \nu'_{n5}=\tfrac{1}{2}\ln[(e^2-1)x_1/H+1].
\end{aligned}$$

The following assumptions are taken into account: stationary, one-dimensional issue, loads are a temperature: $\theta_0|_{x_1=0}=90$°C, $\theta_0|_{x_1=H}=10$°C, and two types of boundary conditions: 1°: $w_1^1\big|_{x_1=0}=0$, $\sigma_1\big|_{x_1=H}=0$, 2°: $w_1^2\big|_{x_1=0}=0$, $w_1^2\big|_{x_1=H}=0$.

Using the tolerance model equations (14) we obtain the equations for considered issue:

$$\begin{aligned}
&\partial(<c_{1111}>\partial w_1+<c_{1111}\partial h>\nu_1)=\partial<b_{11}>\theta+<b_{11}>\partial\theta, && \text{(a)}\\
&<c_{1111}\partial h\partial h>\nu_1+<c_{1111}\partial h>\partial w_1=<b_{11}\partial h>\theta, && \text{(b)}\\
&\partial(<k_{11}>\partial\theta+<k_{11}\partial g>\psi)=0, && \text{(c)}\\
&<k_{11}\partial g>\partial\theta+<k_{11}\partial g\partial g>\psi=0. && \text{(d)}
\end{aligned} \tag{16}$$

Using equations (16d) and (16b) we obtained the formulas for the fluctuation amplitudes of temperature and the fluctuation amplitudes of displacements, respectively:

$$\psi=-\frac{<k_{11}\partial g>\partial_1\theta}{<k_{11}\partial g\partial g>}, \tag{17}$$

$$\nu_1=\frac{<b_{11}\partial h>\theta-<c_{1111}\partial h>\partial_1 w_1}{<c_{1111}\partial h\partial h>}. \tag{18}$$

Then, solving (in analytical way) equations (16c) and (16a), we obtain the formulas for the macrotemperature (after substituting the load conditions) and the macrodisplacements for the 1º and the 2º boundary conditions, respectively:

$$\theta(x_1)=\theta_0+(\theta_H-\theta_0)\frac{k_1x_1+(k_2-k_1)\int_0^{x_1}\nu'(s)ds}{k_1H+(k_2-k_1)\int_0^{H}\nu'(s)ds}, \tag{19}$$

$$\begin{aligned}
w_1^1(x_1)=w_0+(w_H-w_0)&\frac{\int_0^{x_1}<c^{-1}>(s)ds}{\int_0^{H}<c^{-1}>(s)ds}+\\
+\int_0^H\theta(s)<bc^{-1}>(s)ds&\left(\frac{\int_0^{x_1}\theta(s)<bc^{-1}>(s)ds}{\int_0^{H}\theta(s)<bc^{-1}>(s)ds}-\frac{\int_0^{x_1}<c^{-1}>(s)ds}{\int_0^{H}<c^{-1}>(s)ds}\right),
\end{aligned} \tag{20}$$

$$w_1^2(x_1)=w_0+\theta_H<bc^{-1}>(H)\int_0^{x_1}\left[\frac{\theta(s)<bc^{-1}>(s)}{\theta(H)<bc^{-1}>(H)}-\frac{<c^{-1}>(s)}{<c^{-1}>(H)}\right]ds+\sigma_H\int_0^{x_1}\frac{<c^{-1}>(s)}{<c^{-1}>(H)}ds. \tag{21}$$

After taking into account the specific distribution functions of materials, we obtain the formulas and graphs for the displacement and the temperature.
For the distribution function of material $\nu'_{n0}=0.5$:

$$\theta(x_1)=\theta_0+\frac{\theta_H-\theta_0}{H}x_1, \tag{22}$$

$$w_1^1(x_1)=x_1\left[\frac{<b_{11}><c_{1111}\partial h\partial h>-<b_{11}\partial h><c_{1111}\partial h>}{2(<c_{1111}><c_{1111}\partial h\partial h>-<c_{1111}\partial h>^2)}(\theta_H-\theta_0)\ \frac{x_1}{H}+\theta_0\right], \tag{23}$$

$$w_1^2(x_1)=\frac{<b_{11}><c_{1111}\partial h\partial h>-<b_{11}\partial h><c_{1111}\partial h>}{2(<c_{1111}><c_{1111}\partial h\partial h>-<c_{1111}\partial h>^2)}(\theta_H-\theta_0)\ x_1\left(\frac{x_1}{H}-1\right). \tag{24}$$

For the distribution function of material $\nu'_{n1}=x_1/H$:

$$\theta(x_1)=\theta_0+(\theta_H-\theta_0)\frac{k_1x_1+\frac{1}{2}(k_2-k_1)(x_1)^2H^{-1}}{k_1H+\frac{1}{2}(k_2-k_1)H}, \tag{25}$$

$$w_1^i(x_1)=\frac{\mathrm{A}}{c_2}x_1+\mathrm{A}\frac{c_2-c_1}{c_1c_2}\tfrac{1}{2}\frac{(x_1)^2}{H}+\frac{b_2}{c_2}[\theta_0x_1+\frac{\theta_H-\theta_0}{k_1H+(k_2-k_1)\frac{1}{2}H}(\tfrac{1}{2}k_1(x_1)^2+\tfrac{1}{6}\frac{(k_2-k_1)(x_1)^3}{H})]+$$
$$+\frac{b_1c_2-b_2c_1}{c_1c_2}(\tfrac{1}{4}\frac{\mathrm{a}_1}{H}(x_1)^4+\tfrac{1}{3}\frac{\mathrm{a}_2}{H}(x_1)^3+\tfrac{1}{2}\frac{\mathrm{a}_3}{H}(x_1)^2)+\mathrm{B}, \tag{26}$$

where: A, B are constants, which depend on the boundary conditions; $\mathrm{a}_1=\frac{\theta_H-\theta_0}{H^2}\frac{k_2-k_1}{k_1+k_2}$, $\mathrm{a}_2=\frac{2(\theta_H-\theta_0)k_1}{H(k_1+k_2)}$, $\mathrm{a}_3=\theta_0$.

For the distribution function of material $\nu'_{n2}=(x_1)^2H^{-2}$:

$$\theta(x_1)=\theta_0+(\theta_H-\theta_0)\frac{k_1x_1+\frac{1}{3}(k_2-k_1)(x_1)^3H^{-2}}{k_1H+\frac{1}{3}(k_2-k_1)H}, \tag{27}$$

$$w_1^i(x_1)=\frac{\mathrm{A}}{c_2}x_1+\mathrm{A}\frac{c_2-c_1}{c_1c_2}\tfrac{1}{3}\frac{(x_1)^3}{H^2}+\frac{b_2}{c_2}[\theta_0x_1+\frac{\theta_H-\theta_0}{k_1H+(k_2-k_1)\frac{1}{3}H}(\tfrac{1}{2}k_1(x_1)^2+\tfrac{1}{12}\frac{(k_2-k_1)(x_1)^4}{H^2})]+$$
$$+\frac{b_1c_2-b_2c_1}{c_1c_2}(\tfrac{1}{6}\frac{\mathrm{a}_1}{H^2}(x_1)^6+\tfrac{1}{5}\frac{\mathrm{a}_2}{H^2}(x_1)^5+\tfrac{1}{4}\frac{\mathrm{a}_3}{H^2}(x_1)^4+\tfrac{1}{3}\frac{\mathrm{a}_4}{H^2}(x_1)^3)+\mathrm{B}, \tag{28}$$

where: A, B are constants, which depend on the boundary conditions; $\mathrm{a}_1=\frac{\theta_H-\theta_0}{H^3}\frac{k_2-k_1}{2k_1+k_2}$, $\mathrm{a}_2=0$, $\mathrm{a}_3=\frac{3(\theta_H-\theta_0)k_1}{H(2k_1+k_2)}$, $\mathrm{a}_4=\theta_0$.

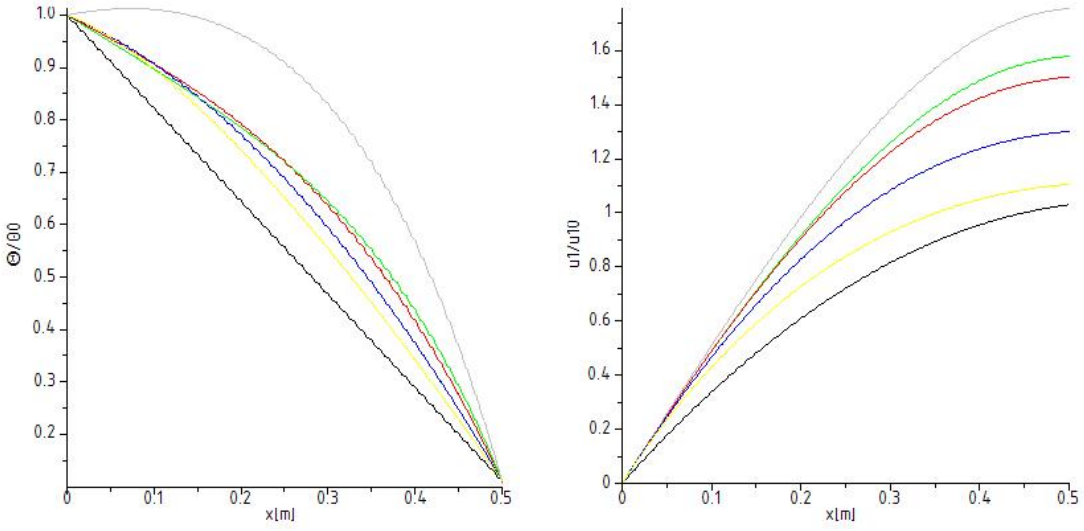


Figure 2: a. The temperature; b. The displacements - for the boundary conditions 1° and various distribution functions of material properties: ν'_{n0}, ν'_{n1}, ν'_{n2}, ν'_{n3}, ν'_{n4}, ν'_{n5}

For the distribution function of material $\nu'_{n3}=(x_1)^3H^{-3}$:

$$\theta(x_1)=\theta_0+(\theta_H-\theta_0)\frac{k_1x_1+\frac{1}{4}(k_2-k_1)(x_1)^4H^{-3}}{k_1H+\frac{1}{4}(k_2-k_1)H}, \tag{29}$$

$$w_1^i(x_1)=\frac{\mathrm{A}}{c_2}x_1+\mathrm{A}\frac{c_2-c_1}{c_1c_2}\tfrac{1}{4}\frac{(x_1)^4}{H^3}+\frac{b_2}{c_2}[\theta_0 x_1+\frac{\theta_H-\theta_0}{k_1H+(k_2-k_1)\frac{1}{4}H}(\tfrac{1}{2}k_1(x_1)^2+\tfrac{1}{20}\frac{(k_2-k_1)(x_1)^5}{H^3})]+$$
$$+\frac{b_1c_2-b_2c_1}{c_1c_2}(\tfrac{1}{8}\frac{\mathrm{a}_1}{H^3}(x_1)^8+\tfrac{1}{7}\frac{\mathrm{a}_2}{H^3}(x_1)^7+\tfrac{1}{6}\frac{\mathrm{a}_3}{H^3}(x_1)^6+\tfrac{1}{5}\frac{\mathrm{a}_4}{H^3}(x_1)^5+\tfrac{1}{4}\frac{\mathrm{a}_5}{H^3}(x_1)^4)+\mathrm{B}, \tag{30}$$

where: A, B are constants, which depend on the boundary conditions; $\mathrm{a}_1=\frac{\theta_H-\theta_0}{H^4}\frac{k_2-k_1}{3k_1+k_2}$, $\mathrm{a}_2=0$, $\mathrm{a}_3=0$, $\mathrm{a}_4=\frac{4(\theta_H-\theta_0)k_1}{H(3k_1+k_2)}$, $\mathrm{a}_5=\theta_0$.

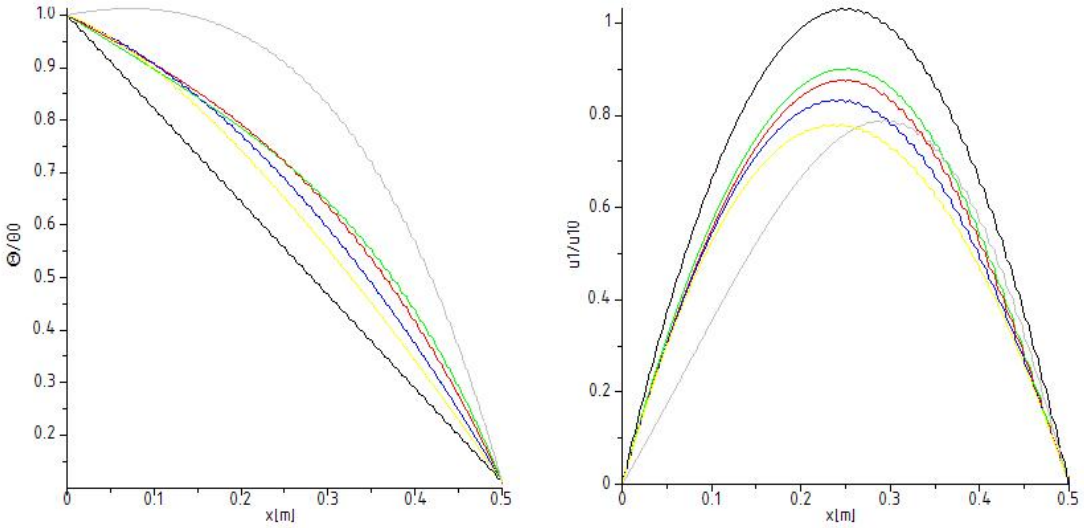


Figure 3: a. The temperature; b. The displacements - for the boundary conditions 2° and various distribution functions of material properties ν'_{n0}, ν'_{n1}, ν'_{n2}, ν'_{n3}, ν'_{n4}, ν'_{n5}

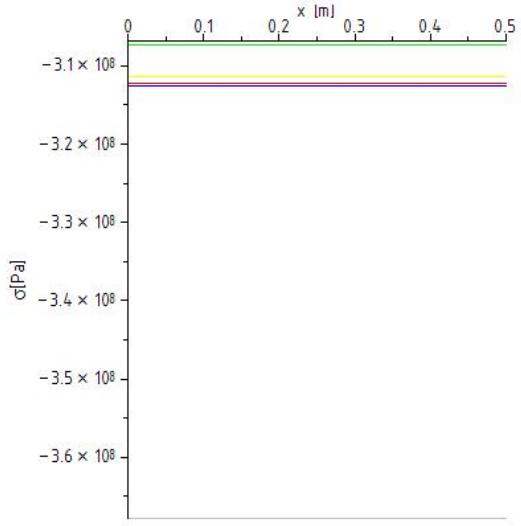


Figure 4: The stresses for the boundary conditions 2°

For the distribution function of material $\nu'_{n4}=(1-\mathrm{e}^{2x_1/H})/(1-\mathrm{e}^2)$:

$$\theta(x_1)=\theta_0+(\theta_H-\theta_0)\frac{k_1x_1+(k_2-k_1)(\,x_1-\frac{1}{2}He^{\frac{2x_1}{H}}+\frac{1}{2}H)}{k_1H+(k_2-k_1)(\frac{3}{2}H-\frac{1}{2}He^2)}. \tag{31}$$

For the distribution function of material $v'_{n5}=\frac{1}{2}\ln[(e^2-1)x_1H^{-1}+1]$:

$$\theta(x_1)=\theta_0+(\theta_H-\theta_0)\frac{k_1x_1+(k_2-k_1)\mathrm{C}}{k_1H+(k_2-k_1)\mathrm{D}}, \tag{32}$$

where:

$$\mathrm{C}=\frac{1}{2}\frac{H+(e^2-1)x_1}{e^2-1}\ln(H+(e^2-1)x_1)-\frac{1}{2}x_1(1+\ln H)-\frac{1}{2}\frac{H}{e^2-1}\ln H,$$

$$\mathrm{D}=\frac{1}{2}\frac{H+(e^2-1)H}{e^2-1}\ln(H+(e^2-1)\mathrm{H})-\frac{1}{2}H(1+\ln H)-\frac{1}{2}\frac{H}{e^2-1}\ln H.$$

Then, the formula for the stresses has the form:

$$\sigma_{11}=<c_{1111}>\partial w_1+<c_{1111}\partial h>v_1-<b_{11}>\theta. \tag{33}$$

The stresses for the 1° boundary conditions are equal zero.

7 REMARKS

Under the above considerations some general remarks can be formulated:

- using *the tolerance modelling* it is possible to replace the differential equations of thermo-elasticity for transversally graded laminates with highly-oscillating, non-continuous coefficients by the differential equations with smooth, slowly-varying coefficients;
- *the tolerance model* equations take into account *the effect of the microstructure size*;
- this effect is omitted in *the asymptotic model* equations;
- the equations for both the models, tolerance and asymptotic, can be applied in the analysis of specific cases, namely where distribution of the ingredients is functional but non-periodic;
- distributions of the temperature and the displacements depend on the distribution functions of material properties.

REFERENCES

[1] Suresh, S. and Mortensen, A. *Fundamentals of functionally graded materials*. The University Press, Cambridge, (1998).

[2] Jędrysiak, J. On the tolerance modelling of thermoelasticity problems for transversally graded laminates. *Arch. Civ. Mech. Eng.* (2011) 11:61−74.

[3] Reiter, T., Dvorak, G.J. and Tvergaard, V. Micromechanical models for graded compositematerials. *J. Mech. Phys. Solids* (1997) 45:1281−1302.

[4] Jikov, V.V., Kozlov, C.M. and Oleinik, O.A. *Homogenization of differential operators and integral functionals*. Springer Verlag, Berlin-Heidelberg, (1994).

[5] Matysiak, S.J. On certain problems of heat conduction in periodic composites. *Z. Angew. Math. Mech.* (1994) 71:524–528.

[6] Kim, K.S. and Noda, N. A Green's function approach to the deflection of a FGM plate

under transient thermal loading. *Arch. Appl. Mech.* (2002) 72:127–137.
[7] Ootao, Y. and Tanigawa, Y. Transient thermoelastic analysis for a functionally graded hollow cylinder. *J. Therm. Stresses* (2006) 29:1031–1046.
[8] Aboudi, J., Pindera, M. J. and Arnolds, M. Higher-order theory for functionally graded materials. *Composites* (1999) 30:777–832.
[9] Bansal, Y. and Pindera, M. J. Efficients reformulation of the thermoelastic higher-order theory for functionally graded materials. *J. Therm. Stresses* (2003) 26:1055–1092.
[10] Chiba, R. Stochastic thermal stresses in an FGM annular disc of variable thickness with spatially random heat transfer coefficients. *Meccanica* (2009) 44:159–176.
[11] Schmauder, S. and Weber, U. Modelling of functionally graded materials by numerical homogenization. *Arch. Appl. Mech.* (2001) 71:182–192.
[12] Goldberg, R.K. and Hopkins, D.A. Thermal analysis of a functionally graded material subject to a thermal gradient using the boundary element method. *Comp. Eng.* (1995) 5:793–806.
[13] Sladek, J., Sladek,. V. and Zhang, Ch. Transient heat conduction analysis in functionally graded materials by the meshless local boundary integral equation method. *Comput. Mater. Sci.* (2003) 28:494–504.
[14] Sadowski, T., Ataya, S. and Nakonieczny, K. Thermal analysis of layered FGM cylindrical plates subjected to sudden cooling process at one side. Comparison of two applied methods for problem solution. *Comput. Mater. Sci.* (2009) 45:624–632.
[15] Woźniak, C. and Wierzbicki, E. *Averaging techniques in thermomechanics of composite solids*, Publishing House of Częstochowa Univ. Techn., Częstochowa, (2000).
[16] Jędrysiak, J. *Thermomechanics of laminates, plates and shields with functional graded properties* (in Polish), Publishing House of Łódź Univ. Techn., Łódź, (2010).
[17] Woźniak, C., et al, (eds.) *Mathematical modelling and analysis in continuum mechanics of microstructured media*, Publishing House of Silesian Univ. Techn., Gliwice, (2010).
[18] Woźniak, C., Michalak, B. and Jędrysiak, J. (eds.) *Thermomechanics of microheterogeneous solids and structures. Tolerance averaging approach*, Publishing House of Łódź Univ. Techn., Łódź, (2008).
[19] Ostrowski,. P. and Michalak, B. Non-stationary heat transfer in hollow cylinder with functionally graded material properties. *J. Theor. Appl. Mech.* (2011) 49.2:385−397.
[20] Ostrowski, P. Thermoelasticity in a two-phase hollow cylinder with longitudinally graded material properties, *Shell Structures. Theory and Applications* (2014) 3:133−136.
[21] Jędrysiak, J. and Radzikowska, A. On the modelling of heat conduction in a non-periodically laminated layer. *J. Theor. Appl. Mech.* (2007) 45:239–257.
[22] Baczyński, Z.F. Dynamic thermoelastic processes in microperiodic composites. *J. Therm. Stresses* (2003) 26:55−66.
[23] Ignaczak, J. A spatial decay estimate for transient thermoelastic process in a composite semispace. *J. Therm. Stresses* (2000) 23:1−14.

AUTHORS INDEX

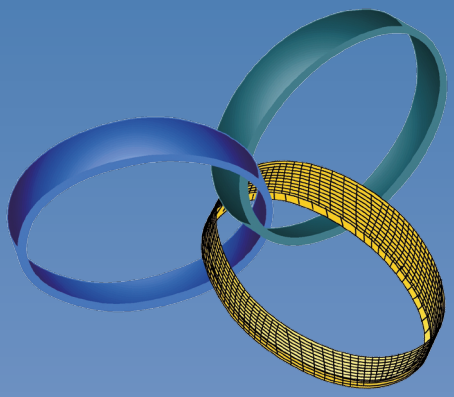
Abbiati, G. 254, 266, 290
Abdel-Gawad, S. 440
Abdelsabour Fahmy, M. 382
Agosti, A. 300
Aguilar, A.M. 1273
Alauzet, F. 1197
Ali, A. 512
Andrushkevich, V. V. 976
Ardianti, A. 1171
Armand, G. 1251
Asfour, S. 35
Asinari, P. 593
Assous, F. 635
Astafev, V.I. 734
Athanassoulis, G.A. 893
Auzins, J. 859
Averkova, O.A. 677, 1291
Axotla, J.G. 1273
Azuhata, T. 278
Baake, E. 657
Baba, A. 1023
Badu, S. 576
Bahai, H. 1037
Barnel, N. 822
Bay, N. 322
Beer, G. 358
Belibassakis, K.A. 893, 1110
Benner, P. 1029
Bensihoum, M. 566
Bermúdez, B. 812
Bertucco, A. 416
Bijl, H. 544, 556
Birken, P. 452
Bittner, K. 1029
Blom, D.S. 544, 556
Bluhm, J. 871
Bochev, P. 23
Boutoudj, M. S. 1101
Brachtendorf, H.-G. 1029
Brandely, A. 1058
Brauer, H. 657
Briffaut, M. 71
Brun, M. 242
Budyn, E. 566
Bursi, O.S. 254, 266
Cai, S-G. 1130
Campana, E.F. 952
Canar, A.E. 677
Cardellini, A. 593
Carter, J.P. 1209
Castellazzi, G. 334
Castiglione, T. 666
Cavallaro, C. 666
Cazzador, E. 266
Ceccato, F. 82
Cerfontaine, B. 1185
Cermak, M. 743
Chabriac, P.A. 59
Charlier, R. 1185
Cherunova, I. 1303
Chiavazzo, E. 593
Chmilenko, F.V. 976
Christopoulos, C. 230
Coco, M. 1003
Collin, F. 1185
Combescure, A. 242
Dal Pont, S, ... 71
Dammann, C. 713
de Miranda, S. 334
Deiterding, R. 489
Deleu, F. 1029
Demidovich, V.B. 976
Deprez, E. 566
Dervieux, A. 1197
Deteix, J. 1138
Dhone, M. 1303
Di Barba, P. 404
Didier, E. 1159
Diebels, S. 94
Dieudonné, A-C. 1185
Diez, M. 137, 952
Diez, P. 1123
Doi, Y. 1171
Drahoš, P. 1015
Duenser, C. 358
Dufour, F, ... 71
Dughiero, F. 312, 404
Duriez, J. 1232
Eken, A. 116
Emamy, N. 753
Engels-Putzka, A. 645
Ertl, T. 544
Ezzedine El Dandachy, M. 71
Fabbri, A. 59
Fasano, M. 593
Favier, J. 1130
Feng, J.J. 725
Feng, L. 1029
Fernandes, O. 544
Ferrari, M. 583
Fiorenzato, V. 416
Fontana, M. 222
Formaggia, L. 300
Formica, G. 334
Fortin, A. 1138
Forzan, M. 312, 404

Frey, C. 645
Frey, S. 544
Fujii, S. 1171
Fukuda, K. 760, 772
Galunin, S. 392
García, O.R. 1273
García Anguita, F.J. 783
Garzon, M. 689
Gauci, E. 1197
Gerasimov, K.V. 191
Ghasemi, F. 464
Giomo, M. 416
Giovanardi, B. 300
Gourdon, E. 59
Gourvenec, S. 1218
Gravouil, A. 242
Gray, L.J. 689
Grenestedt, J. 137
Guarnieri, M. 416
Guerrero, W.F. 812
Guilmineau, E. 1089
Gustin, A. 181
Göransson, P. 881
Görke, U-J. 149
Halama, R. 743
Hammelmüller, F. 1080
Hashiguchi, K. 837, 1243
Hassanizadeh, S.M. 106
Hatic, V. 181
Hinata, K. 849
Hirschberg, A. 930
Hoarau, Y. 1130
Horie, T. 1023
Horie, Y. 849
Hrabovský, J. 993, 1015
Hu, F. 964
Huang, X. 230
Humer, A. 1080
Huybrechts, N. 1147
Huyghe, J.M. 106
Hüppe, A. 524
Iakovlev, S. 1123
Inukai, M. 278
Ishihara, D. 1023
Ishihara, M. 985
Jędrysiak, J. 1324
Janssen, R. 1029
Jaros, M. 125
Jendoubi, A. 1138
Jiang, Z. 701, 964
Jonszta, T. 125
Jung, A. 94
Kabanikhin, S. 921
Kaltenbacher, M. 346, 524
Kameo, Y. 985
Karasek, T. 125, 743
Karcher, M. 753
Kardani, M. 1209
Karperaki, A.E. 893, 1110
Kasatkin, A.E. 734
Kashima, T. 278
Kawada, T. 614
Khennane, A. 1312
Kim, D-H. 137
Kireenkov, A.A. 203
Klimach, H. 556
Knobloch, M. 222
Kojic, M. 583
Kolditz, O. 149
Kondratyuk, A. 1047
Kornev, N. 1303
Kosenko, I.I. 191
Kosnik, N. 181
Kozubek, T. 125
Krivorotko, O. 921
Krupp, V. 556
Kuehnel, W. 524
Kurihara, K. 625
Kutiš, V. 993, 1015
Kwon, O-H. 230
Labbé, O. 501
Langer, S. C. 370
Lattari, K. 1312
Le Blanc, A. 635
Le Marrec, L. 930
Lee, S. 605
Lefrancois, E. 1058
Lemoine, B. 930
León, F.R. 1273
Leotardi, C. 952
Leroyer, A. 1089
Li, Y. 1261
Li, S. 701
Linde, F. 1147
Lindner, F. 477
Lindner, B. 358
Logachev, I.N. 677, 1291
Logachev, K.I. 677, 1291
Loseille, A. 1197
Mackie, K.R. 290
Madani, S.H. 1037
Mahnken, R. 713
Majorana, A. 1003
Majorana, C.E. 47
Mannemare, F. 566
Marchevsky, I.K. 532

Marchevsky, I.K. 800
Markolefas, S.I. 1110
Marsan, T. 566
Martins, P.A.F. 322
Mascali, G. 1003
Matharu, H. 428
Matsumoto, J. 625
Matsumoto, S. 625
Mavric, B. 160, 181
Mehl, M. 477
Melnik, R. 576
Michaelides, A.M. 428
Milosevic, M. 583
Miyashita, K. 1243
Mizuno, Y. 760, 772
Moe, S. 23
Mojiri, S. 230
Molari, L. 334
Molina, F.J. 266
Monge, A. 452
Morel, J.C. 59
Moreva, V.S. 800
Moro, F. 416
Mousavi, R. 753
Mramor, K. 169
Muramatsu, M. 614
Murín, J. 993, 1015
Muscato, O. 666
Mutsuda, H. 1171
Nacke, B. 392
Nagahashi, K. 1072
Nagata, T. 760, 772
Nagel, T. 149
Nakao, W. 849
Nastase, A. 939
Nazem, M. 1209
Neuenschwander, M. 222
Ngo, D.T. 59
Nielsen, C.V. 322
Niho, T. 1023
Nikanorov, A. 392, 657
Nikooee, E. 106
Nita, S. 1072
Nonomura, T. 760, 772
Nordström, J. 464
Nowak, S. 791, 909
Oberlack, M. 753
Ootao, Y. 985
Osada, T. 849
Ota, H. 1072
Ouahsine, A. 1130, 1147
Ozaki, S. 849
Ozbulut, M. 725
Palermo, D. 822
Paolacci, F. 266
Papadimitriou, C. 1282
Papadimitriou, D.I. 1282
Papathanasiou, T.K. 893, 1110
Paulech, J. 993, 1015
Pazera, E. 1324
Pegon, P. 266
Peng, K. 701
Peterson, K. 23
Petite, H. 566
Peyret, C. 501
Pineda, M.B. 1273
Prabhakar, S. 576
Pulch, R. 1029
Puzikova, V.V. 532, 800
Queutey, P. 1089
Radu, J-P. 1185
Rahmat, A. 725
Rammerstorfer, F. G. 346
Randolph, M.F. 1218
Rastvorova, I.I. 976
Reimann, T. 512
Reppenhagen, A. 524
Ridzal, D. 23
Riley, C.P. 428
Rojicek, J. 743
Roller, S. 556
Romano, V. 1003
Rosati, L. 35
Rubio, J. 1123
Rumpler, R. 881
Sadaoui, A. 1312
Sadat-Hosseini, H. 137
Sahin, M. 116
Saito, T. 278
Sakharov, A. 210
Sallet, F. 59
Salomoni, V. 47, 82
Sánchez-Badorrey, E. 783
Šarler, B. 160, 169, 181
Sasnouski, S. 566
Sato, S. 1243
Sato, T. 1072
Sawant, N.P. 428
Schauer, M. 370
Scheufele, K. 477
Schlesselmann, D. 392
Schoenmaker, W. 1029
Schröder, J. 871
Schulthess, P. 222
Schäfer, M. 512, 791, 909, 1047
Schörgenhumer, M. 1080

Scotti, A. 300
Sergent, P. 1147
Serpieri, R. 35
Sethian, J.A. 689
Seyedi, D. 1251
Sieni, E. 404
Simic, V. 583
Simonini, P. 82
Skukis, E. 859
Smaoui, H. 1130
Specht, S. 871
Spezzapria, M. 312
Spiezia, N. 47, 82
Staab, D. 791
Stavrovskiy, M.E. 191
Stern, F. 137
Sternel, D. C. 512, 791, 909
Stojadinovic, B. 254, 290
Strakos, P. 125
Sweijen, T. 106
Takada, N. 625
Takahashi, S. 760, 772
Takase, S. 614
Tauc, P. 566
Tebbiche, H. 1101
Teixeira, P.R.F. 1159
ter Maten, J. 1029
Terada, K. 614
Thodal, R. 137
Tischendorf, C. 1029
Tofighi, N. 725
Tondini, N. 254
Topping, P.J. 428
Toth, F. 346
Travascio, F. 35
Trenchea, C. 1261
Ubertini, F. 334
Uekermann, B. 477
van Zuijlen, A.H. 544, 556
Vassilev, D.H. 428
Vavra, P. 125
Venskus, A. 428
Vertnik, R. 169, 181
Visonneau, M. 1089
Volpi, S. 137
Vu, M-N. 1251
Wale, J.D. 428
Wan, R. 1232
Wang, C. 440
Wang, D. 964, 1218
Wang, X.D. 440
Wearing, E. 428
Weil, C. 657
Wheeler, M.F. 605
Whyte, C. 254, 290
Wick, T. 605
Wieers, A. 1029
Wissink, J. 1037
Wong, H. 59
Wood, S.L. 489
Wu, Z. 964
Yashiro, K. 614
Yildiz, M. 725
Yu, H.C. 440
Yvin, C. 1089
Zehetner, C. 1080
Zhang, W. 322, 701
Ziemys, A. 583

This volume contains the full papers of the VI International Conference on Coupled Problems in Science and Engineering (Coupled 2015) held in San Servolo, Venice, Italy on May 18 - 20, 2015.



CIMNE
International Center
for Numerical Methods in Engineering