ICCM Proceedings

Proceedings
of the International Conference
on Computational Methods

(Vol.1, 2014)

5th ICCM2014, 28th-30th July 2014, Fitzwilliam College
Cambridge, England

Editors: **G. R. Liu and Z. W. Guan**

ICCM2014

Proceedings of the International Conference on Computational Methods (Vol.1, 2014)

5th ICCM2014, 28th-30th July 2014, Cambridge, UK

Edited by

G. R. Liu
University of Cincinnati, USA

Zhongwei Guan
University of Liverpool, UK

PREFACE

Dear Colleagues

On behalf of the organising committees, we are delighted to welcome you to the 5th International Conference on Computational Methods (ICCM2014) at Cambridge, UK.

The ICCM2014 is an international conference that provides an international forum for the exchange of ideas on recent advances in areas related to computational methods, numerical modelling & simulation, as well as their applications in engineering and science. It will accommodate presentations on a wide range of topics to facilitate inter-disciplinary exchange of ideas in science, engineering and related disciplines, and foster various types of academic collaborations internationally. All papers accepted for publication in the proceedings have been peer reviewed. Papers may also be selected and invited to be developed into a full journal paper for publication in special issues of the journals.

The conference series originated in Singapore in 2004 by Professor GR Liu, followed by ICCM2007 in Hiroshima, Japan, ICCM2010 in Zhangjiajie, China, and ICCM2012 in Golden Coast, Australia. This year, the ICCM2014 conference programme covers over 270 oral presentations in 47 technical sessions, including 3 Plenary talks, 9 Thematic Plenary talks, and a number of Keynote talks in technical sessions. These presentations cover a broad range of topics related to computational mechanics, including formulation theory, computational methods and techniques, modelling techniques and procedures, materials, deformation processing, materials removal processes, processing of new and advanced materials, welding and joining, surface engineering and other related processes.

We would like to express my gratitude to all the members of the Local Organizing Committee, International Scientific and Organization Committee, Honorary Chairmen and Co-Chairmen, who have provided advices and guidance timely in planning and executing this conference. We also would like to use this opportunity to express my gratitude to the School of Engineering of the University of Liverpool, and to colleagues for their strong support and encouragement. Sincere thanks and appreciation go to some 100 international reviewers for their prompt review reports on the submitted papers. Our appreciation goes also to all the Mini-Symposium Organizers for their efforts and contributions in the organization. A vote of thanks also goes to members at the Scientech Publisher, USA for their professional services and management of the conference website and timely coordination with our participants.

G.R. Liu, University of Cincinnati, USA
Zhongwei Guan, University of Liverpool, UK

ORGANIZATION COMMITTEES

Local Organization Committee

Chairmen: Z.W. Guan (University of Liverpool)

Co-Chairmen: Ivan Au (University of Liverpool), Michael Beer (University of Liverpool), Shuguang Li (University of Nottingham), Carlo Sansour (University of Nottingham), Jianqiao Ye (University of Lancaster)

Jiye Chen (University of Portsmouth), Steve Jones (University of Liverpool), Ioannis Kougioumtzoglou (University of Liverpool), Ming Li (University of Liverpool), Alex Diaz De La O (Secretary, University of Liverpool), Edoardo Patelli (University of Liverpool), Jianguo Zhou (University of Liverpool)

Honorary Chairmen

Gui-Rong Liu (USA)   Andy Long (UK)   Dai-Ning Fang (China)   Mingwu Yuan (China)

International Organization Committee

Chairman: Z.W. Guan (UK)

Co-Chairmen

Xi-Qiao Feng (China)	Hiroshi Okada (Japan)	Ming Li (China)	Sung-Kie Youn (Korea)
Yuantong Gu (Australia)	Carlo Sansour (UK)	Jinsong Leng (China)	Weihong Zhang (China)

- Leonardo Alves (Brazil)
- Jorge A. C. Ambrósio (Potugal)
- Soowon Chae (Korea)
- Jiye Chen (UK)
- Weiqiu Chen (China)
- Changzheng Cheng (China)
- Maeng Hyo Cho (Korea)
- Marcelo Colaco (Brazil)
- Roberto Contro (Italy)
- Fangsen Cui (Singapore)
- Guangze Dai (China)
- Raj Das (New Zealand)
- George S. Dulikravich (USA)
- Xu Guo (China)
- Wanlin Guo (China)
- Wei Guo (Australia)
- Zaoyang Guo (China)
- S. Hagihara (Japan)
- Xu Han (China)
- Annette Harte (Ireland)
- Muneo Hori (Japan)
- Ning Hu (China)
- Zheng-Ming Huang (China)
- Hung Nguyen Xuan (Vietnam)
- Chao Jiang (China)
- Hiroshi Kanayama (Japan)
- H. Kawai (Japan)
- Cheol Kim (Korea)
- Moon Ki Kim (Korea)
- Yoon Young Kim (Korea)
- Yun-jae Kim (Korea)

- Alexander Korsunsky (UK)
- S. Koshizuka (Japan)
- Oh Joon Kwon (Korea)
- Ik-Jin Lee (Korea)
- Chenfeng Li (UK)
- San-Yih Lin (Taiwan)
- Moubin Liu (China)
- Yijun Liu (USA)
- Yinghua Liu (China)
- Zhaomiao Liu (China)
- Lianchun Long (China)
- Ping Lu (USA)
- Hitoshi Matsubara (Japan)
- Karol Miller (Australia)
- Toshio Nagashima (Japan)
- Yuji Nakasone (Japan)
- Perumal Nithiarasu (UK)
- Yang-Yao Niu (Taiwan)
- Zhongrong Niu (China)
- Masao Ogino (Japan)
- Makoto Ohsaki (Japan)
- Takeshi Omori (Japan)
- Marc Oudjene (France)
- Joe Petrolito (Australia)
- Dong Qian (USA)
- Qinghua Qin (Australia)
- Alessandro Reali (Italy)
- Božidar Šarler (Slovenia)
- Sandeep M Shiyekar (India)
- Leandro Alcoforado Sphaier (Brazil)
- Yoshio Suzuki (Japan)

- Nasser Hassan Sweilam (Egypt)
- Yuichi Tadano (Japan)
- Shintaro Takeuchi (Japan)
- Y. Cengiz Toklu (Turkey)
- Biao Wang (China)
- Jizeng Wang (China)
- Yue-Sheng Wang (China)
- Hengan Wu (China)
- Mao S. Wu (Singapore)
- Yang Xiang (Australia)
- Yi-Min (Mike) Xie (Australia)
- Qing-sheng Yang (China)
- Weian Yao (China)
- Zhenjun Yang (China)
- K Yashiro (Japan)
- Hongling Ye (China)
- Yongjin Yoon (Singapore)
- S. Yoshimura (Japan)
- Rena C. Yu (Spain)
- Julien Yvonnet (France)
- Kang Zhan (China)
- Xiong Zhang (China)
- Qing Zhang (China)
- Lihai Zhang (Australia)
- Xianfeng Zhang (China)
- Zhiqian Zhang (Singapore)
- Ya-Pu Zhao (China)
- Zheng Zhong (China)
- Huanlin Zhou (China)
- Kun Zhou (Singapore)
- Yufeng Zhou (Singapore)
- Jihong Zhu (China)

CONTENTS

ODE-Solver-Oriented Computational Method for the Structural

Dynamic Analysis of Super Tall Buildings

*Yaoqing Gong[1], and Xiancheng Wang[2]

[1]Yaoqing Gong, School of Civil Engineering, Henan Polytechnic University, China
[2]Department of Civil and Material Engineering, Henan University of Urban Construction, China

*Corresponding author: gongyq@hpu.edu.cn

Abstract

The paper is to introduce a computational methodology that is based on ordinary differential equations (ODE) solver for the structural systems adopted by super tall buildings in their preliminary design stage so as to facilitate the designers to adjust the dynamic properties of the adopted structural systems. The construction of the study is composed by following aspects. The first aspect is the modelling of a structural system. As a typical example, a mega frame-core-tube structural system adopted by some famous super tall buildings such as Taipei 101 building, Shanghai World financial center, is employed to demonstrate the modelling of a computational model. The second aspect is the establishment of motion equations constituted by a group of ordinary differential equations for the analyses of free vibration and resonant response. The solutions of the motion equations (that constitutes the third aspect) resorted to ODE-solver technique. Finally, some valuable conclusions are summarized.

Keywords: ODE-solver-oriented computational methodology, tall building structures, structural dynamic analysis, computational model of a mega frame-core-tube structural system, free vibration and resonant response, ODE solver

Introduction

Nowadays, we are experiencing an unprecedented level of activity in the design and construction of super tall buildings because of the limitation of land resources and advanced construction technology, ad hoc in China [X. zhao et al. (2011)]. The world architecture history has been rewritten by the multiformity of structural systems, the complexity of component arrangements and the variation of architectural styles of current super tall buildings. However, the analytical level for the investigation of dynamic properties of various structural systems adopted by super tall buildings lags behind their construction level. Both computational models and numerical methods for the dynamic analyses proposed hitherto by existing literatures are quite limited in their ability to model and to determine the three-dimensional motion of the structural systems.

For instance, Reza Kamgar, Mohammad Mehdi and Saadatpour [Reza et al. (2011)] developed a simple mathematical model based on Euler-Bernoulli beam theory to determine the first natural frequency of tall buildings including a framed tube, a shear core, a belt truss and an outrigger system with multiple jumped discontinuities in the cross section of the framed tube and shear core. Hong Fan, Q.L. Li, Alex Y. Tuan and Lihua Xu [Hong Fan et al. (2009)] investigated the seismic analysis of the structural system of Taipei 101, a mega-frame system with a central braced core connected to perimeter columns on each building face, by employing a 5-storey frame computational model composed by 3-D beams, 3-D columns and floor slabs. Wen-Hae Lee [Lee (2007)] simplified a tube-in-tube tall-building system as an Euler-

Bernoulli beam with variable cross-sections and then formulated an approximate solution procedure for the free vibration analysis.

In order to render the computational model of a super tall building system closer to practical engineering as well as the corresponding numerical method more efficient, the purpose of this paper is to present an ODE-solver-oriented computational methodology for the structural systems adopted by super tall buildings in their preliminary design stage so as to facilitate the designers to adjust the dynamic properties of the adopted structural systems. The construction of the study is composed by following aspects. The first aspect is the modelling of a structural system adopted by a super tall building. As a typical example, a mega frame-core-tube structural system as showed in Figure 1(a) adopted by some famous super tall buildings such as Taipei 101 building, Shanghai World financial center, is employed to demonstrate the modelling of a computational model. The second aspect is the establishment of motion equations constituted by a group of ordinary differential equations (ODE) for the analyses of free vibration and resonant response. The establishment utilized semi-discretization, displacement quantification and motion-field quantification techniques. The solutions of the motion equations (that constitutes the third aspect) resorted to an ODE solver technique (Yuan Si [Yuan (1991, 1993)]). Finally, some valuable conclusions are summarized.

1. Modelling of a super-tall building system

Figure 1(a) shows a mega frame-core-tube system adopted by some famous super tall buildings such as Taipei 101 building, Shanghai World financial center, etc. On structural aspects, the space mega frame is composed by two grades of members. The first grade is mega columns and beams, and the second grade is interiorly supplementary frames in the mega frame. The mega columns are generally made by tubes or other mega-substructures, which are jointed by the giant beams in every several floors. Since the geometric dimension (cross sectional area and inertial moment, etc.) of the members of the mega frame is very large, comparing with that of the supplementary ones, the characteristic makes this kind of structure has great load bearing capacity, strong sidesway stiffness. By analyzing the structural performance of the mega frame-core-tub system shown in Figure 1(a), we can conduct following two basic assumptions:

(1) Rigid floor slab assumption, that is, each floor is infinite rigid in its own plane;
(2) Strain state assumption, that is, the axial strain of a mega beam is negligible comparing with that of a mega column.

(a)　　　　(b)
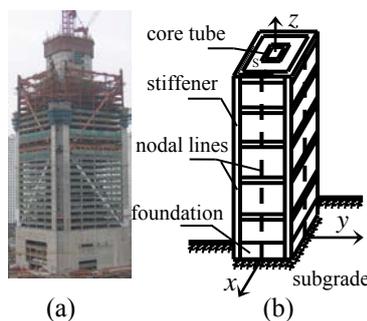
Figure 1: A mega-frame-core-tube system and its computational model
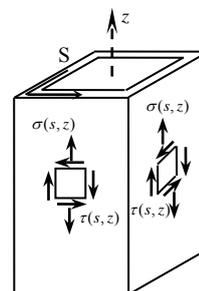
Figure 2: Stress state of the tube

Based on the two assumptions, we might simplify a mega frame-core-tube system shown in Figure 1(a) as a generalized equivalent continuous stiffened thin-walled tube-in-tubes as shown in Figure 1(b), and conclude that the wall of the thin-walled tubes is subjected to a plane stress state of longitudinal normal stress and horizontal as well as vertical shear stress, which are the functions with respect to the curved coordinate S, along the direction of centerline of the thin-walled tubes, and the longitudinal coordinate (vertical axis Z) of the tubes, as showed in Figure 2.

The subgrade or the foundation soil of the structural system is idealized as a semi-infinite elastic body, whose equivalent stiffness equations in the normal and tangential direction at the bottom and walls of a foundation pit have been formulated by employing the displacement equations of Mindlin [Mindlin (1936)] in a semi-infinite elastic body subject to a concentrated force acting in the interior of the semi-infinite elastic body. Using the equivalent stiffness equations [Gong (2007)], the interactions between the foundation and the subgrade (foundation soil) can be readily taken into considerations quantitatively.

Consequently, the computational model of a mega frame-core-tube system of a super tall building will be a generalized equivalent continuous tubular shell constituted by stiffened thin-walled tubes-in-tubes supported on a semi-infinite elastic body as shown in Figure 1(b).

2. Formulation of motion equations

2.1 *Semi-discretization technique and displacement quantification*

As shown in Figure 1(b), if we use one-variable functions $v_{0x}(z)$, $v_{0y}(z)$ and $\theta(z)$ defined on the vertical axis Z of the tubular shell, which are piecewise functions in most cases, the transverse displacements of the cross section of the tubular shell in the X and Y directions, and the rotation around the longitudinal axis Z will be represented as

$$\{v_0(z)\} = \left\{ \begin{Bmatrix} v_{0x}(z) \\ v_{0y}(z) \\ \theta(z) \end{Bmatrix}_j \right\}. \tag{1}$$

Similarly, if we make a semi-discretization along the cross-section central line S by the vertical lines named nodal lines and employ the one-variable functions $w_j^{in}(z)$, and $w_j^{ex}(z)$, respectively defined on the inner and outer nodal lines, and interpolation functions $\varphi_j(s_{in})$, and $\varphi_j(s_{ex})$ between the inner and outer nodal lines, the axial displacement or longitudinal warping of the tubular shell will be expressed as

$$\{u(s,z)\} = \begin{Bmatrix} [\varphi(s_{in})]\{w_{in}(z)\} \\ [\varphi(s_{ex})]\{w_{ex}(z)\} \end{Bmatrix}_j. \tag{2}$$

where $j = 1, 2, \cdots n$ is the segment number of the nodal lines in the longitudinal direction Z (1 may represent foundation and 2 to 6 may stand for the first to fifth floors and so on, for example), and the segment number depending upon the property variation of the building system up the height is the intersection number between nodal lines and the curvilinear coordinate (central line of the cross section) S; $\{v_0(z)\}$ and $\{u(s,z)\}$ are function sets, constituted by all of the basic unknown functions; $[\varphi(s)]$ a row vector, and $\{w(z)\}$ a column vector, respectively.

2.2 *Motion-field quantification*

For free vibration analysis, the longitudinal and transverse dynamic displacements of the structural system (the tubular shell) can be respectively written in Galerkin's form as

$$\{u(s,z,t)\} = \left\{ \begin{Bmatrix} [\varphi(s_{in})]\{w_{in}(z)\} \\ [\varphi(s_{ex})]\{w_{ex}(z)\} \end{Bmatrix}_j e^{i\omega t} \right\}, \tag{3}$$

$$\{v_0(z,t)\} = \left\{ \{v_0(z)\}_j e^{i\omega t} \right\}. \tag{4}$$

For forced vibration steady-state response analysis, if giving an arbitrary vertical ground-motion of $\{u_g(t)\}$, an arbitrary horizontal ground-motion of $\{T_g(t)\}$ in the X

and Y directions, and the rotation around the longitudinal axis Z, for instance, the motion field of the computational model can be readily quantified as follow

$$\left.\begin{aligned}\{u^t(s,z,t)\} &= \{u_g(t)\} + \{u(s,z)\}r(t)\\\{v_o^t(z,t)\} &= \{T_g(t)\} + [f(t)]\{v_0(z)\}\end{aligned}\right\}, \tag{5}$$

where $r(t)$ and $f(t)$ are the time functions concluded by means of the time-change law of $u_g(t)$ and $T_g(t)$, respectively.

2.3 *Motion equations or governing equations*

By employing above motion field, the total kinetic energy as well as the potential energy of the structural system including the strain energy stored in the subgrade can be readily estimated. Then, by using a Hamiltonian principle, the governing equations of the structural system can be derived conveniently, which are the ordinary differential equations (ODE) and corresponding boundary conditions. For instance, the motion equations for free vibration will lead to

$$\left.\begin{aligned}\{F_s^u\}_1^{in} + \{F_i^u\}_1^{in} &= \{0\}, \qquad \{F_s^u\}_1^{ex} + \{F_i^u\}_1^{ex} = \{0\}\\\{F_s^v\}_1^{in} + \{F_s^v\}_1^{ex} + \{F_i^v\}_1^{in} + \{F_i^v\}_1^{ex} - \{F_r^t\}_1^{ex} &= \{0\}\end{aligned}\right\}, \tag{6}$$

$$\left.\begin{aligned}\{F_s^u\}_j^{in} + \{F_i^u\}_j^{in} &= \{0\}, \quad \{F_s^u\}_j^{ex} + \{F_i^u\}_j^{ex} = \{0\}\\\{F_s^v\}_j^{in} + \{F_s^v\}_j^{ex} + \{F_i^v\}_j^{in} + \{F_i^v\}_j^{ex} &= \{0\}\end{aligned}\right\}, \tag{7}$$

in which,

$$\{F_s^u\} = E[A]\{w''(z)\} - G[B]\{w(z)\} - G[C]\{v_0'(z)\}, \quad \{F_i^u\} = m\omega^2[A]\{w(z)\},$$

$$\left.\begin{aligned}\{F_s^v\} &= G[D_1]\{v_0''(z)\} + G[C]^T\{w'(z)\}, \{F_i^v\} = m\omega^2[D_2]\{v_0(z)\},\\\{F_r^t\} &= C_r k_{nH}[E]\{v_0(z)\}\end{aligned}\right\}.$$

Equations (6) and (7) are the motion equations for the foundation and other segments of the computational model respectively, and their corresponding boundary conditions at the bottom of the foundation will be

$$\left.\begin{aligned}\{E[A]\{w'(0)\} = k_{zD}[A]\{w(0)\}\}^{in}, \quad \{E[A]\{w'(0)\} &= k_{zD}[A]\{w(0)\}\}^{ex}\\\left[(G[D_1])^{in} + (G[D_1])^{ex}\right]\{v_0'(0)\} + \left\{(G[C]^T\{w(0)\})^{in}\right\}&\\+ \left\{(G[C]^T\{w(0)\})^{ex}\right\} + k_{tD}[S]\{v_0(0)\} &= \{0\}\end{aligned}\right\}. \tag{8}$$

The boundary conditions at the top of the computational model become as

$$\left.\begin{aligned}\{E[A]\{w'(H)\}\}^{in} = \{0\}, \quad \{E[A]\{w'(H)\}\}^{in} &= \{0\}\\\left[(G[D_1])^{in} + (G[D_1])^{ex}\right]\{v_0'(H)\} + \{(G[C]^T\{w(H)\})^{in} + (G[C]^T\{w(H)\})^{ex}\} &= \{0\}\end{aligned}\right\}. \tag{9}$$

Also the displacement consistence and generalized internal force equilibrium conditions at each connection of the computational model must be

$$\left.\begin{aligned}\{w^{in}(H_k)\}_k &= \{w^{in}(H_k)\}_{k+1}, \{w^{ex}(H_k)\}_k = \{w^{ex}(H_k)\}_{k+1}\\\{v_0(H_k)\}_k &= \{v_0(H_k)\}_{k+1}\end{aligned}\right\}, \tag{10}$$

$$\left.\begin{array}{l} \left\{E[A]\{w'(H_k)\}\right\}_k^{in} = \left\{E[A]\{w'(H_k)\}\right\}_{k+1}^{in}, \\[2mm] \left\{E[A]\{w'(H_k)\}\right\}_k^{ex} = \left\{E[A]\{w'(H_k)\}\right\}_{k+1}^{ex}, \\[2mm] \left\{[(G[D_1])^{in} + (G[D_1])^{ex}]\{v_0'(H_k)\}\right\}_k + \left\{G[C]^T\{w(H_k)\}\right\}_k^{in} \\[2mm] \quad + \left\{G[C]^T\{w(H_k)\}\right\}_k^{ex} = \left\{[(G[D_1])^{in} + (G[D_1])^{ex}]\{v_0'(H_k)\}\right\}_{k+1} \\[2mm] \quad + \left\{G[C]^T\{w(H_k)\}\right\}_{k+1}^{in} + \left\{G[C]^T\{w(H_k)\}\right\}_{k+1}^{ex} \end{array}\right\}. \qquad (11)$$

The meaning of the matrices such as [A], [B], etc. can be referred to [Gong (2010)]. It is observed, in mathematic view, that the problem about the free vibration of a super tall building system is an eigenvalue problem, and its governing ordinary differential equations (ODE) can be theoretically solved by an ODE solver such as COLSYS [Ascher (1981)], a general purpose program developed to solve various ODE problems. However, the normal ordinary differential equation solver can only solve the standard ODE problem. Consequently, a computational software package known as EIGENCOL [Yuan (1991, 1993)] has been developed to solve the eigenvalues and corresponding modes efficiently [Yaoqing Gong (2010)].

3. ODE-Solver Method

As mentioned previously, the free vibration of a super tall building system is an eigenvalue problem of a group of ordinary differential equations, which can be theoretically solved by an ODE solver. However, a normal ordinary differential equation solver can only solve a standard ODE problem. In order to find the eigenvalues, a computational software package known as EIGENCOL [Yuan (1991, 1993)] has been developed to solve the eigenvalues and corresponding modes efficiently. According to the technique proposed in the literatures, before the ordinary differential equations with eigenvalues are solved, they should be transformed into the standard ODE forms accepted by COLSYS [Ascher (1981)]. The procedure includes following steps.

3.1 *Coordinate transformation*

The solving interval of standard ordinary differential equations must be $[0,1]$. Thusly, the coordinate transformation must be performed for a practical problem with the solving interval of $[0, H_j]$, for example. At this point, the transformation technique will be

$$\xi = x / H_i, \qquad \frac{d(\ )}{dx} = \frac{d(\ )}{d\xi} \cdot \frac{d\xi}{dx} = \frac{1}{H_i}\frac{d(\ )}{d\xi}.$$

3.2 *Trivial ODE conversion technique*

Because eigenvalues are undetermined constants and also a part of the solution of a group of ODEs, the determination of the unknown constants become a key point for the solution of the group of ODEs. Therefore, a trivial ODE is necessary to convert the ODEs with eigenvalues into a new set of standard ODEs in which an eigenvalue, say $\omega 2$, has been made as an unknown function. In view of the derivative of a constant is zero, the trivial ODE can be thusly established as

$$\lambda' = d(\omega^2) / d\xi = 0. \qquad (12)$$

The addition of equation (12) will lead to one more corresponding boundary condition. Finding the condition introduces another technique, equivalent ODE technique.

3.3 *Equivalent ODE technique*

If we define a normalized function with respect to the forgoing mentioned basic unknown functions as

$$R(\xi) = \frac{\int_0^\xi \left(\{w(\varsigma)\}_{in}^T \{w(\varsigma)\}_{in} + \{w(\varsigma)\}_{ex}^T \{w(\varsigma)\}_{ex} + \{v_0(\varsigma)\}^T \{v_0(\varsigma)\} \right) d\varsigma}{H^2} , \quad (13)$$

where H is the total height of the structural system. Equation (13) can be recognized as a generalized inner production of the basic unknown functions, and obviously

$$\frac{dR}{d\xi} = R'(\xi) = \frac{\left(\{w(\xi)\}_{in}^T \{w(\xi)\}_{in} + \{w(\xi)\}_{ex}^T \{w(\xi)\}_{ex} + \{v_0(\xi)\}^T \{v_0(\xi)\} \right)}{H^2} . \quad (14)$$

Also if we set

$$R(1) = \frac{\int_0^1 \left(\{w(\varsigma)\}_{in}^T \{w(\varsigma)\}_{in} + \{w(\varsigma)\}_{ex}^T \{w(\varsigma)\}_{ex} + \{v_0(\varsigma)\}^T \{v_0(\varsigma)\} \right) d\varsigma}{H^2} = 1, \quad (15)$$

the equation will become a standard normalized condition, and we can find two useful boundary conditions as follows

$$\left. \begin{aligned} R(0) &= 0 \\ R(1) &= 1 \end{aligned} \right\} . \quad (16)$$

By employing above trivial ODE conversion and equivalent ODE techniques, one can transform ordinary differential equations with eigenvalues into a new group of standard ODEs. For instance, equations (6), (7), (12) and (14) constitute a group of standard ODEs, and equations (8), (9), (10), (11) and (16) become their corresponding boundary conditions. The group of ordinary differential equations can be readily solved by a normal ODE solver such as COLSYS [Ascher (1981)].
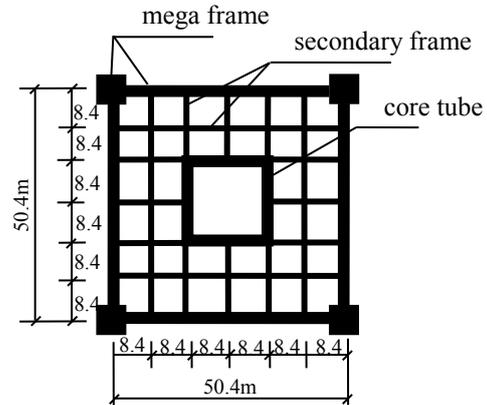
Figure 3. cross-section of a mega frame-core-tube system

4. Example and computational result analysis

The purpose of the section is to demonstrate the numerical determination of resonant response for a super tall building system due to a given complex ground motion.

4.1 *Example*

Figure 3 shows the cross section of a mega frame-core-tube system adopted by a reinforced concrete super tall building under its structural construction as shown in Figure 1 (a). The height of the main superstructure is 261.9 meters, and the height of its foundation structure is 21 meters. The cross-section area of all the mega columns and beams is $2.4 \times 2.4 \text{m}^2$; the cross-section size for all the columns and beams of the secondary frame in the mega frame is $0.7 \times 0.7 \text{m}^2$; the distance between two columns is 8.4m. A box-pile foundation is implemented, and the foundation soil is clayey silt. The equivalent stiffness of soil at the bottom and side faces of the foundation pit is respectively utilized as follows.
$K_{zD} = r_d 32.4 \times 10^5$ KPa/m, $K_{tD} = r_d 25.1 \times 10^5$ KPa/m, $K_{tH} = r_d 25.1 \times 10^5$ KPa/m, $K_{nH} = r_d 32.8 \times 10^5$ KPa/m. r_d is a coefficient that depends on the realistic site conditions (1.0 is used in this example); C_r is the contact coefficient between the foundation and the subgrade (takes 0.5 in this example); the materials used in the structural system are respectively: the thickness of the wall for the foundation tube is

0.6m, the concrete level is $C50$, and the concrete level for the mega frame is $C50$ too; the concrete level executed for the secondary frame is $C35$, the thickness of the wall for the inner tube is 0.40m, 0.45m for the outer tube, and the concrete level for the tubes is $C40$.

4.2 *Computational result analysis*

In the following tables and figures P_x, P_y, P_z and P_θ represent a group of natural periods of the structural system in the X, Y, Z, and θ (around axis Z) directions, respectively. It is implied that if the ground motion periods coincide with the group of natural periods of the structural system, the resonance of the structural system will occur. That is, if $2\pi/\omega_x=P_x$, $2\pi/\omega_y=P_y$, $2\pi/\omega_z=P_z$ and $2\pi/\omega_\theta=P_\theta$, resonance will occur; $\overline{v}_{0x}$, $\overline{v}_{0y}$, θ, $\overline{w}_{in}$, and $\overline{w}_{ex}$ stand for the resonant displacement amplitudes at the top of the structural system in the transvers directions as well as the resonant warping at the top of inner and outer tubes of the computational model.

Table 1. Resonant response of coupling vibration

P_x	P_y	P_z	P_θ	$\overline{v}_{0x}$	$\overline{v}_{0y}$	$\overline{w}_{in}$	$\overline{w}_{ex}$
3.9	4.9	4.2	21	1866.75	52.70	342.66	372.85
4.5	4.9	4.2	1.3	38.66	34.46	13.99	15.07
2.9	3.1	3.1	3.1	1533.69	11.15	319.34	344.75

Table 2. Influence of dynamic-property adjustment on resonant response

P_x	$\overline{v}_{0y}$	$\overline{w}_{in}$	$\overline{w}_{ex}$
7.9	0.45	0.05	0.11
3.5	1.62	0.46	0.28
2.2	33.86	5.05	6.52
1.7	151.96	28.95	28.38
1.5	245.71	45.16	44.71

Table 3. Influence of structural stiffness on resonant response

structural stiffness	$\overline{v}_{0x}$	$\overline{v}_{0y}$	$\overline{\theta}$	$\overline{w}_{in}$	$\overline{w}_{ex}$
3.15	3.51	3.25	-0.82	0.45	2.53
3.5	1.69	-0.16	-0.82	0.26	1.22
4.15	-3.34	-26.24	-0.82	-1.87	-5.25
4.55	9.97	463.79	-0.82	38.67	107.55

Table 4. Influence of foundation stiffness on resonant response

foundation stiffness	$\overline{v}_{0x}$	$\overline{v}_{0y}$	$\overline{\theta}$	$\overline{w}_{in}$	$\overline{w}_{ex}$
2.2	3.58	4.60	-0.82	0.58	2.59
3.2	3.55	4.53	-0.82	0.57	2.56
4.2	3.53	4.49	-0.82	0.57	2.55

5. Conclusions

The structural resonance will occur when the ground motion period in one direction is very close to the natural period of the structural system in the identical direction. The characteristic of the computational result is that the dynamic response value is very large (should be infinite theoretically), as shown in Figure 4, a step change is happening.

A designer must pay attention to the coupling natural periods of a structural system as long as the movement of its foundation soil during an earthquake is very hard to predict or evaluate quantitatively. The adjustment of a structural system, including the change of its material, arrangement of its components, etc., will lead to the change of its dynamic property, especially its coupling natural period that possess many combinations, as listed in table 1. As shown in Figure 5, when the ground motion period changes to a certain degree, the structural system might experience a different coupling resonant state. Also as listed in table 2, they implicitly teach us that the dynamic property improvement of a structural system just in a single direction could render the structural system to stay in a potential coupling resonant state.

The resonant periods or natural periods of a structural system strongly depend on its global rigidity, as shown in table 3. The reduction of global rigidity of a structural system will make its natural period become longer, and vice versa.

The influence of the foundation stiffness of a structural system on its resonant periods or natural periods is not obvious if the superstructure remain unchanged, as shown in table 4. These computational results tell us that it is not a wise way to improve the dynamic property of a structural system by means of increasing the size of the foundation in its aseismic design.

The methodology presented in the paper is helpful for the determination of coupling frequencies or periods of a complex structural system, which are very hard to find in the published literatures hitherto and to determine by utilizing other numerical methods.
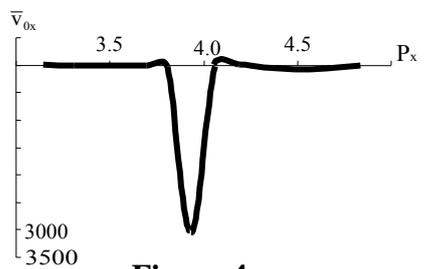
Figure 4.

variation law between $\overline{v}_{0x}$ and P_x
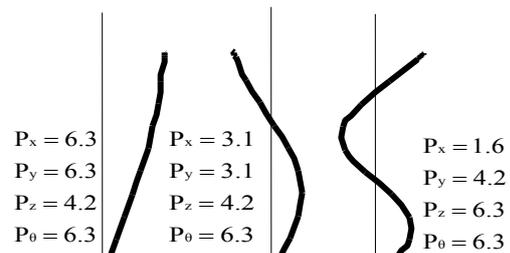
Figure 5.

the first three coupling resonent modes of v_{0x}

Acknowledgments

References

X. Zhao, J. M. Ding and H. H. Sun. (2011) Structural design of shanghai tower for wind loads, *Procedia Engineering* **14**, 1759–1767.

Reza Kamgar, Mohammad Mehdi and Saadatpour. (2011) A simple mathematical model for free vibration analysis of combined system consisting of framed tube, shear core, belt truss and outrigger system with geometrical discontinuities, *Appl. Math. Modell.* doi:10.1016/j.apm.2011.12.029.

Hong Fan, Q.L. Li, Alex Y. Tuan and Lihua Xu. (2009) Seismic analysis of the world's tallest building, *Journal of Constructional Steel Research* **65**, 1206–1215.

Wen-Hae Lee. (2007) Free vibration analysis for tube-in-tube tall buildings, *Journal of Sound and Vibration* **303**, 287–304.

Yuan Si. (1991) ODE conversion techniques and their applications in computational mechanics, *Acta Mechanica Sinica* **7(3)**, 283–288.

Yuan Si and Zhang Yiguo. (1993) Solution of eigenproblems in ODES by standard ODE solvers, *Earthquake Engineering and Engineering Vibration* (in Chinese) **13(2)**, 94–102.

Mindlin Raymond D. (1936) Force at a point in the interior of a semi-infinite solid, *Physics*, **7(5)**, 195–202.

Gong Yaoqing and Zhang Zhengwei. (2007) Formulation of equivalent stiffness for semi-infinite elastic subgrade and its application, *Engineering Mechanics* (in Chinese) **24(5)**, 10–16.

Ascher, U., Christiansen, J., and Russell, R. D. (1981) Collocation Software for Boundary-Value ODEs, *ACM Trans. Math. Software* **7(2)**, 209-222.

Yaoqing Gong. (2010) A New Free Vibration Analysis Method for Space Mega Frames of Super Tall Buildings. *IOP Conf. Series: Materials Science and Engineering* 10 (2010) 012002, doi:10.1088/1757–899X/10/1/012002.

A face-based smoothed XFEM for three-dimensional fracture problems

[*]**Y. Jiang[1], T.E. Tay[2], and Eric Li[3]**

[1]Department of Engineering Mechanics, Institute of High Performance Computing, 1 Fusionopolis Way, Connix, Singapore 138632

[2]Department of Mechanical Engineering, National University of Singapore, 9 Engineering Drive 1, Singapore 117576

[3]School of Aerospace, Mechanical and Mechatronic Engineering, The University of Sydney, Sydney, NSW 2006, Australia

∗ Corresponding author: jiangyo@ihpc.a-star.edu.sg

Abstract

In this work, a face-based smoothed extended finite element method (FS-XFEM) is developed for three-dimensional fracture problems. This method combines the extended finite element method (XFEM) and smoothing technique together. With XFEM, arbitrary crack geometry can be modeled and crack advance can be simulated without remeshing. With face-based smoothing technique, the integration of singular term over the volume around the crack front can be eliminated, thanks to the transformation of volume integration into area integration. Numerical examples are presented to test the accuracy and convergence rate of the FS-XFEM. From the results, it is clear smoothing technique can improve the performance of XFEM for three-dimensional fracture problems

Keywords: three-dimension, face-based smoothed extended finite element method, stress intensity factor.

1. Introduction

The fracture analysis by standard finite element method (FEM) is quite cumbersome and tedious caused by conforming the crack geometry to element boundary. Remeshing, which greatly increases the computation time, is always needed to match the new geometry of the crack surface, when the crack advances. In order to avoid these two disadvantages of FEM, the extended finite element method (XFEM) has been proposed to facilitate the modeling of arbitrary crack geometry and its evolvement [Belytschko et al. (1999); Moes et al. (1999);]. In the XFEM, the displacement field of the standard FEM is enriched by a discontinuous displacement function and the asymptotic displacement field around the crack tip based on a local partition of unity. The most important advantage of XFEM is that it can simulate the crack without conforming the mesh to the crack geometry and crack propagation without remeshing. The method can improve the accuracy by incorporating arbitrary functions into the displacement field of the standard FEM to describe the local behavior around the singular features such as crack tips, notches or corners and thus is a flexible and powerful tool in the field of fracture mechanics. Currently, the XFEM is widely used to simulate two- and three-dimensional elastic and plastic fracture problems [Elguedj et al. (2006); Bordas et al. (2008); Rabczuk et al. (2007); Rabczuk et al. (2009)]. Attracted by the advantages of the XFEM, researchers in other fields of computational physics have also employed it [Chessa et al. (2003); Chopp et al. (2003); Merle et al. (2002); Ji et al. (2002)].

On the other hand, a generalized gradient smoothing technique was introduced by [Chen et al. 2001]. More recently, Liu has established a G space theory and developed weakened weak (W2) formulation which has been the foundation for smoothed finite element methods (SFEM) [Liu et al. (2009); Liu (2010); Liu et al. (2010)]. Using different schemes of smoothing domain formation, cell-based smoothed finite element method (CS-FEM) [Le et al. (2010)], node-based smoothed finite element method (NS-FEM) [Liu et al. (2010)] and edge-based smoothed finite element method (ES-FEM) [Chen et al. (2012)] are developed. With the smoothing technique the domain integration is transformed into boundary integration according to the divergence theory. The shape function derivative is replaced with the shape function multiplied by the component of the outward unit vector along the boundary of the smoothing domain. Thanks to this transformation, the singular term existing in the derivatives of the shape functions for fracture mechanics is eliminated with smoothing technique. Smoothed methods have shown several advantages. For example, NS-FEM can provide upper bound solution [Liu et al. (2010)]. ES-FEM [Chen et al. (2012)] is proved to be more efficient and more accurate. In the ES-FEM, the system stiffness matrix is computed using strains smoothed over the smoothing domains formed based on the edges of the triangles. It is proved that the ES-FEM possesses the following excellent properties: (1) the ES-FEM model possesses a close-to-exact stiffness: it is much softer than the 'overly stiff' FEM and much stiffer than the 'overly soft' NS-FEM model; (2) the results are often found to be superconvergent and ultra-accurate: much more accurate than the linear triangular elements of FEM; (3) the implementation of the method is straightforward and no penalty parameter is used, and the computational efficiency is better than the FEM

using the same set of nodes. These properties of the ES-FEM have been confirmed by many works [[Chen et al. (2012)], Liu et al. (2008); Cui et al. (2011)].

In view of the advantages of XFEM and ES-FEM, an edge-based smoothed XFEM has been developed to combine the advantages of the two methods for two-dimensional fracture problems [Jiang et al. (2013); Chen et al. (2012)]. Although the ES-XFEM has achieved remarkable progress in the simulation of fracture mechanics, the previous works are confined to two-dimensional fracture problems. In this paper, for the first time, the face-based smoothing technique is combined into XFEM to develop three-dimensional face-based smoothed extended finite element method (FS-XFEM).

[Karihaloo et al. (2003); Karihaloo et al. (2001)] from a simplified variational function using a truncated asymptotic crack tip displacement, formulated the hybrid crack element (HCE) for evaluating the SIF but also the coefficients of the higher order terms of the crack tip. But it has not been extended to three-dimensional fracture problem. A direct traction boundary integral equation method (TBIEM) for three-dimensional crack problems is developed in [Xie et al. (2014)]. However, a singular system of equations is always obtained [Aliabadi (1997); Cruse (1988)]. Special methods [Pan (1997)] has to be employed to tackle this problem. The proposed FS-XFEM will not have these problems comparing with the above problems.

This paper is organized as follows: Section 2 provides a brief description of FS-FEM. Section 3 introduces the formation of face-based smoothed XFEM. Section 4 illustrates the computational procedure for three-dimensional stress intensity factor (SIF). Section 5 gives two examples to test the newly developed method and compare the results of FS-XFEM with those of XFEM. The conclusion is made in Section 6.

2. Face-based smoothed FEM (FS-FEM)

2.1. Smoothing domain formation

Due to the excellent features of ES-FEM in two-dimensional problems, the FS-FEM [Nguyen-Thoi et al. (2009a; 2009b)] for three-dimensional problems is developed. In the FS-FEM, linear tetrahedral elements, which are feasible for arbitrarily complicated geometry, are used to mesh the problem domain. Instead of using the edges of the elements in two-dimensional problems, faces of the elements in the FS-FEM are used to create smoothing domains. Therefore, it is named face-based smoothed finite element method. The faces of the elements in three-dimension can be classified into two types: boundary face and interior face. The boundary face lies on the boundary of the domain, while the interior face lies inside the domain. The smoothing domains associated with these two types of faces are formed in different ways. For the interior face, which is shared by two elements, the smoothing domain is formed by connecting the three points of the face to the centroids of the two elements shown in Fig. 1(a). For the boundary face, which belongs to only one element, the smoothing domain is formed by the face and the centroid of the only element. Four points (three from the face and one being the centroid of the element) automatically form a tetrahedral shown in Fig. 1(b).
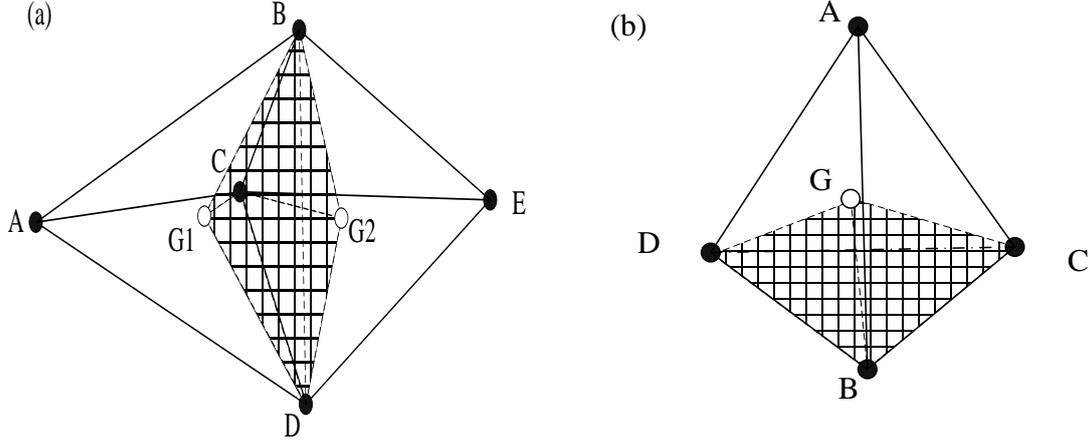
Figure 1. Two types of smoothing domains (a) smoothing domain formed based on interior face (b) smoothing domain formed based on boundary face

2.2. *The formulation of FS-FEM*

In the FS-FEM, the problem domain is meshed with 4-node tetrahedral elements. Based on the above description of smoothing domain formulation, N_{face} smoothing domains in the whole model can be created. Here N_{face} is the number of the faces in the whole problem domain Ω. The smoothing domains satisfy $\Omega = \sum_{k=1}^{N_{face}} \Omega^k$ and $\Omega^i \cap \Omega^j = \varnothing$, $i \neq j$. With face-based smoothing technique, the integration of the derivatives of the shape functions over domain can be transformed into integration of shape functions multiplied with component of outward unit vector of the boundary face. The integration result is then divided by the volume of the smoothing domain. In the setting of FS-FEM, the smoothed strain is obtained as:

$$\bar{\varepsilon} = \overline{\mathbf{B}}\mathbf{u} \tag{1}$$

Here $\mathbf{u} = \begin{bmatrix} u_1 & v_1 & w_1 & \cdots & u_{ns} & v_{ns} & w_{ns} \end{bmatrix}^T$ is the displacement vector with all the displacement components of the nodes belonging to the smoothing domain. $\overline{\mathbf{B}}$ is the strain-displacement relationship matrix in three dimension expressed as:

$$\overline{\mathbf{B}} = \begin{bmatrix} \bar{b}_{ix}(\mathbf{x}_k) & 0 & 0 \\ 0 & \bar{b}_{iy}(\mathbf{x}_k) & 0 \\ 0 & 0 & \bar{b}_{iz}(\mathbf{x}_k) \\ \bar{b}_{iy}(\mathbf{x}_k) & \bar{b}_{ix}(\mathbf{x}_k) & 0 \\ 0 & \bar{b}_{iz}(\mathbf{x}_k) & \bar{b}_{iy}(\mathbf{x}_k) \\ \bar{b}_{iz}(\mathbf{x}_k) & 0 & \bar{b}_{ix}(\mathbf{x}_k) \end{bmatrix} \tag{2}$$

with

$$\bar{b}_{ih}(\mathbf{x}_k) = \frac{1}{V_k^s} \int_{\Gamma^k} n_h^k(\mathbf{x}) N_i(\mathbf{x}) d\Gamma \qquad h = (x, y, z) \tag{3}$$

where Γ^k is the boundary face of the smoothing domain. $n_h^k(\mathbf{x})$ is the h component of the outward normal vector on the boundary Γ^k. $N_i(\mathbf{x})$ is the shape function. V_k^s is the volume of the smoothing domain.

By Gauss quadrature, $\bar{b}_{ih}(\mathbf{x}_k)$ can be further written as:

$$\bar{b}_{ih}(\mathbf{x}_k) = \frac{1}{V_k^s} \sum_{m=1}^{N_{face}^k} \sum_{n=1}^{N_{gau}} n_h^k(\mathbf{x}_n) N_i(\mathbf{x}_n) w_n \quad h = (x, y, z) \tag{4}$$

where N_{gau} is the number of the Gauss points and w_n is the weight of the Gauss point. N_{face}^k is the number of faces attached to the smoothing domain Ω^k. $\mathbf{x}_n$ is the coordinate of the Gauss point on the boundary face. $N_i(\mathbf{x}_n)$ is the ith shape function of the Gauss point $\mathbf{x}_n$. $n_n^k(\mathbf{x}_n)$ is the outward unit normal component.

The set of algebraic equations for FS-FEM can be obtained in the form of matrix:

$$\overline{\mathbf{K}}\overline{\mathbf{d}} = \mathbf{f} \tag{5}$$

Here $\bar{\mathbf{d}}$ is the displacement vector of all the nodes in the model, $\overline{\mathbf{K}}$ is the global stiffness matrix and $\mathbf{f}$ is the nodal force.

The entries in sub-matrices of the stiffness matrix $\overline{\mathbf{K}}$ in Eq. (5) can be expressed as:

$$\overline{\mathbf{K}}_{ij} = \sum_{k=1}^{N_{face}} \overline{\mathbf{K}}_{ij,k}^s \tag{6}$$

Here the summation means an assembly process, $\overline{\mathbf{K}}_{ij,k}^s$ is the stiffness matrix associated with the smoothing domain Ω^k and can be computed by

$$\overline{\mathbf{K}}_{ij,k}^s = \int_{\Omega_k^s} \overline{\mathbf{B}}_i^{\mathrm{T}} \mathbf{D} \overline{\mathbf{B}}_j d\Omega = \overline{\mathbf{B}}_i^{\mathrm{T}} \mathbf{D} \overline{\mathbf{B}}_j V_k^s \tag{7}$$

where V_k^s is the volume of the smoothing domain Ω^k, $\mathbf{D}$ is the matrix of material constants that is defined as follows:

$$\mathbf{D} = \begin{bmatrix} 2G+\lambda & \lambda & \lambda & 0 & 0 & 0 \\ \lambda & 2G+\lambda & \lambda & 0 & 0 & 0 \\ \lambda & \lambda & 2G+\lambda & 0 & 0 & 0 \\ 0 & 0 & 0 & G & 0 & 0 \\ 0 & 0 & 0 & 0 & G & 0 \\ 0 & 0 & 0 & 0 & 0 & G \end{bmatrix} \tag{8}$$

with $G = \dfrac{E}{2(1+\nu)}$, $\lambda = \dfrac{\upsilon E}{(1+\nu)(1-2\nu)}$. Here E is the Young's modulus, and ν is the Poisson's ratio.

3. Face-based smoothed XFEM (FS-XFEM)

3.1. The formulation of FS-XFEM

The displacement of XFEM is composed of three parts: the continuous displacement from standard finite element method, the enrichment part that represents discontinuity across the crack surface and the enrichment part that describes the singular strain field around the crack front. Heaviside function is usually employed as enrichment function for the discontinuity across the crack surface. A set of branch functions, which are derived from the displacement field around the crack front, are used to produce the singular strain field around the crack front. Nodal subtraction is used in FS-XFEM.

$$u(\mathbf{x}) = \sum_{i \in N^{fs-fem}} N_i(\mathbf{x})u_i + \sum_{j \in N^{fs-c}} N_j (H(\mathbf{x}) - H(\mathbf{x}_j))a_j$$
$$+ \sum_{k \in N^{fs-e}} N_k(\mathbf{x}) \sum_{\alpha=1}^{4} (\phi_\alpha(\mathbf{x}) - \phi_\alpha(\mathbf{x}_k))b_k^\alpha \tag{9}$$

Here $N_i(x)$, $N_j(x)$ and $N_k(x)$ are the shape functions associated with different types of nodes and u_i is nodal displacement in standard FEM. N^{fs-fem} is the node set of the whole finite element model. $\mathbf{x}_j$ and $\mathbf{x}_k$ are the coordinates of the jth and kth nodes in the element respectively. $H(\mathbf{x})$ is a Heaviside jump function and is set as follows:

$$H(\mathbf{x}) = \begin{cases} 1 & \text{if } (\mathbf{x} - \mathbf{x}^*) \cdot \mathbf{n} \geq 0 \\ -1 & \text{otherwise} \end{cases} \tag{10}$$

N^{fs-c} is the set of nodes whose support domain is completely cut by the crack surface. a_j is the enriched degree of freedom associated with node set N^{fs-c}. N^{fs-e} is the set of nodes in the vicinity of the crack front. $\phi_\alpha(\mathbf{x})$ are a set of branch functions to model the asymptotic features of the displacement field around the crack front:

$$\phi_\alpha(\mathbf{x})(\alpha = 1 - 4)$$

$$= \left\{ \sqrt{r}\sin(\frac{\theta}{2}) \quad \sqrt{r}\cos(\frac{\theta}{2}) \quad \sqrt{r}\sin(\frac{\theta}{2})\sin(\theta) \quad \sqrt{r}\cos(\frac{\theta}{2})\sin(\theta) \right\} \tag{11}$$

where (r, θ) is the local polar coordinate system, which is defined so that the plane where $\theta = 0$ must be tangent to the crack front. b_k^α is the αth (of the totally four) enriched degree of freedom associated with node set N^{fs-e}.

Caused by different types of enrichment functions, nodes in FS-XFEM can be categorized into three types. (a) common nodes denoted by N^{fs-fem}, which are not enriched by either $H(\mathbf{x})$ or $\phi_\alpha(\mathbf{x})$; (b) $\phi_\alpha(\mathbf{x})$ enriched nodes denoted by N^{fs-e}. As shown in Fig. 2(a), the smoothing domain $ABCG_1G_2$. Here ABC is the face, based on which the smoothing domain is formed. G_1 and G_2 are the centroids of the two elements, which share the face ABC. The crack surface $EFMN$ is in this smoothing domain. But the crack front MN is also inside the smoothing domain, which means that the crack surface does not completely cut the smoothing domain but part of the smoothing domain. $\phi_\alpha(\mathbf{x})$ is used to describe the displacement behavior around the crack front. Therefore, nodes associated with this smoothing domain are enriched by $\phi_\alpha(\mathbf{x})$. (c) $H(\mathbf{x})$ enriched nodes denoted by N^{fs-c}. As shown in Fig. 2(b), the smoothing domain $ABCG_1G_2$ is constructed in the same way as Fig. 2(a). Here EFD is the crack surface. This smoothing domain is completely cut by the crack surface. Therefore, $H(\mathbf{x})$ is used to enrich the nodes associated with the smoothing domain, if the nodes are not enriched by $\phi_\alpha(\mathbf{x})$.
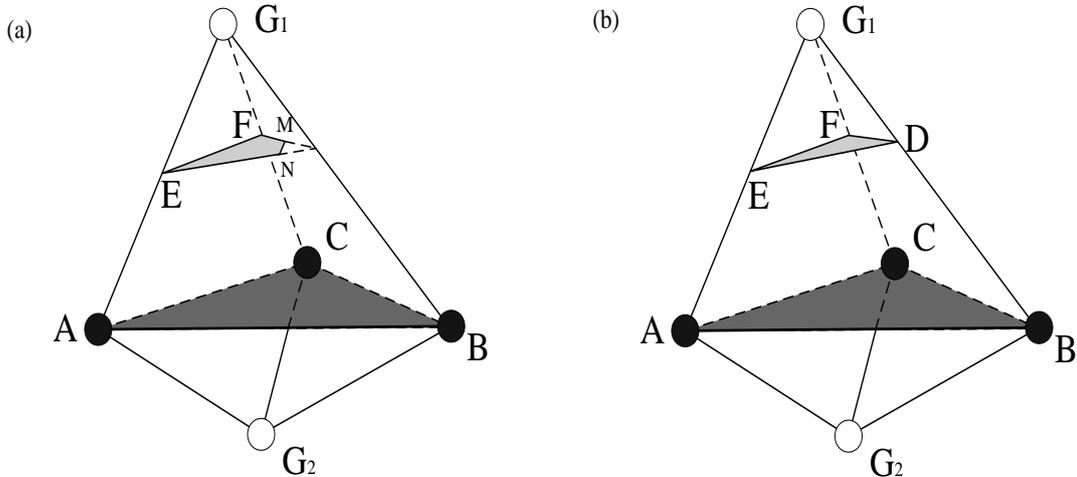
Figure 2. (a) crack-front element (b) crack-cut element

Employing the strain smoothing operation, the smoothed strain over Ω^k from the displacement approximation can be written as:

$$\bar{\varepsilon}_k(\mathbf{x}) = \sum_{i \in N^{fs-fem}} \overline{\mathbf{B}}_i^u(\mathbf{x}_k)\mathbf{u}_i + \sum_{j \in N^{fs-c}} \overline{\mathbf{B}}_j^a(\mathbf{x}_k)\mathbf{a}_j + \sum_{m \in N^{fs-e}} \overline{\mathbf{B}}_m^b(\mathbf{x}_k)\sum_{\alpha=1}^{4} b_k^\alpha \tag{12}$$

where $\overline{\mathbf{B}}_i^u(\mathbf{x}_k)$ is the smoothed strain gradient matrix for the standard FS-FEM part, and $\overline{\mathbf{B}}_j^a(\mathbf{x}_k)$, $\overline{\mathbf{B}}_m^b(\mathbf{x}_k)$ correspond to the Heaviside function and branch functions enriched parts respectively. Those matrices can be written as:

$$\overline{\mathbf{B}}_i^r(\mathbf{x}_k) = \begin{bmatrix} \bar{b}_{ix}^r(\mathbf{x}_k) & 0 & 0 \\ 0 & \bar{b}_{iy}^r(\mathbf{x}_k) & 0 \\ 0 & 0 & \bar{b}_{iz}^r(\mathbf{x}_k) \\ \bar{b}_{iy}^r(\mathbf{x}_k) & \bar{b}_{ix}^r(\mathbf{x}_k) & 0 \\ 0 & \bar{b}_{iz}^r(\mathbf{x}_k) & \bar{b}_{iy}^r(\mathbf{x}_k) \\ \bar{b}_{iz}^r(\mathbf{x}_k) & 0 & \bar{b}_{ix}^r(\mathbf{x}_k) \end{bmatrix} \tag{13}$$

In the above equation, $\bar{b}_{ih}^r(\mathbf{x}_k), h = x, y, z$ and $r = u, a, b$ is computed by:

$$\bar{b}_{ih}^u(\mathbf{x}_k) = \frac{1}{V_k^s}\int_{\Gamma^k} n_h(\mathbf{x})N_i(\mathbf{x})d\Gamma$$

$$\bar{b}_{ih}^a(\mathbf{x}_k) = \frac{1}{V_k^s}\int_{\Gamma^k} n_h(\mathbf{x})N_i(\mathbf{x})(H(\mathbf{x}) - H(\mathbf{x}_i))d\Gamma \qquad (h = x, y, z) \tag{14}$$

$$\bar{b}_{ih}^b(\mathbf{x}_k) = \frac{1}{V_k^s}\int_{\Gamma^k} n_h(\mathbf{x})N_i(\mathbf{x})(\phi_\alpha(\mathbf{x}) - \phi_\alpha(\mathbf{x}_i))d\Gamma$$

Using Gauss quadrature along the segments of boundary, the above equations can be written as:

$$\bar{b}_{ih}^u(\mathbf{x}_k) = \frac{1}{V_k^s}\sum_{m=1}^{N_{face}^k}\sum_{n=1}^{N_{gau}} n_h(\mathbf{x}_n)N_i(\mathbf{x}_n)w_n$$

$$\bar{b}_{ih}^a(\mathbf{x}_k) = \frac{1}{V_k^s}\sum_{m=1}^{N_{face}^k}\sum_{n=1}^{N_{gau}} n_h(\mathbf{x}_n)N_i(\mathbf{x}_n)(H(\mathbf{x}_n) - H(\mathbf{x}_i))w_n \qquad (h = x, y, z) \tag{15}$$

$$\bar{b}_{ih}^b(\mathbf{x}_k) = \frac{1}{V_k^s}\sum_{m=1}^{N_{face}^k}\sum_{n=1}^{N_{gau}} n_h(\mathbf{x}_{m,n})N_i(\mathbf{x}_n)(\phi_\alpha(\mathbf{x}_n) - \phi_\alpha(\mathbf{x}_i))w_n$$

Here N_{face}^k is the number of the boundary faces of the smoothing domain, N_{gau} is the number of the Gauss points used on the boundary face. $\mathbf{x}_n$ is the coordinate of the nth Gauss point on the boundary face.

The stiffness matrix $\overline{\mathbf{K}}$ is yielded by:

$$
\begin{aligned}
\overline{\mathbf{K}}_{ij} &= \sum_{k=1}^{N_s} \overline{\mathbf{K}}_{ij,k}^s \\
&= \sum_{k=1}^{N_s}
\begin{bmatrix}
(\overline{\mathbf{B}}_i^u)^T \mathbf{D}\overline{\mathbf{B}}_j^u V_k^s & (\overline{\mathbf{B}}_i^u)^T \mathbf{D}\overline{\mathbf{B}}_j^a V_k^s & (\overline{\mathbf{B}}_i^u)^T \mathbf{D}\overline{\mathbf{B}}_j^b V_k^s \\
(\overline{\mathbf{B}}_i^a)^T \mathbf{D}\overline{\mathbf{B}}_j^u V_k^s & (\overline{\mathbf{B}}_i^a)^T \mathbf{D}\overline{\mathbf{B}}_j^a V_k^s & (\overline{\mathbf{B}}_i^a)^T \mathbf{D}\overline{\mathbf{B}}_j^b V_k^s \\
(\overline{\mathbf{B}}_i^b)^T \mathbf{D}\overline{\mathbf{B}}_j^u V_k^s & (\overline{\mathbf{B}}_i^b)^T \mathbf{D}\overline{\mathbf{B}}_j^a V_k^s & (\overline{\mathbf{B}}_i^b)^T \mathbf{D}\overline{\mathbf{B}}_j^b V_k^s
\end{bmatrix}
\end{aligned}
\tag{16}
$$

Substituting Eq. (16) into Eq. (5) can produce a set of linear equations. In FS-XFEM, $\mathbf{f}$ is composed of three parts: $\mathbf{f}^u, \mathbf{f}^a$ and $\mathbf{f}^b$. These three vectors can be obtained as follows:

$$
\begin{aligned}
\mathbf{f}_{i,m}^u &= \int_{\Omega_m^e} N_i(\mathbf{x})\mathbf{b}d\Omega + \int_{\Gamma_t} N_i(\mathbf{x})\mathbf{t}_\Gamma d\Gamma \\
\mathbf{f}_{i,m}^a &= \int_{\Omega_m^e} N_i(\mathbf{x})H(\mathbf{x})\mathbf{b}d\Omega + \int_{\Gamma_t} N_i(\mathbf{x})H(\mathbf{x})\mathbf{t}_\Gamma d\Gamma \\
\mathbf{f}_{i,m}^{b\alpha} &= \int_{\Omega_m^e} N_i(\mathbf{x})\phi_\alpha(\mathbf{x})\mathbf{b}d\Omega + \int_{\Gamma_t} N_i(\mathbf{x})\phi_\alpha(\mathbf{x})\mathbf{t}_\Gamma d\Gamma
\end{aligned}
\tag{17}
$$

4. Three-dimensional stress intensity factor calculation

Several numerical techniques, in conjunction with finite-element (FE) analyses, have been developed to calculate fracture mechanics parameters. Three of these techniques are: (1) the virtual crack extension (VCE) method [Parks et al. (1974; 1977); Hellen (1975; 1989)], (2) the virtual crack closure technique [Rybicki et al. (1977); Shivakumar et al.(1988); Raju et al. (1988); Buchholz (1984)], and (3) the *J*-integral method [Rice (1968); Cherepanov (1967; 1969); Eshelby (1956);]. Based on *J*-integral method, an interaction energy integral method is used to calculate stress intensity factor in this work. A cylindrical volume with the radius r_d surrounding a point C located on the crack front is shown in Fig. 3. If the crack surfaces are traction-free, the domain form of the interaction energy integral $I(s)$ can be written as:

$$
I(s) = \frac{-\int_V [\mathrm{tr}(\mathbf{P} \cdot \vec{\nabla}\mathbf{q}) + (\vec{\nabla} \cdot \mathbf{P}^\mathrm{T}) \cdot \mathbf{q}]dV}{\int_{L_c} \Delta a(s)ds}
\tag{18}
$$

where

$$
\mathbf{P} = \boldsymbol{\sigma} : \boldsymbol{\varepsilon}^{\mathrm{aux}}\mathbf{I} - \vec{\nabla}\mathbf{u}^{\mathrm{aux}} \cdot \boldsymbol{\sigma} - \vec{\nabla}\mathbf{u} \cdot \boldsymbol{\sigma}^{\mathrm{aux}}
\tag{19}
$$

Here the superscript aux stands for auxiliary field. $\boldsymbol{\varepsilon}$, $\mathbf{u}$ and $\boldsymbol{\sigma}$ are the actual strain, displacement and stress field respectively.
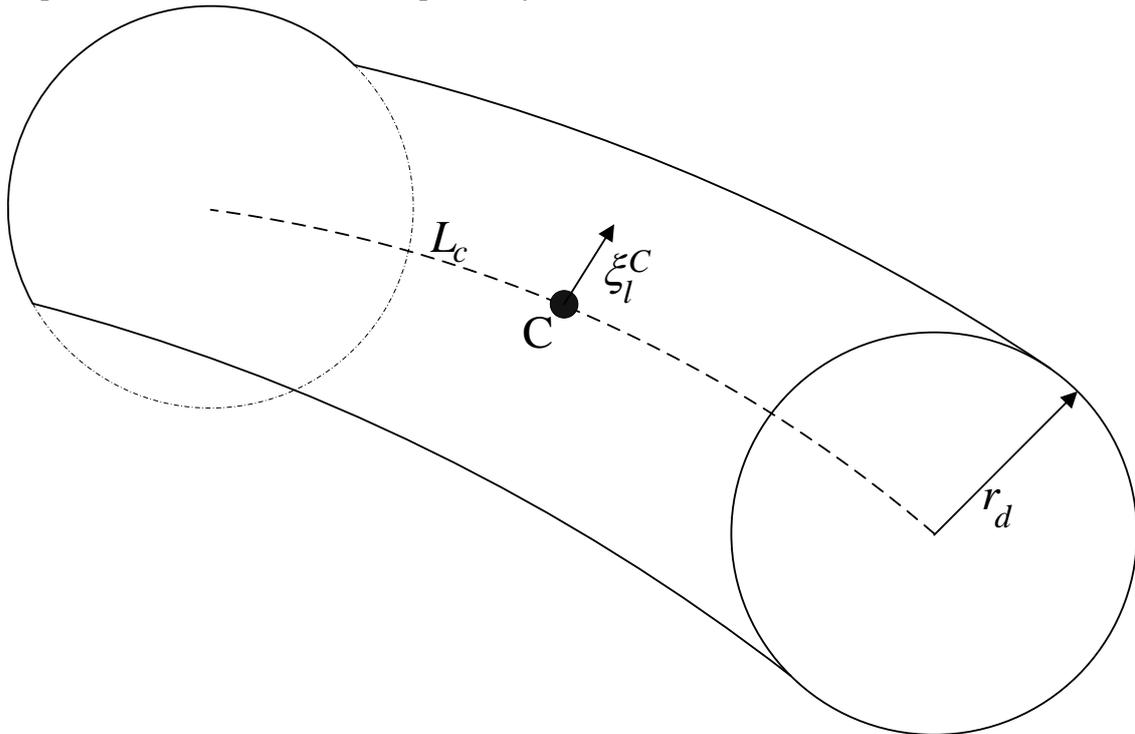
Figure. 3. A cylindrical volume surrounding a point C

The main difficulty in the calculation of interaction energy integral lies in the evaluation of the gradients and higher order gradients of the auxiliary fields that appear in the integrand. As shown in Fig. 4, we define a local orthogonal coordinate system at a point s along the crack front such that the local x_2 axis is perpendicular to the plane of the crack, and the x_1 and x_3 axes lie in the plane of the crack and are normal and tangent respectively to the crack front. To illustrate a convenient procedure to evaluate these gradients, we consider a point p which lies in the local $x_1 - x_2$ plane as shown in Fig. 4. The base vectors $\mathbf{e}_1$, $\mathbf{e}_2$ and $\mathbf{e}_3$ as shown in the figure are constructed by keeping $\mathbf{e}_1$ and $\mathbf{e}_2$ parallel to x_1 and x_2 and moving in the direction of x_3. r, θ are local polar coordinates defined in the figure. The auxiliary fields expressed in this orthogonal curvilinear coordinate system are given as:

$$\sigma_{11}^{aux} = \frac{1}{\sqrt{2\pi r}}\left\{ K_{I}^{aux}\cos\frac{\theta}{2}\left[1-\sin\frac{\theta}{2}\sin\frac{3\theta}{2}\right] - K_{II}^{aux}\sin(\frac{\theta}{2})\left[2+\cos\frac{\theta}{2}\cos\frac{3\theta}{2}\right]\right\}$$

$$\sigma_{22}^{aux} = \frac{1}{\sqrt{2\pi r}}\left\{ K_{I}^{aux}\cos\frac{\theta}{2}\left[1+\sin\frac{\theta}{2}\sin\frac{3\theta}{2}\right] + K_{II}^{aux}\sin\frac{\theta}{2}\cos\frac{\theta}{2}\cos\frac{3\theta}{2}\right\}$$

$$\sigma_{12}^{aux} = \frac{1}{\sqrt{2\pi r}}\left\{ K_{I}^{aux}\sin\frac{\theta}{2}\cos\frac{\theta}{2}\cos\frac{3\theta}{2} + K_{II}^{aux}\cos\frac{\theta}{2}\left[1-\sin\frac{\theta}{2}\sin\frac{3\theta}{2}\right]\right\} \qquad (20)$$

$$\sigma_{23}^{aux} = \frac{K_{III}^{aux}}{\sqrt{2\pi r}}\cos\frac{\theta}{2}$$

$$\sigma_{13}^{aux} = -\frac{K_{III}^{aux}}{\sqrt{2\pi r}}\sin\frac{\theta}{2}$$

$$\sigma_{33}^{aux} = \upsilon(\sigma_{11}^{aux} + \sigma_{22}^{aux})$$

$$u_{1}^{aux} = \frac{1}{8G}\sqrt{\frac{2r}{\pi}}\left\{ K_{I}^{aux}\left[(5-8\upsilon)\cos\frac{\theta}{2}-\cos\frac{3\theta}{2}\right] + K_{II}^{aux}\left[(9-8\upsilon)\sin\frac{\theta}{2}+\sin\frac{3\theta}{2}\right]\right\}$$

$$u_{2}^{aux} = \frac{1}{8G}\sqrt{\frac{2r}{\pi}}\left\{ K_{I}^{aux}\left[(7-8\upsilon)\sin\frac{\theta}{2}-\sin\frac{3\theta}{2}\right] - K_{II}^{aux}\left[(3-8\upsilon)\cos\frac{\theta}{2}+\cos\frac{3\theta}{2}\right]\right\} \qquad (21)$$

$$u_{3}^{aux} = \frac{1}{G}\sqrt{\frac{2r}{\pi}}K_{III}^{aux}\sin\frac{\theta}{2}$$

$$\varepsilon_{11}^{aux} = \frac{\partial u_{1}^{aux}}{\partial \xi_{1}}; \varepsilon_{22}^{aux} = \frac{\partial u_{2}^{aux}}{\partial \xi_{2}}; \varepsilon_{12}^{aux} = \frac{1}{2}(\frac{\partial u_{2}^{aux}}{\partial \xi_{1}} + \frac{\partial u_{1}^{aux}}{\partial \xi_{1}})$$

$$\varepsilon_{13}^{aux} = \frac{1}{2}(\frac{\partial u_{1}^{aux}}{\partial \xi_{3}} + \frac{\partial u_{3}^{aux}}{\partial \xi_{1}}); \varepsilon_{23}^{aux} = \frac{1}{2}(\frac{\partial u_{3}^{aux}}{\partial \xi_{2}} + \frac{\partial u_{2}^{aux}}{\partial \xi_{3}}); \varepsilon_{33}^{aux} = 0 \qquad (22)$$

29
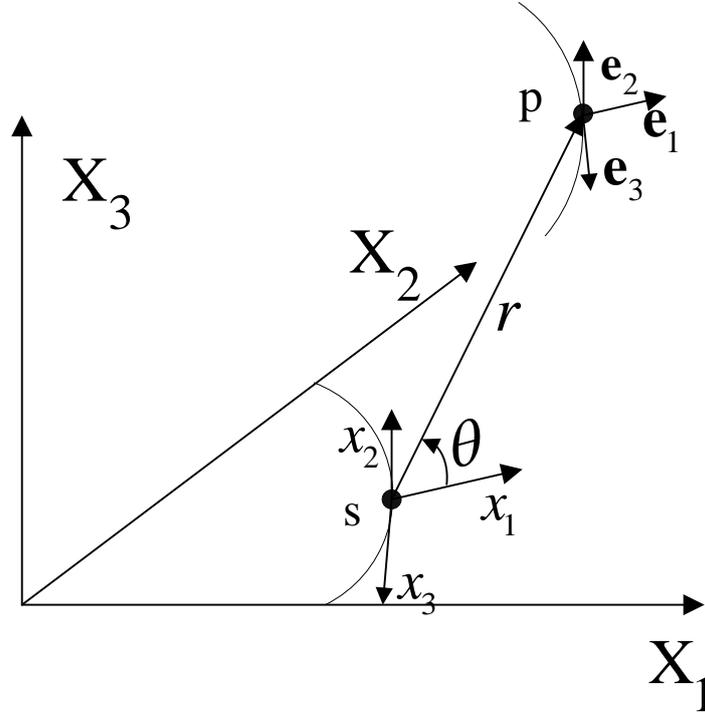
Figure 4. Local orthogonal coordinate system at a point s along the crack front

The weighting function q is defined as follows: A set of elements having at least one node within a cylindrical volume of radius r_d around the crack front are selected. The value of q_c of node associated with the crack tip C is defined:

$$q_c = \begin{cases} \xi_l^C & \text{if } x_3^C = 0 \text{ and } r^C < r_d \\ 0 & \text{otherwise} \end{cases} \tag{23}$$

where ξ_l^C is a unit vector that is perpendicular to the crack front at the crack tip C and lies in the local tangent plane to the crack surface.

Having defined the auxiliary fields, the interaction energy integral $I(s)$ defined by Eq. (18) takes the value

$$I(s) = \frac{2(1-\upsilon^2)}{E}\left[K_{\mathrm{I}}K_{\mathrm{I}}^{\mathrm{aux}} + K_{\mathrm{II}}K_{\mathrm{II}}^{\mathrm{aux}}\right] + \frac{1}{G}K_{\mathrm{III}}K_{\mathrm{III}}^{\mathrm{aux}} \tag{24}$$

Here $K_{\mathrm{I}}^{\mathrm{aux}}$, $K_{\mathrm{II}}^{\mathrm{aux}}$, and $K_{\mathrm{III}}^{\mathrm{aux}}$ are the stress intensity factors associated with the auxiliary fields and K_{I}, K_{II}, and K_{III} are the stress intensity factors associated with the actual fields. The process of evaluating the actual stress intensity factors involves making a judicious choice of the auxiliary stress intensity factors, and then evaluating the interaction energy integral. For example, Substituting $K_{\mathrm{I}}^{\mathrm{aux}} = 1$, and $K_{\mathrm{II}}^{\mathrm{aux}} = K_{\mathrm{III}}^{\mathrm{aux}} = 0$ into Eq. (24) yields:

$$K_{\mathrm{I}}(s) = \frac{E}{2(1-\upsilon^2)} I(s) \tag{25}$$

Similarly, K_{II} can be obtained by substituting $K_{\mathrm{II}}^{\mathrm{aux}} = 1$ and $K_{\mathrm{I}}^{\mathrm{aux}} = K_{\mathrm{III}}^{\mathrm{aux}} = 0$ and K_{III} by substituting $K_{\mathrm{III}}^{\mathrm{aux}} = 1$ and $K_{\mathrm{I}}^{\mathrm{aux}} = K_{\mathrm{II}}^{\mathrm{aux}} = 0$.

5. Numerical examples

Two examples are presented in this work to test our method. One is a plate with a thorough edge crack under tension. The other problem is a cylinder with a penny-shaped crack under remote tension. Strain energy and SIFs are obtained by FS-XFEM and compared with those of XFEM.

5.1. A plate with a thorough edge crack under tension

A plate with a thorough edge crack under tension is first analyzed as shown in Fig. 5. The mesh is plotted in Fig. 6. The dimension of plate is: the height $H = 2\,\mathrm{mm}$ the width $W = 1\,\mathrm{mm}$ and the thickness $t = 0.5\,\mathrm{mm}$ with the crack length $a = 0.3\,\mathrm{mm}$. The load $\sigma = 1\,\mathrm{MPa}$ is applied on the top surface of the plate. All the degrees of freedom on the bottom surface are fixed. The material parameters are: Young's modulus $E = 1\,\mathrm{MPa}$ and the Poisson's ratio $\upsilon = 0.3$.
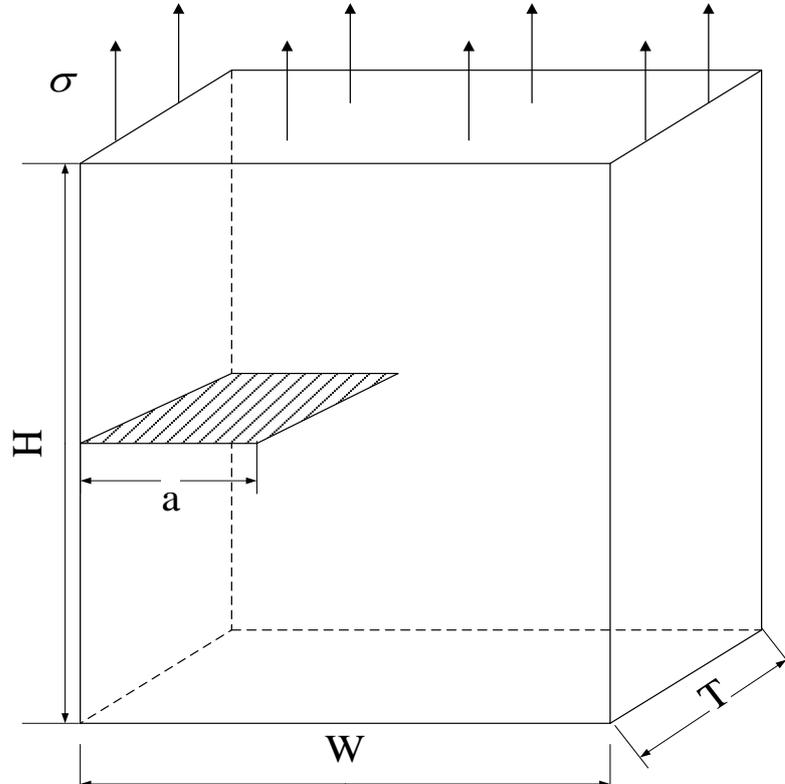
Figure 5. A plate with a thorough edge crack under tension
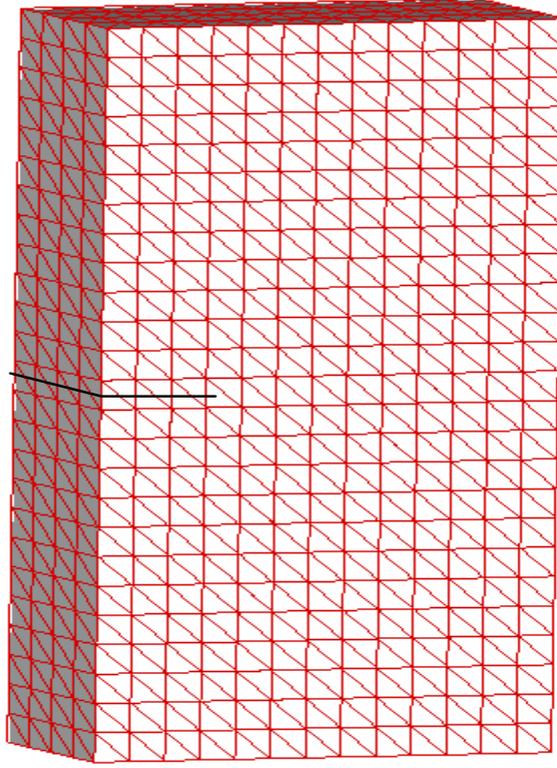
Figure 6. Mesh of the plate with a thorough edge crack

5.1.1 Result

Five types of meshes with linear tetrahedral elements ($13 \times 25 \times 4$, $18 \times 35 \times 4$, $19 \times 37 \times 4$, $31 \times 61 \times 4$, $41 \times 81 \times 4$) are used in the model. A sample mesh is shown in Fig. 6. For comparison, the results are also computed using XFEM. The reference solution of strain energy is obtained using singular FEM with very fine mesh (2,179,458 nodes) in this study. The strain energy is defined as:

$$U = \frac{1}{2} \int_{\Omega} \boldsymbol{\varepsilon}^{T} \mathbf{D} \boldsymbol{\varepsilon} d\Omega \qquad (26)$$

The results of the strain energy produced by FS-XFEM and XFEM are plotted in Fig. 7. From the figure, it can be seen that the numerical results obtained from FS-XFEM are closer to the reference solution than those of XFEM using the same mesh. This is due to integration of face-based smoothing technique into the XFEM.
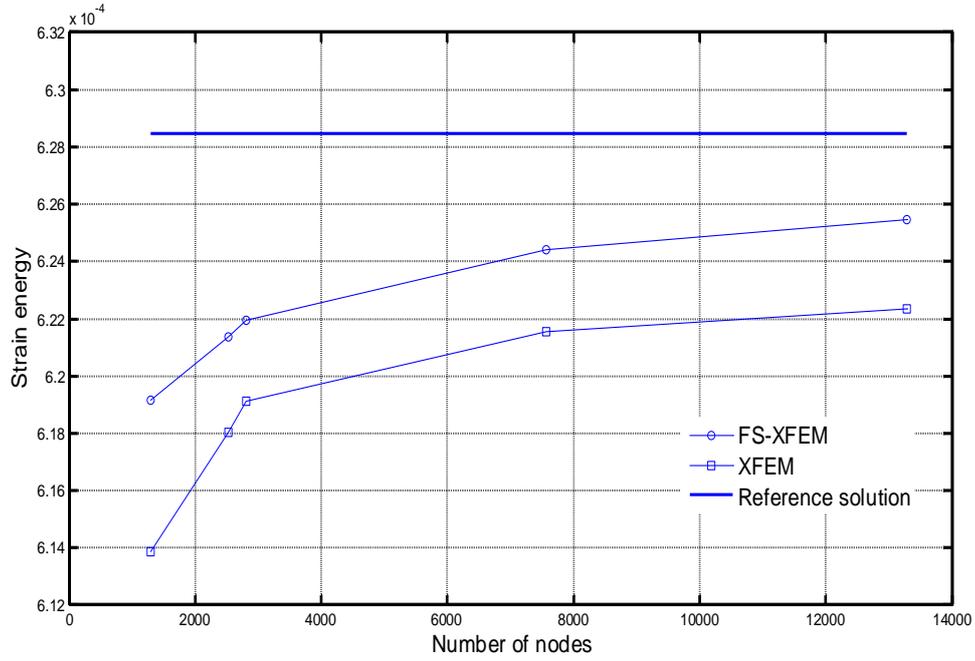
Figure 7. Plate with a thorough edge crack under tension: the variation of strain energy given by XFEM and FS-XFEM with different node numbers

5.1.2 Convergence rate of FS-XFEM

The convergence property of FS-XFEM and XFEM is studied in this section. In order to investigate quantitatively the numerical results, an error indicator in energy norm is defined as follows:

$$E_e = \sqrt{\frac{\left|U_{\text{num}} - U_{\text{ref}}\right|}{U_{\text{ref}}}} \tag{27}$$

where U_{ref} denotes the strain energy of reference solution and U_{num} stands for the strain energy of numerical solution. The errors in strain energy norm against h for this example is plotted in Fig. 8, where h is the average distance between two adjacent nodes. From the figure, it can be seen that the error of FS-XFEM is smaller than that of XFEM with the same mesh. At the same time, FS-XFEM has higher convergence rate than XFEM for this example, which means that FS-XFEM can converge to the reference solution at a higher rate.
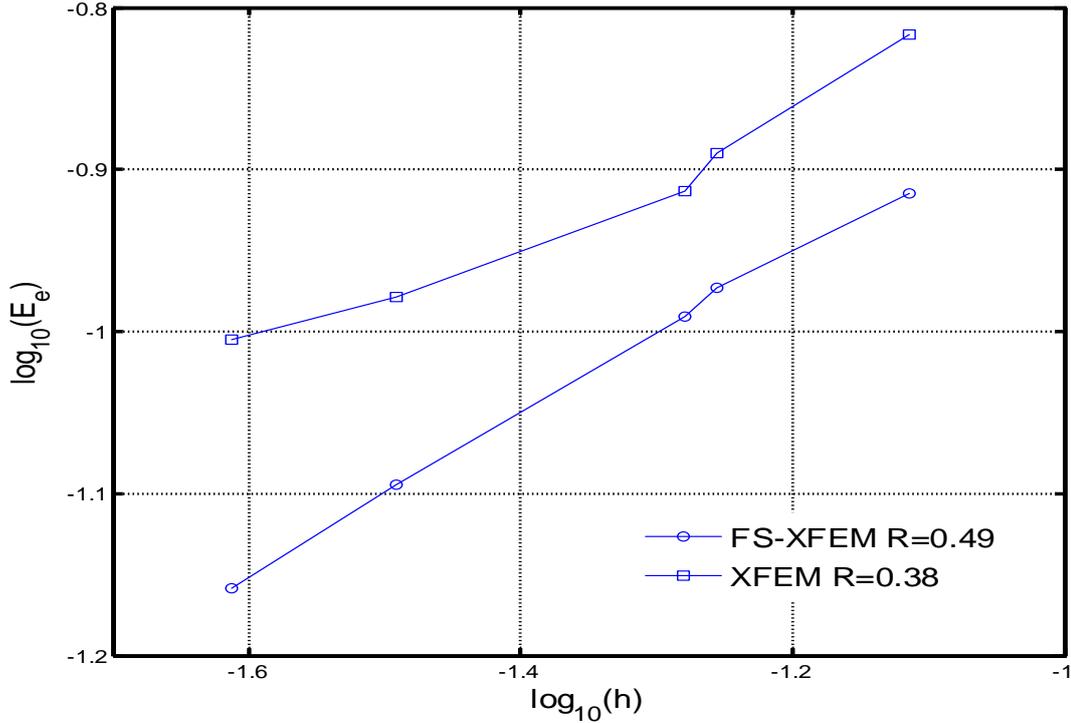
Fig. 8. Convergence rate of XFEM and ES-XFEM for a plate with a thorough crack under tension

5.1.3 Condition number

Another important property of numerical methods is the condition number of the global stiffness matrix, $cond(K)$. The condition number of the global stiffness matrix can affect the number of iterations needed to obtain a converged solution in the manner of $n_{iter} \propto \sqrt{cond(K)}$, when an iteration solver is used to solve the algebraic system equation. The condition number of FS-XFEM and XFEM for thist example against node numbers is listed in Table 1. As it can be seen, FS-XFEM has bigger condition number than XFEM with the same mesh. But the difference is not quite big.

Table 1. Condition numbers of FS-XFEM and XFEM for the first example with different mesh densities

Mesh	13×25×4	18×35×4	31×61×4	41×81×4
XFEM	1.3212e+006	1.8420e+006	1.4312e+007	1.7578e+007
FS-XFEM	1.5427e+006	2.0628e+006	1.7675e+007	2.8165e+007

5.1.4 Efficiency of FS-XFEM

In the assessment of numerical methods, the time cost of different numerical methods should also be taken into consideration. As shown in Table 2, the time consumption for FS-XFEM and XFEM with different meshes is compared. From the table, it is clear that the FS-XFEM takes more time to solve the equation than XFEM for the same mesh. This is in agreement with the condition number comparison between FS-XFEM and XFEM. However, after taking the results accuracy into account and

considering the efficiency, the present FS-XFEM is found to perform much better than XFEM for the results in energy error norms as shown in Fig. 9. From the figure, it is clear that within the same computation time, the results of FS-XFEM are more accurate than XFEM.

Table 2. Time cost of FS-XFEM and XFEM for the first example with different mesh densities

Mesh	13×25×4	18×35×4	31×61×4	41×81×4
XFEM	0.321761s	1.185841s	12.08005s	29.681937s
FS-XFEM	0.67652s	2.579944s	34.70061s	104.7701s
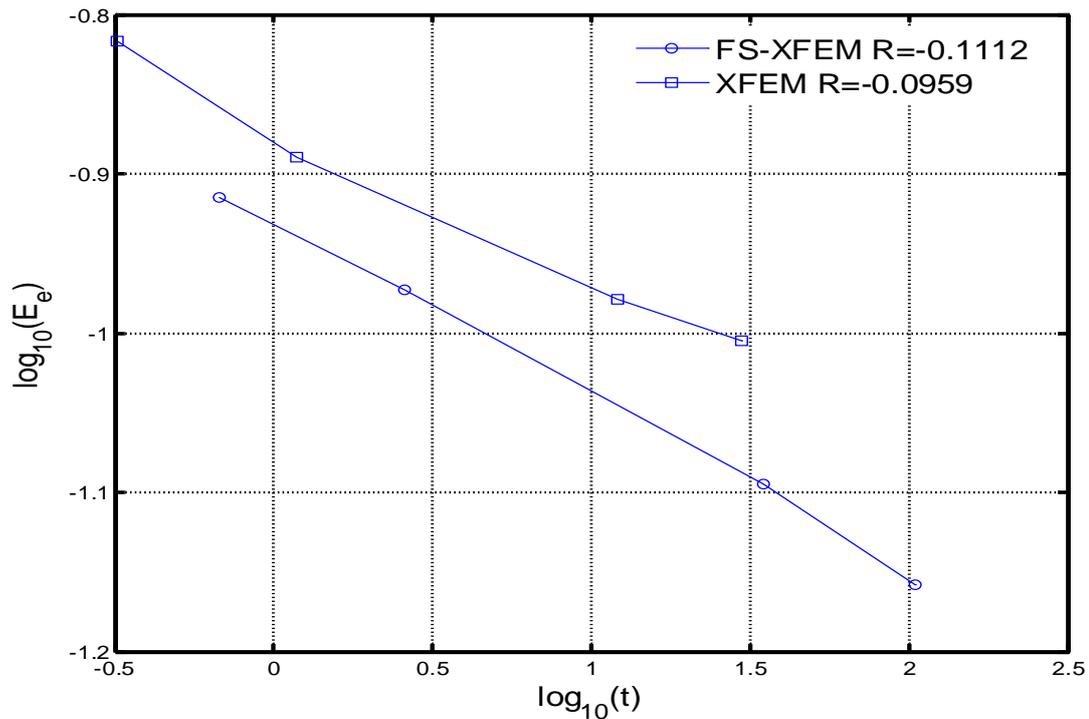
Figure 9. Comparison of computational efficiency of FS-XFEM and XFEM in terms of energy norm for a plate with a thorough crack under tension

5.2. A cylinder with a penny-shaped crack under remote tension

From the first example, it is seen that the FS-XFEM is powerful to simulate a straight crack in three-dimension. In order to extend the applicability of the proposed method, a cylinder with a penny-shaped crack under remote tension is studied. The crack is in the middle of the cylinder, with the radius (of the penny) $a = 0.3$ mm shown in Fig. 10. The remote tension is applied on the top surface of the cylinder. The bottom surface of the cylinder is fixed. The geometrical details are as follows: $H = 12$ mm and $R = 3$ mm. With the ratio $a/R = 0.1$, this problem can be considered as a crack in an infinite body. The solution of stress intensity factor is given by [Anderson

(1995)]: $K_I = \dfrac{2}{\pi}\sigma\sqrt{\pi a}$. This is an axisymmetric problem. Due to the symmetry, only one quarter of the model is simulated with appropriate boundary condition shown in Fig. 11. In this model, symmetrical boundary condition are prescribed on both of the two side surfaces of the quarter-cylinder.
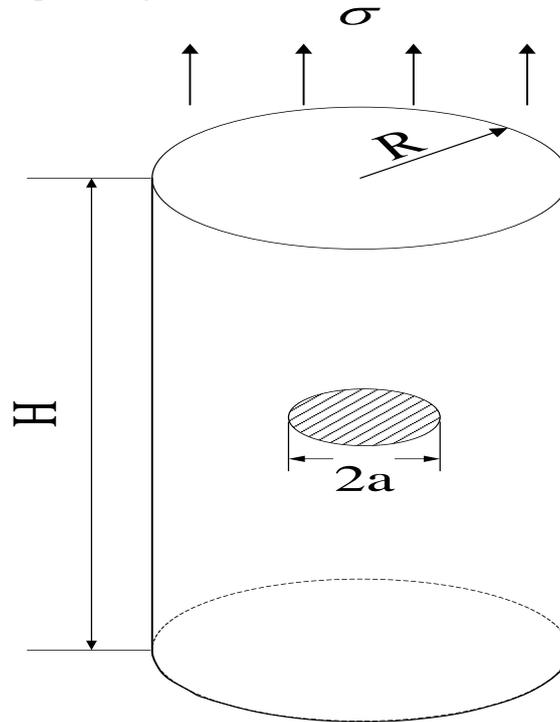
Figure 10. A cylinder with a penny-shaped crack under remote tension
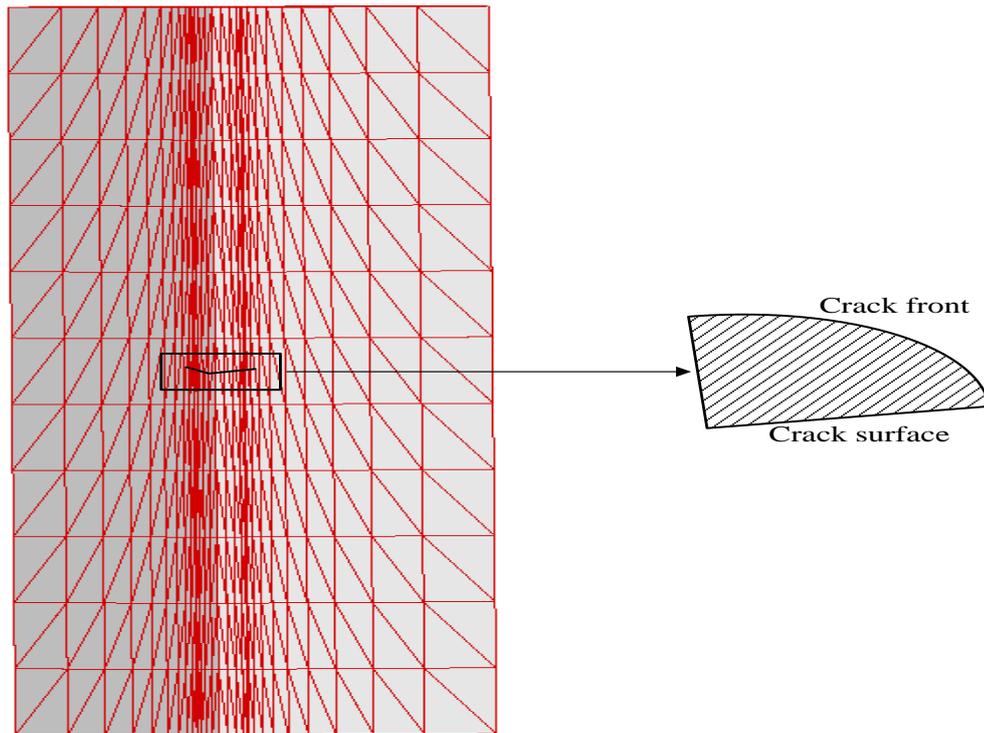
Figure 11. The mesh used for a cylinder with a penny-shaped crack under remote tension

The strain energy for this problem by FS-XFEM and XFEM is plotted in Fig. 12. The reference solution for strain energy is obtained by singular FEM with very fine mesh (1,443,082 nodes). It can be seen that FS-XFEM can produce more accurate results than XFEM with the same mesh. The stress intensity factor (SIF) is also obtained by FS-XFEM and XFEM. The SIFs with error are tabulated in Table 3. From the table, it is noticed that numerical solutions of SIFs using FS-XFEM are closer to the reference solutions than XFEM for the same mesh. This confirms that face-based smoothing technique has a strong value to integrate to XFEM.
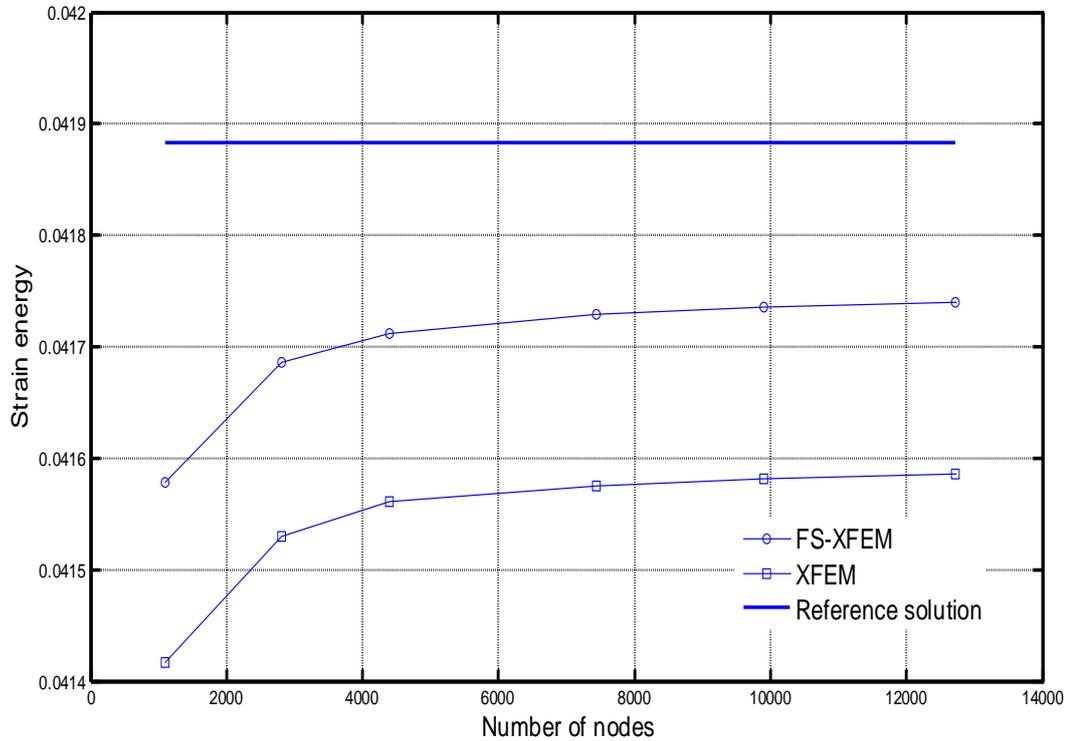
Figure 12.   The variation of strain energy with number of nodes for a cylinder with a penny-shaped crack under remote tension

Table 3. $K_{\mathrm{I}}(\mathrm{MPa}\sqrt{\mathrm{mm}})$ (with error) of FS-XFEM and XFEM for a cylinder with a penny-shaped crack under remote tension with different mesh densities

Mesh	1352	2500	4968	6016	9306
XFEM	0.6097	0.6114	0.6118	0.612	0.6121
	(1.34%)	(1.07%)	(1%)	(0.97%)	(0.95%)
FS-XFEM	0.6121	0.6137	0.6140	0.6143	0.6144
	(0.95%)	(0.70%)	(0.65%)	(0.6%)	(0.58%)

6. Conclusion

In this work, the face-based smoothing technique is combined into extended finite element method (XFEM) to develop face-based smoothed extended finite element method (FS-XFEM) for three-dimensional fracture problems. Two numerical examples are used to test the accuracy, efficiency and convergence rate of FS-XFEM. Through the numerical results some conclusions can be drawn as follows:

1. There are no additional parameters involved in the FS-XFEM, hence, the implementation of FS-XFEM using tetrahedral element that can be generated by many commercial software is quite straightforward.

2. Due to the properly softening effects provided by the face-based smoothing technique, the proposed FS-XFEM possesses a close-to-exact stiffness of the continuous system. Hence, it can provide more accurate results than XFEM using the same tetrahedron mesh in terms of strain energy and stress intensity factors.

3. The convergence rate and computational efficiency of FS-XFEM have been improved significantly compared with XFEM. FS-XFEM also possesses some advantages compared to XFEM. For example, in the calculation of the stiffness matrix, no singular term appears in the integrand. Mapping, which increases the complexity of the calculation, is not needed.

Reference

Belytschko T. and Black T. (1999) Elastic crack growth in finite elements with minimal remeshing. Int J Numer Methods Eng 45, 601-620.

Moes N., Dolbow J. and Belytschko T. (1999) A finite element method for crack growth without remeshing. Int. J. Numer. Methods Eng 46, 131-150.

Elguedj T., Gravouil A. and Combescure A. (2006) Appropriate extended functions for XFEM simulation of plastic fracture mechanics. Comput. Methods Appl. Mech. Engrg 195, 501-515.

Bordas S., Rabczuk T. and Zi G. (2008) Three-dimensional crack initiation, propagation, branching and junction in non-linear materials by an extended meshfree method without asymptotic enrichment. Eng. Fract. Mech 75, 943-960.

Rabczuk T., Bordas S. and Zi G. (2007) A three-dimensional meshfree method for continuous multiple-crack initiation, propagation and junction in statics and dynamics. Comput. Mech 40, 473-495.

Rabczuk T., Zi G., Bordas S. and Nguyen-Xuan H. (2009) A simple and robust three-dimensional cracking-particle method without enrichment. Comput. Methods Appl. Mech. Engrg 199, 2437-2455.

Chessa J. and Belytschko T. (2003) An enriched finite element method for axisymmetric two-phase flow with surface tension. J. Comput. Phys 58, 2041-2064.

Chopp D. and Sukumar N. (2003) Fatigue crack propagation of multiple coplanar cracks with the coupled extended finite element/fast marching method. Int. J. Engrg. Sci 41, 845-869.

Merle R. and Dolbow J. (2002) Solving thermal and phase change problems with the extended finite element method. Comput. Mech 28, 339-350.

Ji, H., Chopp D. and Dolbow J. (2002) A hybrid extended finite element/level set method for modeling phase transformations. Int. J. Numer. Methods Engrg 54, 1209-1233.

Chen J.S., Wu C.T. and Yoon Y. (2001) A stabilized conforming nodal integration for Galerkin meshfree methods. Int. J. Numer. Methods Engrg 50, 435-466.

Liu G.R. and Zhang G.Y. (2009) A normed G space and weakened weak formulation of a cell-based smoothed point interpolation method. Int. J. Comput. Methods 6, 147-179.

Liu G.R. (2010) A G space theory and weakened weak (W2) form for a unified formulation of compatible and incompatible methods, part I: theory and part II: applications to solid mechanics problems. Int. J. Numer. Methods Engrg 81, 1093-1156.

Liu G.R. and Nguyen T.T. (2010) Smoothed Finite Element Method, Taylor & Francis/CRC Press: Boca Raton

Le C.V., Nguyen-Xuan H., Askes H., Bordas. S., Rabczuk T. and Nguyen-Vinh H. (2010) A cell-based smoothed finite element method for kinematic limit analysis. Int. J. Numer. Meth. Engng 83, 1651-1674.

Liu G.R., Chen L., Nguyen T.T., Zeng K.Y. and Zhang G.Y. (2010) A novel singular node-based smoothed finite element method (NS-FEM) for upper bound solutions of fracture problems. Int J Numer Meth Engng 83, 1466-1497.

Chen L., Liu G.R., Nourbakhsh N.N. and Zeng K.Y. (2010) A singular edge-based smoothed finite element method (ES-FEM) for bimaterial interface cracks. Comput Mech 45, 109-125.

Jiang Y., Tay T.E., Chen L. and Sun X.S. (2013) An edge-based smoothed XFEM for fracture in composite materials. Int. J. Fract. 179(1-2), 179-199.

Chen L., Rabczuk T., Bordas S., Liu G.R., Zeng K.Y. and Kerfriden P. (2012) Extended finite element method with edge-based strain smoothing (ESm-XFEM) for linear elastic crack growth. Comput. Methods Appl. Mech. Engrg. 209-212, 250-265.

Liu G.R., Nguyen-Thoi T. and Lam K.Y. (2008) An edge-based smoothed finite element method (ES-FEM) for static, free and forced vibration analyses of solids. J. Sound Vib. 320(4-5), 1100-1130.

Cui X.Y., Liu G.R. and Li G.Y. (2011) Bending and vibration responses of laminated composite plates using an edge-based smoothing technique. Eng. Anal. Bound Elem. 35, 818-826.

Karihaloo B.L. and Xiao Q.Z. (2003) Modelling of stationary and growing cracks in FE framework without remeshing: a state-of-the-art review. Comput. Struct. 81(3), 119-129.

Karihaloo B.L. and Xiao Q.Z. (2001) Accurate determination of the coefficients of elastic crack tip asymptotic field by a hybrid crack element with p-adaptivity. Eng. Fract. Mech 68, 1609-1630.

Xie G.Z., Zhang J.M., Huang C., Lu C.J. and Li G.Y. (2014) A direct traction boundary integral equation method for three-dimension crack problems in infinite and finite domains. Comput. Mech. 53, 575-586.

Aliabadi M.H. (1997) Boundary element formulations in fracture mechanics. Appl. Mech. Rev. 50, 83.

Cruse T.A. (1988) Boundary element analysis in computational fracture mechanics. Vol. 1 Springer Berlin.

Pan E. (1997) A general boundary element analysis of 2-D linear elastic fracture mechanics. Int. J. Fract. 88 (1), 41-59.

Nguyen-Thoi T., Liu G.R., Lam K.Y. and Zhang G.Y. (2009a) A face-based smoothed finite element method (FS-FEM) for 3D linear and geometrically non-linear solid mechanics problems using 4-node tetrahedral elements. Int. J. Numer. Meth. Engng 78, 324-353.

Nguyen-Thoi T., Liu G.R., Vu-Do H.C. and Nguyen-Xuan H. (2009b) A face-based smoothed finite element method (FS-FEM) for visco-elastoplastic analyses of 3D solids using tetrahedral mesh. Comput. Methods Appl. Mech. Engrg 198, 3479-3498

Parks D. M. (1974) A stiffness derivative finite element technique for determination of crack tip stress intensity factor. Int. J. Fracture 10, 487-502

Parks D. M. (1977) The virtual crack extension method for nonlinear material behavior. Comput. Meth. appl. Mech. Engng 12, 353-364.

Hellen T. K. (1975) On the method of virtual crack extension. Int. J. numer. Meth. Engng 9, 187-207.

Hellen T. K. (1989) Virtual crack extension methods for non-linear materials. Int. J. numer. Meth. Engng 28, 929-942.

Rybicki E. F. and Kanninen M. F. (1977) A finite element calculation of stress intensity factors by modified crack closure integral. Engng Fracture Mech. 9, 931-938

Shivakumar K. N., Tan P. W. and Newman J. C., (1988).Virtual crack closure technique for calculating stress-intensity factors for cracked three-dimensional bodies. Int. J. Fracture 36, R43-R50

Raju I. S., Shivakumar K. N. and Crews J. H. Jr. (1988).Three-dimensional elastic analysis of a composite double cantilever beam specimen. AIAA Jl 26, 1493-1498

Buchholz F. G. (1984) Improved formulae for the finite element calculation of the strain energy release rate by the modified crack closure integral method, in Accuracy Reliability and Training in FEM-Technology (Edited by J. Robinson), pp. 650-659. Robinson and Associates, Dorset

Rice J. R. (1968) A path-independent integral and the approximate analyses of strain concentration by notches and cracks. J. appl. Mech. 35, 376-386

Cherepanov G. P. (1967) Crack propagation in continuous media. J. appl. Math. Mech. 31, 503-512

Eshelby J. D. (1956) The continuum theory of lattice defects, in Solid State Physics (Edited by F. Seitz and D. Turnbull), Vol. III, pp. 79-114. Academic Press, New York.

Cherepanov G. P. (1979) Mechanics of Brittle Fracture (translated and Edited by R. W. de Wit and W. C. Cooly). McGraw-Hill, New York.

Anderson T. L. (1995) Fracture mechanics: Fundamentals and applications. Taylor & Francis Group

DISCONTINUOUS GALERKIN FINITE VOLUME ELEMENT METHODS FOR ELLIPTIC OPTIMAL CONTROL PROBLEMS

RUCHI SANDILYA * AND SARVESH KUMAR*

ABSTRACT. In this paper, we have discussed a one parameter family of discontinuous Galerkin finite volume element methods for the approximation of the solution of distributed optimal control problems governed by a class of second order linear elliptic equations. In order to approximate the control problem, the method of variational discretization is used. By following the analysis of Kumar *et. al.* [Numer. Meth. Part. Diff. Eqns. 25 (2009), pp. 1402–1424], optimal order of convergence in L^2-norm for state, costate and control variables are derived. Moreover, optimal order of convergence in broken H^1-norm are also derived for state and costate variables. Several numerical experiments are presented to validate the theoretical order of convergence.

Keywords: Optimal control; variational discretization; discontinuous Galerkin finite volume element methods; order of convergence; numerical experiments.

1. INTRODUCTION

This paper is concerned with the discontinuous Galerkin finite volume element (DGFVE) approximation of the elliptic optimal control problem of the following type : Find y, u such that

$$\min_{u \in U_{ad}} \frac{1}{2} \|y - y_d\|^2_{L^2(\Omega)} \quad + \quad \frac{\lambda}{2} \|u\|^2_{L^2(\Omega)}, \tag{1.1}$$

subject to

$$-\nabla.(K\nabla y) = Bu + f \quad in \, \Omega, \tag{1.2}$$

$$y = 0 \quad on \, \Gamma. \tag{1.3}$$

where, $\Omega \subset \mathbb{R}^2$ is a convex, bounded and polygonal domain and Γ is the boundary of Ω, λ is a positive number, $f, y_d \in L^2(\Omega)$ or $H^1(\Omega)$, $K = (k_{ij}(x))_{2\times 2}$ denotes a real valued, symmetric and uniformly positive definite matrix in Ω, i.e., there exists a positive constant α_0 such that

$$\xi^T K \xi \geq \alpha_0 \xi^T \xi \quad \forall \xi \in \mathbb{R}^2.$$

B is a bounded continuous linear operator and U_{ad} is denoted by

$$U_{ad} = [u \in L^2(\Omega) : a \leqslant u(x) \leqslant b, \, a.e. \, in \, \Omega, \, a, b \in \mathbb{R}].$$

The numerical solutions of such kind of elliptic problems have been investigated by many researchers, since these problems have lots of applications in mathematical and physical problems. Finite element methods extensively used for the approximation of the control problems and for the error analysis of finite element methods (FEM) applied to elliptic control problems, we refer to [3, 4, 5, 6, 7, 15] and references therein. In most of these papers, the state and costate variables are discretized by continuous linear elements and control variable by piecewise constant or piecewise linear polynomials. More recently, Hinze given a new direction for approximating the control problem in which a new variational discretization approach is introduced for

---

*Department of Mathematics, Indian Institute of Space Science and Technology, Thiruvananthapuram 695 547, Kerala, India. Email:sarvesh@iist.ac.in and ruchisandilya.12@iist.ac.in.

linear-quadratic optimal control problems whereas the control set is not discretized explicitly and obtained improved convergence order for optimal control, for more details, kindly see [8].

Because of local conservative property of the finite volume element (FVE) methods, these methods are very popular in computational fluid dynamics and (FVE) methods have also been used to solve fluid optimal control problems. In [12], the author has used the *optimize-then-discretize* approach and FVE discretizations to approximate elliptic optimal control problems.

It is well known the discontinuous Galerkin (DG) methods which was introduced by Arnold in [1] does not demand the inter element continuity criteria and has some attractive features such as: high order accuracy, localizabilty and suitable for parallel computing easily handle the boundary conditions. Keeping in mind the advantages of FVE methods and DG methods, in [16], Ye introduced discontinuous Galerkin finite volume element (DGFVE) methods for elliptic problems. Later Kumar *et. al.* [9] have discussed a one parameter family of DGFVE methods for the approximation of the elliptic problem. Recently, Kumar extended the analysis of [9] for approximation of miscible displacement problems, see [10].

In this paper, in order to obtain an optimal system, first we apply Lagrange multiplier method to the problem (1.1Introductionequation.1.1)-(1.3Introductionequation.1.2) and obtain an optimal system. Then we use DGFVE methods to discretize the state and adjoint equation of the system. For the optimal condition, we use variational discretization approach introduced in [8] to obtain the control.This paper is organized as follows: While the Section 1 is introductory, Section 2 is devoted to the DGFVE formulation for the optimal control problem. In Section 3, we discuss the convergence analysis of DGFVE in different norms and finally in Section 4, we present some numerical experiments to support the theoretical results obtained in Section 3.

2. Discontinuous Galerkin Finite Volume Element Formulation

We assume that our optimal control problem admits a unique control u, since U_{ad} is bounded, convex and closed. For the subsequent standard existence, uniqueness and first-order optimality results we refer to [14]. We can then write the first-order optimality condition in the following form:

$$(\lambda u + B^* p, v - u) \geq 0 \quad \forall v \in U_{ad}, \tag{2.1}$$

where the function p is called *adjoint state* (or *costate*) associated with u and solution of the *adjoint equation*

$$-\nabla.(K \nabla p) = y - y_d, \quad in \, \Omega \tag{2.2}$$

$$p = 0, \quad on \, \Gamma. \tag{2.3}$$

Let τ_h be a regular, quasi-uniform triangulation of $\bar{\Omega}$ into closed triangles T with $h = \max_{T \in \tau_h}(h_T)$, where h_T is the diameter of the triangle T. The dual partition τ_h^* of τ_h is constructed as follows: divide each triangle $T \in \tau_h$ into three triangles by joining the barycenter B and the vertices of T as shown in Figure 1A triangular partition and its dualfigure.1. Let τ_h^* consists of all these triangles T_i^*. We define the finite dimensional Trial (V_h) and test space (W_h) associated with τ_h and τ_h^*, respectively as follows:

$$V_h = \{v_h \in L^2(\Omega) : v_h|_T \in P_1(T) \quad \forall T \in \tau_h\}$$
$$W_h = \{w_h \in L^2(\Omega) : w_h|_{T^*} \in P_0(T^*) \quad \forall T^* \in \tau_h^*\}.$$

where $P_m(T)$ or $P_m(T^*)$ denotes the space of all polynomials of degree less than or equal to m defined on T or T^*, respectively. Let $V(h) = V_h + H^2(\Omega) \cap H_0^1(\Omega)$. To connect the trial space and test space, we define a transfer operator $\gamma : V(h) \longrightarrow W_h$ as:

$$\gamma v|_{T^*} = \frac{1}{h_e} \int_e v|_{T^*} ds, \quad T^* \in \tau_h^*,$$
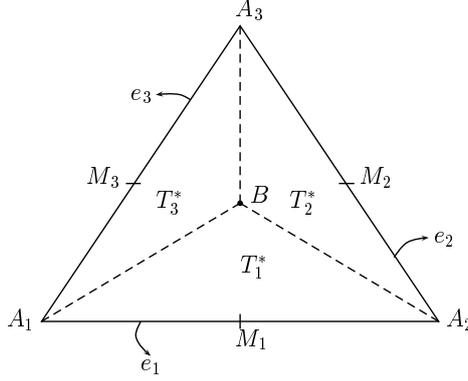
FIGURE 1. A triangular partition and its dual

where e is an edge in T, T^* is the dual element in τ_h^* containing e, and h_e is the length of the edge e.

Multiply (1.2Introductionequation.1.2) and (2.2Discontinuous Galerkin Finite Volume Element Formulationequation.2.2) by γv_h, integrate over the control volumes and an application of Gauss divergence methods leads the following DGFVE formulation: Find $(y_h, p_h, u_h) \in V_h \times V_h \times U_{ad}$ such that

$$A_h(y_h, w_h) = (Bu_h + f, \gamma w_h) \quad \forall w_h \in V_h, \tag{2.4}$$

$$A_h(p_h, q_h) = (y_h - y_d, \gamma q_h) \quad \forall q_h \in V_h, \tag{2.5}$$

$$(\lambda u_h + B^* p_h, v - u_h) \geq 0 \quad \forall v \in U_{ad}, \tag{2.6}$$

where the bilnear form $A_h(\cdot, \cdot)$ defined as

$$A_h(\Phi_h, \Psi_h) = -\sum_{T \in \tau_h} \sum_{j=1}^{3} \int_{A_{j+1}BA_j} (K\nabla\Phi_h.\mathbf{n})\gamma\Psi_h ds + \theta \sum_{e \in \Gamma} \int_e [\gamma\Phi_h].\langle K\nabla\Psi_h\rangle ds$$

$$-\sum_{e \in \Gamma} \int_e [\gamma\Psi_h].\langle K\nabla\Phi_h\rangle ds + \sum_{e \in \Gamma} \int_e \frac{\alpha}{h_e^\beta}[\Phi_h].[\Psi_h]ds \quad \forall \Phi_h, \Psi_h \in V_h.$$

Here, the symbols $[\cdot]$ and $\langle\cdot\rangle$ used for jump and average respectively and $\theta \in [-1, 1]$, α and β are penalty parameters, for more details kindly see [9]. Let $y_h(u)$ and $p_h(y)$ be the solutions of

$$A_h(y_h(u), w_h) = (Bu + f, \gamma w_h) \quad \forall w_h \in V_h, \tag{2.7}$$

and

$$A_h(p_h(y), q_h) = (y - y_d, \gamma q_h) \quad \forall q_h \in V_h, \tag{2.8}$$

respectively. A norm $\|.\|$ on V(h) is defined by

$$\|v\|^2 = |v|_{1,h}^2 + \sum_{e \in \Gamma} \frac{1}{h_e^\beta} \int_e [v]^2 ds,$$

where $|v|_{1,h}^2 = \sum_{T \in \tau_h} |\nabla v|_{0,T}^2$. Using the coercivity and boundedness of the bilinear form $A_h(\cdot, \cdot)$ which is proved in [9, pp. 1410–1413] and noting that $y_h = y_h(u_h)$ and $p_h = p_h(y_h)$ we have the following result.

Lemma 2.1. *Let $y_h(u)$ and $p_h(y)$ be the solutions of (2.7Discontinuous Galerkin Finite Volume Element Formulationequation.2.7) and (2.8Discontinuous Galerkin Finite Volume Element Formulationequation.2.8) respectively. Then the following results hold :*

$$\|p_h(y) - p_h\| \leq C\|y - y_h\| \quad and \quad \|y_h(u) - y_h\| \leq C\|u - u_h\|.$$

The result easily follows by using (Theorem 2.3, [9]) and Cauchy-Schwarz inequality.

We emphasis that throughout the article C is a generic positive constant (also appeared in Lemma 2.1theorem.2.1) which is independent of the mesh size h but may depend on the bounds of f, u, y, p and size of the domain Ω.

3. CONVERGENCE ANALYSIS

3.1. Convergence in L^2-norm.

Theorem 3.1. *Assume that $K \in W^{1,\infty}(\Omega)$ and $u, f, y_d \in L^2(\Omega)$. Let $(y, p, u) \in (H^2(\Omega) \cap H_0^1(\Omega)) \times (H^2(\Omega) \cap H_0^1(\Omega)) \times U_{ad}$ be the exact solutions and $(y_h, p_h, u_h) \in V_h \times V_h \times U_{ad}$ be the solutions of (2.5Discontinuous Galerkin Finite Volume Element Formulationequation.2.4)-(2.6Discontinuous Galerkin Finite Volume Element Formulationequation.2.6). Then there exists an $h_0 > 0$ such that for all $0 < h \le h_0$*

$$\|u - u_h\| \le Ch. \tag{3.1}$$

Moreover, if $K \in W^{2,\infty}(\Omega)$ and $u, f, y_d \in H^1(\Omega)$, then there exists an $h_0 > 0$ such that for all $0 < h \le h_0$

$$\|u - u_h\| \le Ch^2. \tag{3.2}$$

The above theorem can be proved by using the variational inequalities (2.1Discontinuous Galerkin Finite Volume Element Formulationequation.2.1) and (2.6Discontinuous Galerkin Finite Volume Element Formulationequation.2.6) with the functions u and u_h, using (Lemma 2.4, Theorem 3.2, [9]) and Lemma 2.1theorem.2.1. For more details, we refer to [11].

Now, using triangle inequality, (Theorem 3.2, [9]), Lemma 2.1theorem.2.1 and Theorem 3.1theorem.3.1, we have the following theorem.

Theorem 3.2. *Assume that $K \in W^{1,\infty}(\Omega)$ and $u, f, y_d \in L^2(\Omega)$. Let $(y, p, u) \in (H^2(\Omega) \cap H_0^1(\Omega)) \times (H^2(\Omega) \cap H_0^1(\Omega)) \times U_{ad}$ be the exact solutions and $(y_h, p_h, u_h) \in V_h \times V_h \times U_{ad}$ be the solutions of (2.5Discontinuous Galerkin Finite Volume Element Formulationequation.2.4)-(2.6Discontinuous Galerkin Finite Volume Element Formulationequation.2.6). Then there exists an $h_0 > 0$ such that for all $0 < h \le h_0$*

$$\|y - y_h\| \le Ch, \quad \|p - p_h\| \le Ch. \tag{3.3}$$

Moreover, if $K \in W^{2,\infty}(\Omega)$ and $u, f, y_d \in H^1(\Omega)$, then there exists an $h_0 > 0$ such that for all $0 < h \le h_0$

$$\|y - y_h\| \le Ch^2, \quad \|p - p_h\| \le Ch^2. \tag{3.4}$$

Following the proof lines of (Theorem 3.1, [9]) and using Theorem 3.1theorem.3.1, Theorem 3.2theorem.3.2 together with Lemma 2.1theorem.2.1, we can derive the following error estimates in the H^1-norm. For a detailed proof, we refer to [11].

3.2. Convergence in broken H^1-norm.

Theorem 3.3. *Assume that $K \in W^{1,\infty}(\Omega)$ and $u, f, y_d \in L^2(\Omega)$. Let $(y, p, u) \in (H^2(\Omega) \cap H_0^1(\Omega)) \times (H^2(\Omega) \cap H_0^1(\Omega)) \times U_{ad}$ be the exact solutions and $(y_h, p_h, u_h) \in V_h \times V_h \times U_{ad}$ be the solutions of (2.5Discontinuous Galerkin Finite Volume Element Formulationequation.2.4)-(2.6Discontinuous Galerkin Finite Volume Element Formulationequation.2.6). Then there exists an $h_0 > 0$ such that for all $0 < h \le h_0$*

$$\|y - y_h\| \le Ch, \quad \|p - p_h\| \le Ch. \tag{3.5}$$

4. Numerical Experiments

In this section, we present two numerical examples in order to discuss the performance of the DGFVE for the approximation of the elliptic optimal control problem (1.1Introductionequation.1.1)-(1.3Introductionequation.1.2). The method holds true for any value of $\theta \in [-1, 1]$ but in particular, for the numerical experiments we take $\theta = $-1,0,1, as these values of θ leads to different interesting schemes in the context of discontinuous finite element methods, kindly see [13]. We will investigate the order of convergence of state, costate and control variables in L^2-norm and order of convergence of state and costate variables in the broken norm $\|.\|$.

Example 1. We consider the following elliptic control problem with Dirichlet boundary value condition:

$$\min_{u \in U_{ad}} \frac{1}{2} \|y - y_d\|_{L^2(\Omega)}^2 + \frac{1}{2} \|u\|_{L^2(\Omega)}^2,$$

$$\begin{aligned}
-\triangle y &= u & in\,\Omega, \\
y &= 0 & in\,\Gamma, \\
u &\geq 0,
\end{aligned}$$

where $\Omega = [(x_1, x_2) : 0 \leq x_1 \leq 1, 0 \leq x_2 \leq 1]$, Γ denotes the boundary of Ω. The exact state y is $sin(\pi x_1)sin(\pi x_2)$, $y_d = (4\pi^4 + 1)sin(\pi x_1)sin(\pi x_2)$, $p = -2\pi^2 sin(\pi x_1)sin(\pi x_2)$ and $u = max(0, -p)$.
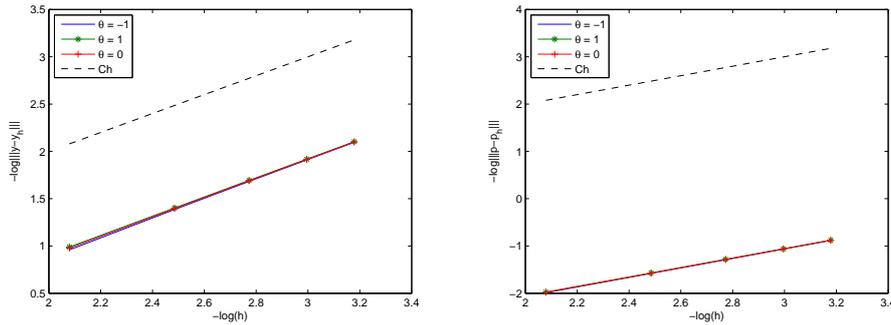
FIGURE 2. Order of convergence in broken H^1-norm for state and costate variables for Example 1.

$\|y - y_h\|$			$\|p - p_h\|$		
θ=-1	θ=0	θ=1	θ=-1	θ=0	θ=1
0.3719918	0.37650518	0.38334068	7.21776447	7.24618657	7.29030071
0.2460884	0.24772621	0.25033399	4.82156990	4.83286579	4.85124495
0.1839629	0.18473849	0.18600306	3.61641018	3.62227458	3.63199956
0.1469065	0.14733895	0.14805334	2.89231763	2.89585972	2.90178678
0.1222834	0.12255207	0.12299930	2.40948560	2.41184101	2.41580092

TABLE 1. Numerical results of broken H^1 error for θ=1, θ=-1 and θ=0 with β=1 and $\alpha = 10$ for Example 1.
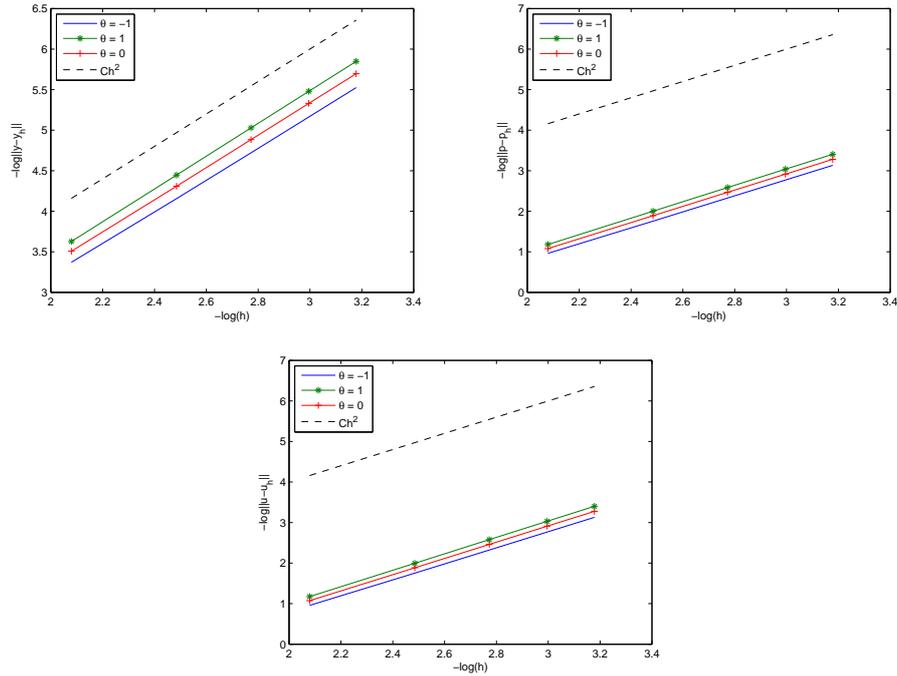
FIGURE 3. Order of convergence in L^2-norm for state and costate and control variables for Example 1.

$\|y - y_h\|$				$\|p - p_h\|$			
Dof	θ=-1	θ=0	θ=1	Dof	θ=-1	θ=0	θ=1
384	0.02657497	0.02999063	0.03437191	384	0.30726316	0.33952145	0.38187541
864	0.01172431	0.01344455	0.01568472	864	0.13500845	0.15109265	0.17258162
1536	0.00654811	0.00756756	0.00890502	1536	0.07535372	0.08485681	0.09766827
2400	0.00416848	0.00483938	0.00572336	2400	0.04797144	0.05421803	0.06268522
3456	0.00288324	0.00335721	0.00398345	3456	0.03318749	0.03759811	0.04359833

$\|u - u_h\|$		
Dof	θ=-1	θ=0	θ=1
384	0.30968478	0.34176897	0.38392985
864	0.13604558	0.15206116	0.17347429
1536	0.07586822	0.08533866	0.09811434
2400	0.04825989	0.05448858	0.06293629
3456	0.03336438	0.03776417	0.04375267

TABLE 2. Numerical results of L^2 error for θ=1, θ=-1 and θ=0 with $\beta = 1$ and $\alpha = 10$ for Example 1.

In the next example we take desired state y_d to be zero and include desired control u_0.

Example 2. We consider the following problem

$$\min_{u \in U_{ad}} \frac{1}{2} \|y - y_d\|_{L^2(\Omega)}^2 + \frac{1}{2} \|u - u_0\|_{L^2(\Omega)}^2,$$

$$
\begin{aligned}
-\triangle y &= u + f \quad in\,\Omega, \\
y &= 0 \quad in\,\Gamma, \\
u &\geq 1 \quad in\,\Omega.
\end{aligned}
$$

In this example we have,
$\Omega = [(x_1, x_2) : 0 \leq x_1 \leq 1, 0 \leq x_2 \leq 1]$, $u_0 = 1 - sin(\pi x_1/2) - sin(\pi x_2/2) + s$, $y_d = 0$,
$p = Z(x_1, x_2)$, $f = 4\pi^4 Z - u$, where $Z = sin(\pi x_1)sin(\pi x_2)$ and

$$
s = \begin{cases} 0.5 & if\,x_1 + x_2 > 1.0 \\ 0.0 & if\,x_1 + x_2 \leq 1.0 \end{cases}
$$

. The exact solution of this problem is $y = 2\pi^2 Z$, $u = max(u_0 - p, 1)$.
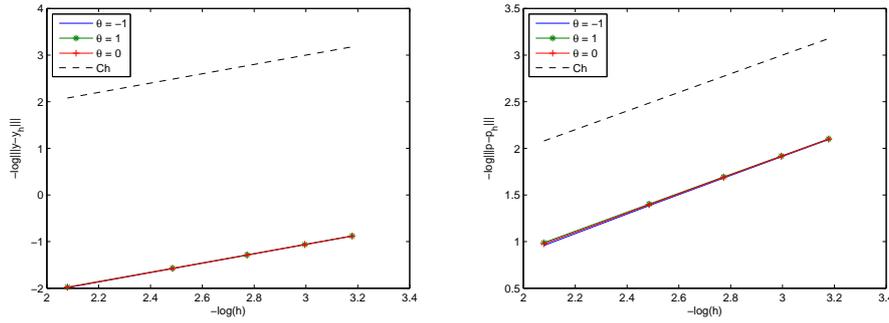
FIGURE 4. Order of convergence in broken H^1-norm for state and costate variables for Example 2.

	$\|y - y_h\|$			$\|p - p_h\|$		
Dof	θ=-1	θ=0	θ=1	Dof	θ=-1	θ=0	θ=1
384	7.21797574	7.24658386	7.29098777	384	0.37204143	0.37657411	0.38343755
864	4.82162819	4.83298774	4.85147100	864	0.24610328	0.24774803	0.25036647
1536	3.61643379	3.62232663	3.63209930	1536	0.18396914	0.18474788	0.18601743
2400	2.89232940	2.89588651	2.90183911	2400	0.14690967	0.14734380	0.14806087
3456	2.40949230	2.41185655	2.41583166	3456	0.12228530	0.12255488	0.12300372

TABLE 3. Numerical results of broken H^1 error for θ=1, θ=-1 and θ=0 with $\beta = 1$ and $\alpha = 10$ for Example 2.

The errors in broken H^1-norm for the DGFVEM solution of state and costate variables are presented in Tables 1Numerical results of broken H^1 error for θ=1, θ=-1 and θ=0 with β=1 and $\alpha = 10$ for Example 1table.1 and 3Numerical results of broken H^1 error for θ=1, θ=-1 and θ=0 with $\beta = 1$ and $\alpha = 10$ for Example 2table.3 for examples 1 and 2, respectively whereas the errors in L^2-norm for the DGFVEM solution of state, costate and control variables for examples 1 and 2 are presented in Tables 2Numerical results of L^2 error for θ=1, θ=-1 and θ=0 with $\beta = 1$ and $\alpha = 10$ for Example 1table.2 and 4Numerical results of L^2 error for θ=1, θ=-1 and θ=0 with $\beta = 1$ and $\alpha = 10$ for Example 2table.4 respectively.
Figures 2Order of convergence in broken H^1-norm for state and costate variables for Example 1figure.2, 3Order of convergence in L^2-norm for state and costate and control variables for Example 1figure.3 (for Example 1) and 4Order of convergence in broken H^1-norm for state and costate variables for Example 2figure.4, 5Order of convergence in L^2-norm for state and costate and control variables for Example 2figure.5 (for Example 2) indicate that the computed orders of convergence match the theoretical orders of convergence in L^2-norm and broken H^1-norm.
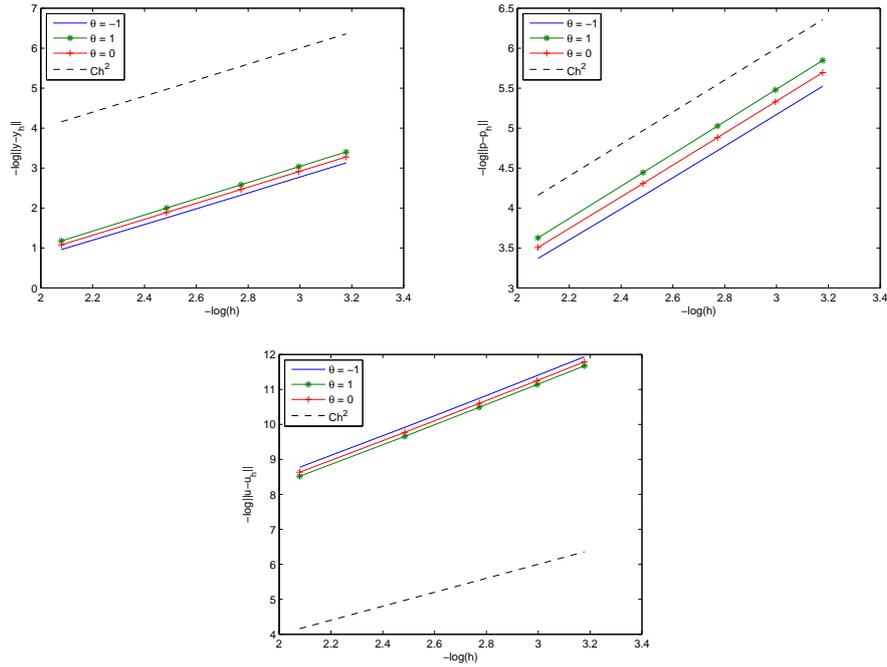
FIGURE 5. Order of convergence in L^2-norm for state and costate and control variables for Example 2.

$\|y - y_h\|$				$\|p - p_h\|$			
Dof	θ=-1	θ=0	θ=1	Dof	θ=-1	θ=0	θ=1
384	0.30825840	0.34069258	0.38326672	384	0.02663156	0.03005540	0.03444685
864	0.13545263	0.15162708	0.17323066	864	0.01174998	0.01347458	0.01572027
1536	0.07560241	0.08515929	0.09803949	1536	0.00656257	0.00758466	0.00892548
2400	0.04812983	0.05441189	0.06292457	2400	0.00417772	0.00485037	0.00573660
3456	0.03329703	0.03773274	0.04376519	3456	0.00288965	0.00336485	0.00399269

$\|u - u_h\|$		
Dof	θ=-1	θ=0	θ=1
$1.0e - 003 *$		
384	0.20013750	0.17824991	0.15447196
864	0.06389361	0.05698655	0.04942176
1536	0.02785402	0.02484840	0.02154210
2400	0.01452166	0.01295391	0.01122471
3456	0.00849969	0.00758109	0.00656606

TABLE 4. Numerical results of L^2 error for θ=1, θ=-1 and θ=0 with $\beta = 1$ and $\alpha = 10$ for Example 2.

REFERENCES

[1] D. N. ARNOLD,*An interior penalty finite element method with discontinous elements*, SIAM J. Numer. Anal. (1982), **19**:742–760.
[2] S. BRENNER AND R. SCOTT, *The mathematical theory of finite element methods*, Springer, New York, 2002.

[3] E. Casas and F. Tröltzsch, *Error estimates for the finite-element approximation of a semilinear elliptic control problem*, Control and Cybernetics (2002), **31**:695–712.

[4] E. Casas and F. Tröltzsch, *Error estimates for linear-quadratic elliptic control problems*, IFIP: Anal. and Optimization of Diff. Systems (2003), **121**:89–100.

[5] R.S. Falk, *Approximation of a class of optimal control problems with order of convergence estimates*, Journal of Mathem. Anal. and Applications (1973), **44**:28–47.

[6] T. Geveci, *On the approximation of the solution of an optimal control problem governed by an elliptic equation*, ESIAM: Mathem. Modelling and Numer. Anal.(1979), **13**:313–328.

[7] W. Gong and N. Yan, *Robust error estimates for the finite element approximation of elliptic optimal control problems*, J. Comput. Appl. Math.(2011), **236**:1370–1381.

[8] M. Hinze, *A variational discretization concept in control constrained optimization: the linear-quadratic case* , Computational Optimization and Applications (2005), **30**:45–61.

[9] S. Kumar, N. Nataraj and A.K. Pani, *Discontinuous Galerkin finite volume element methods for second order linear elliptic problems*, Numer. Meth. Part. Diff. Eqns. (2009), **25**:1402–1424.

[10] S. Kumar, *A mixed and discontinuous Galerkin finite volume element method for incompressible miscible displacement problems in porous media*, Numer. Meth. Part. Diff. Eqns. (2012), **28**:1354–1381.

[11] R. Sandilya and S. Kumar, *Convergence analysis of discontinuous finite volume methods for elliptic control problems*, International Journal of Computational Methods (2015), **13**:1640012–20.

[12] X. Luo, Y. Chen and Y. Huang, *Some error estimates of finite volume element approximation for elliptic optimal control problems*, Int. J. Numer. Anal. Model. (2013), **10**:697–711.

[13] S. Prudhomme, F. Pascal, J.T. Oden and A. Romkes, *A priori error estimation for discontinuous Galerkin Methods*, TICAM Report, (2000).

[14] F. Tröltzsch, *Optimal control of partial differential equations: theory, methods and applications*, American Math. Society, book series Graduate Studies in Mathematics, to appear 2010.

[15] F. Tröltzsch, *On finite element error estimates for optimal control problems with elliptic PDEs*, Springer, Large Scale Scientific Computing, Lecture Notes (2010), **5910**:40–53.

[16] X. Ye, *A new discontinuous finite volume method for elliptic problems*, SIAM J. Numer. Anal. (2004), **42**:1062–1072.

A POD-based reduction approach for multiscale nonlinear structural design

*Liang Xia, and Piotr Breitkopf

Laboratoire Roberval, UMR 7337 UTC-CNRS, Université de Technologie de Compiègne,
Centre de Recherches de Royallieu, CS 60319, 60203, Compiègne Cedex, France

*Corresponding author: liang.xia@utc.fr

Abstract

This paper presents POD-based reduction approach for structural optimization design considering microscopic material nonlinear microstructures. This work introduces Reduced Order Model (ROM) to alleviate the heavy computational demand of nonlinear nested multiscale procedures, particularly in an optimization framework which requires multiple loops involving similar computations. The surrogate model constructed using Proper Orthogonal Decomposition (POD) and Diffuse Approximation reduces the computational effort for solving the microscopic boundary value problems. Multiscale analysis model (FE^2) is applied to link structure and microstructures in the two scales. Maximum stiffness design of the macroscopic structure is realized using a discrete level-set topology optimization model. It is shown by means of numerical tests that the reduced multiscale model provides reasonable designs as compared to those obtained by the unreduced model while with a significantly reduced computational effort.

Keywords: Model reduction, Diffuse Approximation, Multiscale analysis, Topology optimization

Introduction

Optimization techniques for structural size, shape, topology designs have been widely developed and employed in engineering applications. One of its most prominent applications is designing lightweight structures for aircrafts. An increasing number of optimized structures, parts and components appear in the latest models of Airbus and Boeing. Most of present optimization algorithms are developed within frameworks of numerical analysis with the assumption that the considered structure is constituted by one-scale linear elastic materials. However, due to the fast development made in the field of material science, advanced fiber-reinforced composites are increasingly used in both aerospace and military applications. More advanced structural analysis models are required such that the structural influences from microscopic heterogeneities can be considered. As a response, multiscale incremental homogenization approaches or the so called FE^2 approach have been proposed and largely developed in the last decade [Feyel and Chaboche (2000), Kouznetsova et al. (2001)]. Generally speaking, this type of approach solves two nested boundary value problems, one at the macroscopic scale and another at the microscopic scale. The FE^2 approach is able to evaluate the macroscopic responses of heterogeneous material with an accurate account for micro characteristics and evolution of the morphology. The challenges of the FE^2 approach are due to high computational effort. Therefore, there is an increasing research demand of bridging structural optimization models and FE^2-type analysis models.

This paper introduces Reduced Order Model (ROM) to perform multiscale topological optimization design. The multiscale analysis model FE^2 [Feyel and Chaboche (2000)] is applied to link the macroscopic structure and the corresponding RVE microstructures in the microscopic level. The optimization process requires multiple design loops involving similar or even repeated computations of the RVE which perfectly suits the ROM learning process. In the present work the considered RVE is assumed to be the same for all marcoscopic integration points. Maximum stiffness design of the macroscopic structure is performed using a discrete level-set topology

optimization model [Challis (2010)]. The reduction is performed in an adaptive non-intrusive manner which is an alternative to the intrusive approach [Yvonnet and He (2007)}]. The reduced basis is extracted using Proper Orthogonal Decomposition (POD) and the surrogate model is constructed using Diffuse Approximation [Nayroles et al. (1992)}, variant of Moving Least Squares [Lancaster and Salkauskas (1981)].

The remainder of this paper is organized in the following manner: firstly the FE^2 approach is briefly reviewed which links the macroscopic structure and microscopic microstructure RVE; secondly, the discrete level-set model for structural topology optimization design is presented; thirdly, a bi-level reduced surrogate model is developed for microscopic RVE solution using POD and Diffuse Approximation; the presented model is then showcased by one numerical test example; finally, the paper ends with concluding comments and suggestions for future work.
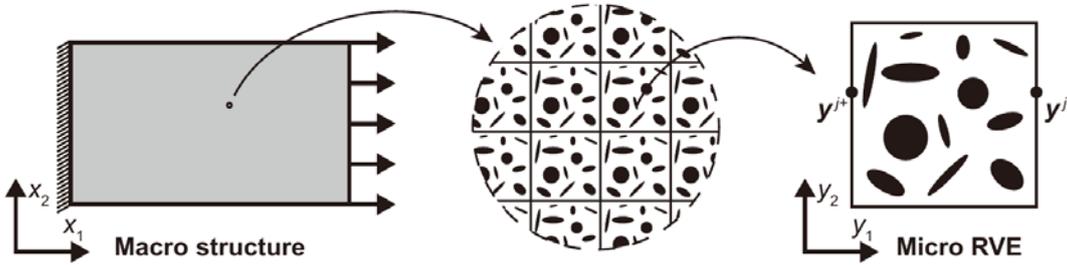
Figure 1. Illustration of the selection of a typical 2D representative volume element (RVE).

FE^2 approach

The FE^2 approach [Feyel and Chaboche (2000)] is chosen here to bridge the macroscopic structure and the corresponding microscopic RVE to perform structural topology optimization. The key hypothesis of FE^2 consists in the separation of macroscopic and microscopic scales. It is assumed that the microscopic length scale is large enough to be considered in the framework of continuum mechanics, and at the same time much smaller than the macroscopic length scale considered in terms of periodically ordered pattern [Kouznetsova et al. (2001)], as illustrated in Fig. 1.

The principal concept of the FE^2 approach assumes that each macroscopic material point is attributed with a RVE so that the macroscopic stress and strain for the considered point can be estimated by averaging the corresponding stress and strain fields of the RVE. Thereafter, there is no need to specify the macroscopic constitutive behavior and we only need to define the constitutive behavior for each material phase of the RVE. Let x and y denote the position of a point at the macroscopic and microscopic scales, respectively. At the macroscopic scale, stress and strain fields are denoted by $\Sigma(x)$ and $E(x)$, which are evaluated as the average of the corresponding microscopic fields $\sigma(x,y)$ and $\varepsilon(x,y)$ over the RVE of region Ω_x corresponding to the material point x. The FE^2 performs the following steps:

- evaluate the macroscopic strain field $E(x)$ with an initially defined elastic tensor C_0;
- define boundary conditions on the RVE at material point x upon the value of $E(x)$;
- evaluate the stress field $\sigma(x,y)$ through periodic homogenization analysis on the RVE;
- compute the macroscopic stress tensor $\Sigma(x)$ at material point x via averaging $\sigma(x,y)$;
- update the structural displacement field $u(x)$ using iterative Newton-Raphson method;
- repeat above procedures until the macroscopic force equilibrium is achieved.
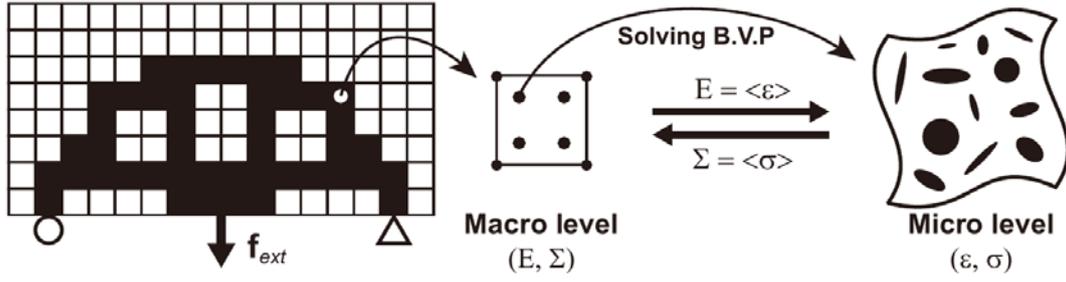
Figure 2. Illustration of the implementation of FE^2 in the framework of FEA.

A schematic view of the FE^2 algorithm is depicted in Fig. 2, where each Gauss integration point is attributed with an RVE within the context of finite element analysis (FEA). In case of nonlinear elasticity, the displacement solution at the macroscopic scale is solved using the iterative Newton-Raphson method.

Nonlinear structural design using level-set method

In order to avoid defining a pseudo-relationship between the intermediate values and the considered RVE, we choose to use the discrete version of level-set topology optimization model [Challis (2010)] to straightforwardly link RVEs to the solid region of the structure. An initial level-set function $\psi(x, t_0)$ is constructed as a signed distance function upon the discretized initial structural topology following

$$\begin{cases} \psi(\mathbf{x}_e, t_0) < 0 & \text{if } \rho_e = 1 \\ \psi(\mathbf{x}_e, t_0) > 0 & \text{if } \rho_e = 0 \end{cases} \tag{1}$$

where $\mathbf{x}_e$ denotes the center of the eth element and ρ_e is its pseudo-density. The initialized level-set function $\psi(\mathbf{x}_e, t_0)$ is then be updated to $\psi(\mathbf{x}_e, t)$ corresponding a new structural topology by solving the ``Hamilton-Jacobi'' evolution equation

$$\frac{\partial \psi(x, t)}{\partial t} + v_n |\nabla \psi(x, t)| = 0 \tag{2}$$

where t is a pseudo-time defined corresponding to different optimization iterations. The normal velocity field v_n determines geometric motion of the boundary of the structure and is chosen based on the shape derivative of the design objective. Within the context of multiscale analysis, the optimization objective corresponding to stiffness maximization or compliance minimization can be written in terms of $\boldsymbol{\rho}(\psi)$

$$\begin{aligned} \min_{\boldsymbol{\rho}(\psi)}: \quad & c(\boldsymbol{\rho}(\psi)) = \mathbf{f}_{ext}^T \mathbf{u} \\ s.t.: \quad & \mathbf{R}(\mathbf{u}, \boldsymbol{\rho}(\psi)) = \mathbf{0} \\ & V(\boldsymbol{\rho}(\psi)) = \sum_{e=1}^{N} \rho_e = V_{req} \\ & \rho_e = 0 \text{ or } 1, \forall e = 1, \ldots, N, \end{aligned} \tag{3}$$

where $\boldsymbol{\rho} = (\rho_1, \ldots, \rho_N)$ is the vector of the element pseudo-densities. In the following, we will denote $\boldsymbol{\rho}(\psi)$ by $\boldsymbol{\rho}$ to alleviate the notation. The objective $c(\boldsymbol{\rho}(\psi))$ is twice of the strain energy. The macroscopic structural stiffness is maximized in terms of minimizing the global strain energy. $V(\boldsymbol{\rho})$ is the total number of solid elements and V_{req} is the required number of solid elements. $\mathbf{u}$ is the final converged displacement solution. $\mathbf{R}(\mathbf{u}, \boldsymbol{\rho})$ stands for the force residual at the macroscopic scale

$$\mathbf{R}(\mathbf{u}, \boldsymbol{\rho}) = \mathbf{f}_{ext} - \sum_{e=1}^{N} \int_{\Omega_e} \mathbf{B}^T \rho_e \langle \boldsymbol{\sigma}(\mathbf{x}, \mathbf{y}) \rangle d\Omega_e. \tag{4}$$

An augmented Lagrangian method is applied to convert the original constrained optimization problem Eq. (3) into an unconstrained problem as presented in [Belytschko et al. (2003)].

Bi-level reduced surrogate model

A bi-level reduced surrogate model is constructed coupling the POD and Diffuse Approximation procedures. The first level of reduction is achieved by Proper Orthogonal Decomposition (POD), allowing to expand a displacement field as a linear combination of the truncated modes. Secondly, the surrogate model based on Diffuse Approximation is built to express the POD projection coefficients as functions of the average micro strain tensors.

Proper Orthogonal Decomposition of RVE displacement field

We consider a D-dimensional (D = 2 or 3) RVE of N points subjected to a time-dependent loading $\mathbf{E}(t) = \langle \boldsymbol{\varepsilon} \rangle(t)$ during a time interval $I = [0, T]$ discretized by M instants $\{t_1, t_2, \ldots, t_M\}$. Let $\mathbf{u}_i \in \mathbb{R}^{DN}$ denote the DN-dimensional nodal displacement vector recorded at the instant t_i. The reduced order displacement vector $\mathbf{u}^R(t) \in \mathbb{R}^{DN}$ may be written

$$\mathbf{u}^R(t) = \mathbf{u}_0 + \sum_{i=1}^{m} \boldsymbol{\phi}_i \alpha_i(\langle \boldsymbol{\varepsilon} \rangle(t)), \tag{5}$$

where $m \ll \min(M, DN)$, $\mathbf{u}_0 = 1/M \sum_{i=1}^{M} \mathbf{u}_i$ $\boldsymbol{\phi}_i \in \mathbb{R}^{DN}$ are constant vectors and coefficients $\alpha_i(\langle \boldsymbol{\varepsilon} \rangle(t))$ are scalar functions of pseudo-time t. $\boldsymbol{\phi}_i$ are the eigenvectors of the eigenvalue problem

$$\mathbf{C}_u \boldsymbol{\phi}_i = \lambda_i \boldsymbol{\phi}_i, \tag{6}$$

where $\mathbf{C}_u$ is the covariance matrix

$$\mathbf{C}_u = \sum_{i=1}^{M} (\mathbf{u}_i - \mathbf{u}_0)(\mathbf{u}_i - \mathbf{u}_0)^T \tag{7}$$

The size of the truncated basis m is chosen in consideration of the projection error ϵ induced by the POD procedure

$$\epsilon = 1 - \frac{\sum_{i=1}^{m} \lambda_i}{\sum_{j=1}^{M} \lambda_j} < \delta, \tag{8}$$

where δ is a prescribed tolerance.

Diffuse Approximation of the projection coefficients

The surrogate model of the projection coefficients $\alpha_i, i = 1, \ldots, m$, with respect to average stain $\langle \boldsymbol{\varepsilon} \rangle$ in Eq. (15) is constructed using the method of Diffuse Approximation

$$\widetilde{\alpha}(\langle \boldsymbol{\varepsilon} \rangle) = \mathbf{p}^T \mathbf{a}, \tag{9}$$

where $\mathbf{p} = [p_1, p_2, \ldots]^T$ is the polynomial basis vector. In 2D case, the polynomial basis vector expressed in terms of the average strain in 2D case is

$$\mathbf{p} = [1, \langle \boldsymbol{\varepsilon} \rangle_{11}, \langle \boldsymbol{\varepsilon} \rangle_{22}, \langle \boldsymbol{\varepsilon} \rangle_{12}, \ldots]^T, \tag{10}$$

Bi-level reduced Surrogate
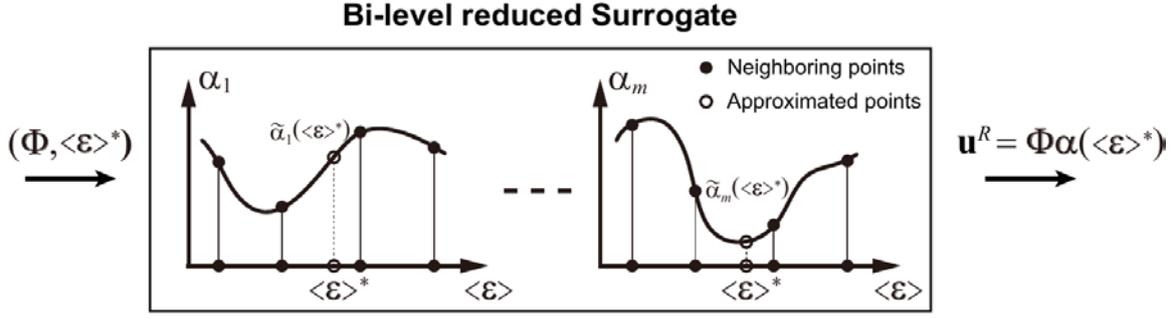
Figure 3. Illustration of the approximation procedure of the surrogate model.

The vector of coefficients $\mathbf{a} = [a_1, a_2, \ldots]^T$ are the minimizers of functional defined by

$$J(\mathbf{a}) = \frac{1}{2}\sum_{k=1}^{M} w_k (\mathbf{p}^T \mathbf{a} - \alpha(\langle \varepsilon \rangle_k))^2, \tag{11}$$

in which w_k are the weights of Euclidean distance defined following [Breitkopf et al. (2004)].

Bi-level reduced model

An illustrative flowchart of the approximation procedure is given in Fig. 3. With a given admissible value of average micro strain $\langle \varepsilon \rangle^*$, the corresponding approximated POD projection coefficients from α_1 to α_m are locally interpolated using Diffuse Approximation. Thereafter, we have the reduced order solution of the displacement filed

$$\mathbf{u}^R = \mathbf{u}_0 + \mathbf{\Phi}\widetilde{\alpha}(\langle \varepsilon \rangle^*), \tag{14}$$

where $\mathbf{\Phi} = \{\phi_1, \ldots, \phi_m\}$ is the reduced basis obtained through POD of RVE displacement fields.

The surrogate model is applied to replace full FEA in microscopic analysis. Computations during the first time step of the first optimization iteration are performed using full FEA to initialize the surrogate model. The surrogate model is then used to replace full FEA in solving the micro problem in the following computations when there are enough neighboring points to perform the approximation. When there is no enough points within the local influence zone, the micro problem is solved using full FEA and the results are used to update the POD basis $\mathbf{\Phi}$ and enrich the surrogate.

Numerical example

The benchmark cantilever problem is considered with anisotropic material defined at microscopic scale. As illustrated in Fig. 4, the macroscopic structure is discretized into 32×20 four-node plane strain elements where each element has four Gauss integration points. Each Gauss point in the macroscopic structure corresponds to a considered RVE in the microscopic scale. The material property of the solid phase in the RVE is assumed to be isotropic with a nonlinear elastic constitutive behavior as shown in Fig. 4. Conventional unreduced FE^2 approach requires $32 \times 20 \times 4$ independent RVE analysis in the microscopic scale for one time evaluation at the macroscopic scale. For the sake of simplicity, the initial elastic stiffness matrix have been kept during the Newton-Raphson iterative resolution procedure. In order to perform sensitivity analysis, tangent stiffness matrix is evaluated using the perturbation method at the converged moment of each design iteration.
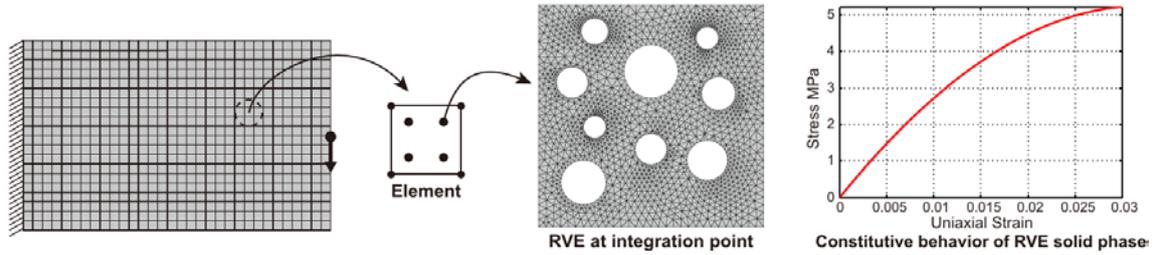
Figure 4. Illustration of the test example.

Mode 1, λ_1/λ_1=1.00   Mode 2, λ_2/λ_1=2.05.10^{-1} Mode 3, λ_3/λ_1=1.93.10^{-1} Mode 4, λ_4/λ_1=1.25.10^{-4}

Mode 5, λ_5/λ_1=3.08.10^{-5} Mode 6, λ_6/λ_1=6.79.10^{-6} Mode 7, λ_7/λ_1=1.32.10^{-6}
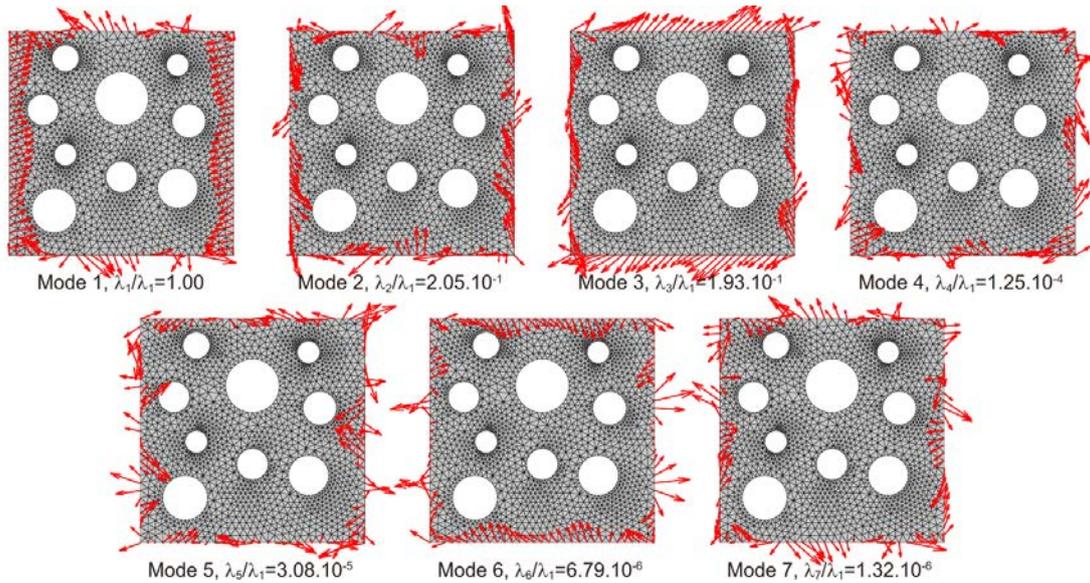
Figure 5. Resultant tractions of the first 7 POD modes after the optimization.

The external loading force is set to 1.5 N and the considered volume ratio constraint is set to 32%. The tolerance error in Eq. (8) is set as in the previous case of $\delta = 10^{-6}$. The extracted POD modes vary adaptively during the optimization procedure and the size of the reduced basis is 6 after the first iterations and then increase to 7 during the following iterations until the end. The resultant tractions of the first 7 of the final POD modes are shown in Fig. 5 together with their associated normalized eigenvalues.

The structural topological evolution in the macroscopic scale is given in Fig. 6. The convergence histories of the strain energy and the volume ratio are demonstrated in Figs. 7(a) and (b), respectively. During the loading phase of the first optimization iteration, the periodic homogenizations of the RVE in the microscopic scale are performed using full FEA. Since the second optimization iteration, both FEA and the surrogate model are used for the microscopic analysis. Fig. 7(c) gives the percentage of FEA usage in each optimization iteration. It can be seen that less than 4% microscopic analysis require full FEA except a jump from 2% in iteration 20 to 17% in iteration 21. It can be seen that a branch of the structure splits in iteration 21. Such a severe topological variation results in a large variation of the structural physical response and hence the surrogate built according to the previous calculations is no longer accurate enough. Therefore, an increased number of full FEA is required to recompute the set of the reduced basis. The surrogate model is updated thereafter and the usage ratio of FEA drops back below 4% and decreases to 0% in the following iterations as the structural topology converges, meaning that all computations are performed with the surrogate.
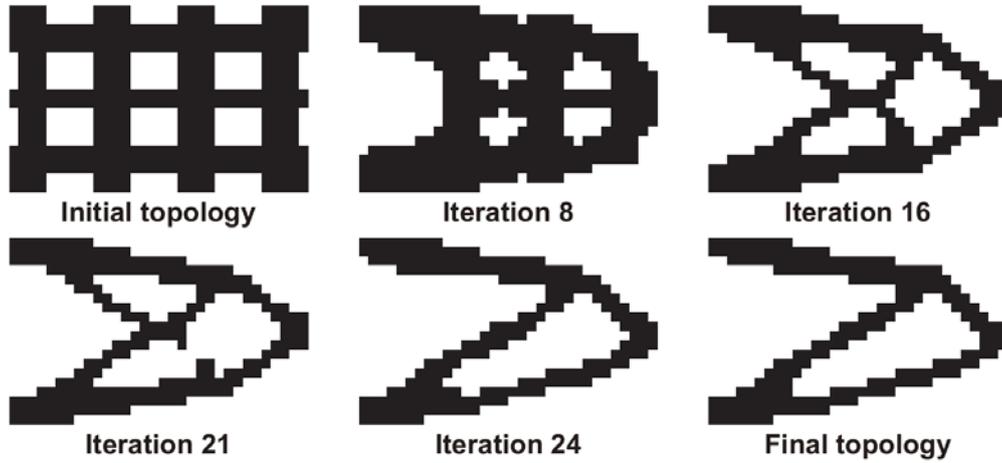
Figure 6. Structural topology variations during the optimization process.
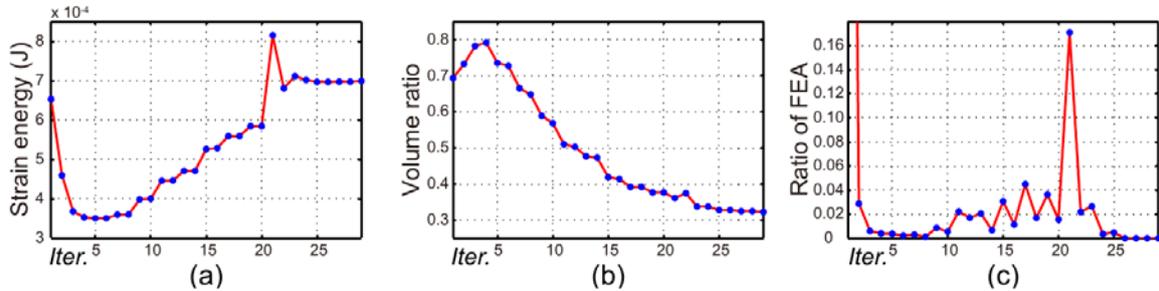
Figure 7. Optimization history: (a) convergence history of the strain energy, (b) convergence history of the volume ratio, (c) the ratio of FEA usage in each iteration.

The same optimization design has also been performed without using the surrogate. The unreduced FE^2 approach gives an exactly the same optimization design result as the reduced model where the relative errors of the objective are less than 10^{-5} . Generally speaking, it requires around two hours of computing for each optimization iteration on a HP Z420 Workstation when using the unreduced sequential FE^2. In contrast, the reduced FE^2 approach requires only ten minutes of computing on average for each design iteration apart from the first design iteration. More saving in computation can be expected using the reduced approach when larger scale problems are considered.

Fig. 8 depicts the equivalent strain distributions in the microscopic scale at selected points where the nodal displacements are scaled 20 times for the purpose of illustration. One may note that the existence of the holes in the RVE concentrates much higher strains and hence stresses in the microscopic scale than the homogenized macroscopic values. The micro strain distributions clearly manifest the difference of the loading status in different structural branches. The micro strain distributions at points b and c are quite similar because they are located in the same branch of the structure. The higher stress concentration may lead to the initial material failure or crack at the micro scale which cannot be detected when using the conventional one scale fracture analysis.

Conclusions

In this work, we have proposed a reduced multiscale model for macroscopic structural design considering microscopic material nonlinear microstructures. Several established techniques have been applied: the structural design is realized using a discrete level-set topology optimization model, the multiscale analysis is performed using the FE^2 approach, and the surrogate model is constructed

using POD and Diffuse Approximation. The surrogate model is constructed in an on-line manner: initially built during the first optimization design iteration is then updated in the following design iterations. It has been observed that the surrogate model can significantly reduce the computational cost, particularly when multiple loops involving similar computations are required. Further improvement of the proposed model could be the employment of the advanced models of any of the applied techniques, such as considering nucleation in level-set topology optimization in order to avoid an artificially defined initial topology, considering the size effect in multiscale analysis, and other possible strategies to perform model reduction either in an intrusive manner or non-intrusive manner using different approaches to construct the surrogate.
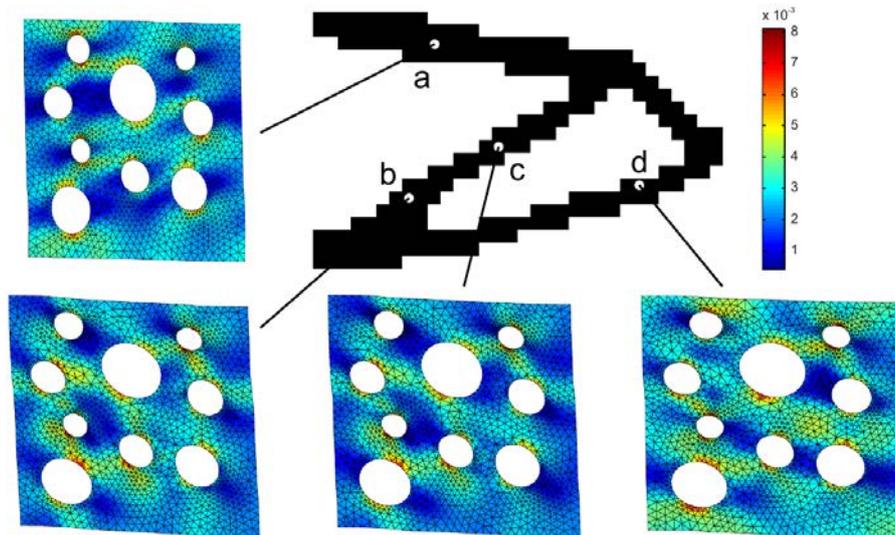
Figure 8. Equivalent strain distribution at selected points.

Acknowledgements

References

Belytschko T., Xiao S., Parimi C. (2003), Topology optimization with implicit functions and regularization, *International Journal for Numerical Methods in Engineering* **57** 1177–1196.

Challis V. J. (2010), A discrete level-set topology optimization code written in matlab, *Structural and Multidisciplinary Optimization* **41** 453–464.

Feyel F. and Chaboche J. (2000) FE^2 multiscale approach for modelling the elastoviscoplastic behaviour of long fibre sic/ti composite materials, *Computer Methods in Applied Mechanics and Engineering* **183** 309–330.

Kouznetsova V., Brekelmans W. A. M., Baaijens F. P. T. (2001), An approach to micro-macro modeling of heterogeneous materials, *Computational Mechanics* **27** 37–48.

Lancaster P., Salkauskas K. (1981), Surfaces generated by moving least squares methods, *Mathematics of Computation* **87** 141–158.

Nayroles B., Touzot G., Villon P. (1992), Generalizing the finite element method: Diffuse approximation and diffuse elements, *Computational Mechanics* **10** 307–318.

Breitkopf P., Rassineux A., Savignat J. M., Villon P. (2004), Integration constraint in diffuse element method, *Computer Methods in Applied Mechanics and Engineering* **193** 1203–1220.

Yvonnet J., He Q. C. (2007), The reduced model multiscale method (r3m) for the non-linear homogenization of hyperelastic media at finite strains, *Journal of Computational Physics* **223** 341–368.

Convergency of viscoelastic constraints to Nonholonomic Idealization

*Alexander Fidlin[1] and Jens Deppler[1] and Björn Braun[1]
[1]Institute for Mechanical Engineering, Karlsruhe Institute of Technology
*Corresponding author: alexander.fidlin@kit.edu

Abstract

Rolling contacts are usual in various technical systems and yield usually non-holonomic constraints. A new regularization method motivated by physical considerations is investigated in the present paper. The convergence of the spring-damper regularization for the so called principal damping, which is motivated by the critical damping in the linear case, is proven. The solutions of the DAEs and the corresponding ODEs converge if a certain condition on the regularization parameters is fulfilled. A rolling disc on the flat plane and a skate on an inclined plane are analyzed as numerical examples. It is demonstrated firstly that the optimal choice of the regularization parameters corresponds to the principle damping and secondly that the sufficient convergence condition obtained in the proof is valid for the numeric simulations.

1. Introduction

In most cases the constraint equations on velocity level enforcing a rolling motion cannot be integrated, yielding nonholonomic constraint equations. Usually the nonholonomic constraints can be incorporated into the equations of motion by the method of Lagrange multipliers. This formulation leads to index-2 differential algebraic problems. In the present paper we investigate a new viscoelastic idealization of nonholonomic constraints, that is motivated by physical considerations. Pure rolling is equal to a sticking state, with a kinematically repositioned contact point. Usually sticking is modeled by introducing an elasticity in the contact as demonstrated by [Vielsack, 1996]. Here the constraint is enforced by the elastic and dissipative terms, that help to avoid numerical oscillations in the contact. In an earlier work [Stamm, 2011] applied this kind of viscoelastic formulation to a tangential contact law, extending the classical laws of friction, like the Coulomb model, to distributed contacts, in order to circumvent the problem of indeterminacy in the sticking state. However a description of a contact law by means of viscoelastic forces is sensible only if it approximates the idealized rigid formulation in case of infinitely stiff chosen viscoelastic parameters. Thus the objective of this work is to show the convergence of the viscoelastic description to the idealized nonholonomic rigid description in a mathematical sense.

2. Statement of problem

Consider the general multibody system with m nonholonomic constraint equations, as given in definition 1.

Definition 1 (Differential algebraic initial value problem)**.** *Let $I = [t_0, t_e]$ be a closed interval. Then the equations of motion can be described by the following differential algebraic initial value problem*

$$M(q)\ddot{q} = F(q, \dot{q}, t) - G^{\mathsf{T}}(q)\Lambda, \tag{1}$$
$$0 = G(q)\dot{q} \tag{2}$$

with the consistent initial conditions $q(t_0) = q_0$, $\dot{q}(t_0) = \dot{q}_0$. Furthermore holds $M(q) \in \mathbb{R}^{n \times n}$ is symmetric and positive definite. The functions $G(q) \in \mathbb{R}^{m \times n}$ and $F(q, \dot{q}, t) \in \mathbb{R}^n$ are sufficiently smooth, the matrix $G(q)$ is assumed to have full rank m.

Usually deformations occur in a contact area due to local deformations of asperities and the elasticity of the bodies itself. A sensible physical description of a contact should take these effects into account. Thus the constraint forces, that enforce the constraint equation, are replaced by applied forces in form of a viscoelastic force element, which leads to the viscoelastic description of the given multibody system as stated in definition 2.

Definition 2 (Viscoelastic description)**.** *Let $I = [t_0, t_e]$ be a closed interval. Then for $t \in I$ and for a fixed $\varepsilon_f \in (0, \varepsilon_0]$ the equation of motion of the viscoelastic description is given by*

$$M(q)\ddot{q} = F(q, \dot{q}, t) - G^{\mathsf{T}}(q)\Lambda, \tag{3}$$
$$\dot{z} = G(q)\dot{q} \tag{4}$$

along with the initial conditions $q(t_0) = q_0$, $\dot{q}(t_0) = \dot{q}_0$ and $z(t_0) = z_0$, where the Lagrange multiplier Λ is replaced by

$$\Lambda = \frac{c}{\varepsilon_f} z + \frac{d}{\varepsilon_f^\kappa} \dot{z}. \tag{5}$$

The functions $G(q) \in \mathbb{R}^{m \times n}$ and $F(q, \dot{q}, t) \in \mathbb{R}^n$ are sufficiently smooth, the matrix $G(q)$ is assumed to have full rank m.

The parameter κ is chosen as $\frac{1}{2}$, inspired by the critical damping in the linear case. The given forms of the underlying systems are not suitable, in order to obtain an estimate of the distance between the corresponding solutions. However by applying appropriate transformations, they can be transformed to a standard singular perturbation form. Transforming the systems given in definition 1 and definition 2 to autonomous systems and introducing the new variable θ according to

$$\sqrt{\varepsilon}\dot{\Lambda} = \theta + h(y, \Lambda) + S(\Lambda), \tag{6}$$

where $S(\Lambda) = \Lambda$ and $h(y, \Lambda) = c(\frac{d}{c})^{\frac{1}{1-\kappa}}(\dot{G}y_2 + GM^{-1}F - GM^{-1}G^\mathsf{T}\Lambda)$. The equivalent systems in standard singular perturbation form can be obtained.

$$\begin{aligned}
\dot{y} &= f(y, \Lambda), \\
\sqrt{\varepsilon}\dot{\Lambda} &= \theta + h(y, \Lambda) + S(\Lambda) &&:= g_1(y, \theta, \Lambda), \\
\sqrt{\varepsilon}\dot{\theta} &= h(y, \Lambda) - \tfrac{\mathrm{d}S}{\mathrm{d}\Lambda}(\theta + h(y, \Lambda) + S(\Lambda)) &&:= g_2(y, \theta, \Lambda).
\end{aligned}$$

Setting $\varepsilon = 0$ leads to the reduced problem in form of differential algebraic equations

$$\begin{aligned}
\dot{y} &= f(y, \Lambda), \\
0 &= \theta + h(y, \Lambda) + S(\Lambda) &&:= g_1(y, \theta, \Lambda), \\
0 &= h(y, \Lambda) - \tfrac{\mathrm{d}S}{\mathrm{d}\Lambda}(\theta + h(y, \Lambda) + S(\Lambda)) &&:= g_2(y, \theta, \Lambda).
\end{aligned}$$

with the column matrices $\lambda = [\Lambda, \theta]$, $g = [g_1, g_2]$ and $f(y, \lambda) = [y_2, M^{-1}(F - G^\mathsf{T})\Lambda, 1]$ the problems can be written conveniently. The differential algebraic equation can be represented in the following form:

Definition 3 (Differential algebraic equation)**.**

$$\begin{aligned}
\dot{y} &= f(y, \lambda), \\
0 &= g(y, \lambda), \\
y(0) &= y_0^0.
\end{aligned} \tag{7}$$

The viscoelastic approximation reads as:

Definition 4 (Viscoelastic description in singular perturbation standard form)**.**

$$\begin{aligned}
\dot{y} &= f(y, \lambda), \\
\sqrt{\varepsilon}\dot{\lambda} &= g(y, \lambda), \\
y(0) &= y_0^0 + \sqrt{\varepsilon}y_1^0 + \sqrt{\varepsilon}^2 y_2^0 + \dots, \quad \lambda(0) = \lambda_0^0 + \sqrt{\varepsilon}\lambda_1^0 + \sqrt{\varepsilon}^2\lambda_2^0 + \dots.
\end{aligned} \tag{8}$$

3. Proof of convergency

In the underlying form, standard singular perturbation approaches can be used in order to obtain an estimate of the distance of the solution of the problems given in definition 3 and definition 4.

In order to construct a solution of the initial value problem eq. (8) in form of an infinite asymptotic power series expansion the following theorem by Hairer and Wanner [Hairer and Wanner, 2010] can be applied.

Theorem 1. *Let f and g be sufficiently smooth functions. Consider the initial value problem given in eq. (8)*

$$\begin{aligned}
\dot{y} &= f(y, \lambda), \\
\sqrt{\varepsilon}\dot{\lambda} &= g(y, \lambda), \\
y(0) &= y_0^0 + \sqrt{\varepsilon}y_1^0 + \sqrt{\varepsilon}^2 y_2^0 + \dots, \qquad \lambda(0) = \lambda_0^0 + \sqrt{\varepsilon}\lambda_1^0 + \sqrt{\varepsilon}^2\lambda_2^0 + \dots.
\end{aligned}$$

Introducing the time scale $\tau = \frac{t}{\sqrt{\varepsilon}}$ enables the construction of the solutions in form of an infinite asymptotic series expansion according to

$$y(t) = \sum_{j=0}^{\infty}\sqrt{\varepsilon}^j y_j(t) + \sqrt{\varepsilon}\sum_{j=0}^{\infty}\sqrt{\varepsilon}^j\eta_j(\tau), \quad \lambda(t) = \sum_{j=0}^{\infty}\sqrt{\varepsilon}^j\lambda_j(t) + \sum_{j=0}^{\infty}\sqrt{\varepsilon}^j\zeta_j(\tau). \tag{9}$$

The functions $\eta_j(\tau)$ and $\zeta_j(\tau)$ satisfy the conditions

$$\|\eta_j(\tau)\| \le K_j\mathrm{e}^{-\kappa_j\tau}, \quad \|\zeta_j(\tau)\| \le C_j\mathrm{e}^{-\kappa_j\tau}.$$

59

The proof of this theorem is given in [Hairer and Wanner, 2010]. Instead of the infinite power series expansion, the truncated power series expansion will be used instead. Special interest is devoted to the series truncated at $N = 0$ since the resulting zeroth approximation corresponds to the differential algebraic equation. Thus the target is to find an estimation of the error made when using the truncated series expansion instead of the full series expansion. This question is answered by the following theorem from Hairer and Wanner [Hairer and Wanner, 2010].

Theorem 2. *Let f and g be sufficiently smooth functions. Consider the viscoelastic formulation in form of initial value problem* (8). *Suppose that the logarithmic norm $\mu(g_\lambda) < -1$ holds in an ε independent neighborhood of the solution $y_0(t)$, $\lambda_0(t)$ of the differential algebraic equation* (7) *with the initial condition $y_0(0) = y_0^0$, satisfying the constraint equation, on the interval $0 < t < T$. If the initial values y_0^0 and λ_0^0 lie in this neighborhood, then the initial value problem* (8) *has a unique solution for ε sufficiently small and for $0 < t < T$, which is of the form*

$$y(t) = y_{tr}(t) + \mathcal{O}(\sqrt{\varepsilon}^{N+1}) = \sum_{j=0}^{N} \sqrt{\varepsilon}^j y_j(t) + \sqrt{\varepsilon} \sum_{j=0}^{N} \sqrt{\varepsilon}^j \eta_j(\tfrac{t}{\sqrt{\varepsilon}}) + \mathcal{O}(\sqrt{\varepsilon}^{N+1}), \tag{10}$$

$$\lambda(t) = \lambda_{tr}(t) + \mathcal{O}(\sqrt{\varepsilon}^{N+1}) = \sum_{j=0}^{N} \sqrt{\varepsilon}^j \lambda_j(t) + \sum_{j=0}^{N} \sqrt{\varepsilon}^j \zeta_j(\tfrac{t}{\sqrt{\varepsilon}}) + \mathcal{O}(\sqrt{\varepsilon}^{N+1}). \tag{11}$$

The coefficient functions $\eta_j(\tau)$ and $\zeta_j(\tau)$ satisfy $\|\eta_j(\tau)\| \leq K_j e^{-\kappa_j \tau}$ and $\|\zeta_j(\tau)\| \leq C_j e^{-\kappa_j \tau}$. The error between the solution of the differential algebraic equation (7), *which corresponds to the truncated series at $N = 0$, and the viscoelastic formulation* (8) *can be estimated above according to*

$$\|y - y_0\| \leq M_1 \sqrt{\varepsilon}, \quad \|\lambda - \lambda_0\| \leq M_2 \sqrt{\varepsilon}.$$

The proof of this theorem is given in Hairer and Wanner [Hairer and Wanner, 2010]. Thus the solution of the viscoelastic contact formulation is in an $\mathcal{O}(\sqrt{\varepsilon})$ vincinity of the solution of the differential algebraic equation. The request that the logarithmic norm $\mu(g_\lambda) < -1$ leads to the following condition on the eigenvalues of the matrix $-GM^{-1}G^{\mathsf{T}}$:

$$\lambda_{max}^{-GM^{-1}G^{\mathsf{T}}} \leq -\tfrac{2c}{d^2}, \tag{12}$$

where $\lambda_{max}^{-GM^{-1}G^{\mathsf{T}}}$ denotes the maximum eigenvalue of the matrix $-GM^{-1}G^{\mathsf{T}}$.

4. Numerical experiments

In order to confirm the theoretical results, numerical experiments were carried out. Therefore a classical mechanical system of a skate sliding down an inclined plane under the influence of gravity is considered. The model is shown in fig. 1. Mathematically the constraint equation is given by the demand, that the velocity of the contact point is always parallel
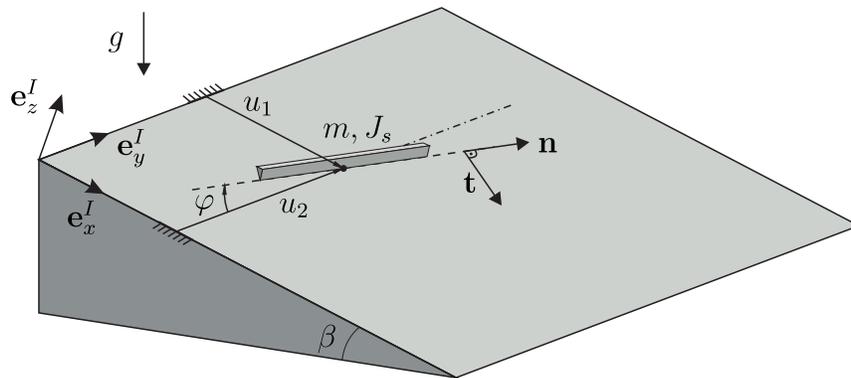
Figure 1: disk rolling on a flat support

to the skid, which can be expressed in the following fashion

$$\mathbf{v} \cdot \mathbf{t} = 0,$$

(a) Condition $\lambda_{max}^{-GM^{-1}G^{\mathsf{T}}} \leq -\frac{2c}{d^2}$ fulfilled.

(b) Condition $\lambda_{max}^{-GM^{-1}G^{\mathsf{T}}} \leq -\frac{2c}{d^2}$ hurt.
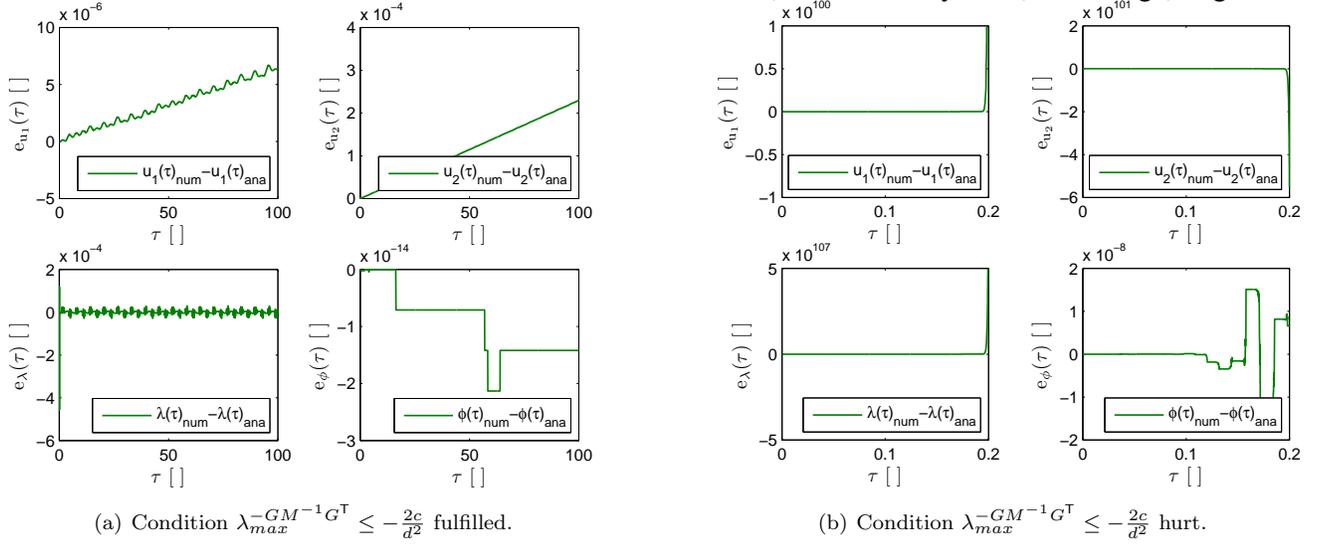
Figure 2: Deviations of the solutions of differential algebraic equation and viscoelastic formulation in case (12) is fulfilled (left) and not fulfilled (right), where the deviation grows exponentially fast.

where $\mathbf{v}$ denotes the velocity of the contact point and $\mathbf{t}$ the vector perpendicular to the skid. This finally results in the scalar constraint equation

$$0 = \underbrace{\begin{bmatrix} -\sin\varphi & \cos\varphi & 0 \end{bmatrix}}_{G(q)} \underbrace{\begin{bmatrix} \dot{u}_1 \\ \dot{u}_2 \\ \dot{\varphi} \end{bmatrix}}_{\dot{q}}.$$

5. Discussion and conclusion

Convergency of the viscoelastic description of contact forces is proven for nonholonomic constraints in general form. The proof is performed for the principal damping exponent. The solutions of the DAE and the corresponding ODE converge if the condition

$$\mu(GM^{-1}G^{\mathsf{T}}) > \frac{2c}{d^2}$$

is fulfilled. Numerical experiments were made to verfiy the statement of the proof. They confirm the optimum performance for this choice of the viscoelastic parameters. In the future the described approach will enable a consistent modelling of sticking, sliding and rolling contacts in multibody dynamics.

References

[Hairer and Wanner, 2010] Hairer, E. and Wanner, G. (2010). *Solving ordinary differential equations II.* Springer.

[Stamm, 2011] Stamm, W. (2011). *Modellierung und Simulation von Mehrkörpersystemen mit flächigen Reibkontakten.* PhD thesis, Karlsruher Institut für Technologie.

[Vielsack, 1996] Vielsack, P. (1996). Regularisierung des Haftzustandes bei Coulombscher Reibung. *Z. angew. Math. Mech.*, 76:439–446.

Numerical modeling of curvilinear corrugated-core sandwich structures subjected to low velocity impact loading

*T. Boonkong[1,2], Y.O. Shen [3], Z.W. Guan[1] and W.J. Cantwell[4]

[1] School of Engineering, University of Liverpool, L69 3GH, UK.
[2] Department of Engineering, The Royal Thai Naval Dockyard, Royal Thai Navy, Bangkok, Thailand
[3] School of Aerospace Engineering and Applied Mechanics, Tongji University, China
[4] Department of Aerospace Engineering, Khalifa University of Science, Technology and Research (KUSTAR), P.O. Box 127788, Abu Dhabi, United Arab Emirates

*Corresponding author: t.boonkong@liv.ac.uk

In this paper, simulations of low velocity impact characteristics of curvilinear corrugated-core sandwich structures were presented, which were validated against the corresponding experimental data. Two different configurations of lightweight aluminium sandwich panels from Metawell® Company in Germany were tested using drop-weight impact tower with spherical indenter to evaluate their energy-absorbing characteristics and to identify the associated failure mechanisms under vary of impact loading conditions.

Here, two panel configurations were studied based on the finite element analysis by using commercial finite element code Abaqus/Explicit developing numerical models to cover the most representative cases. A good degree of correlation was obtained, which indicates the finite element models developed are capable of predicting the dynamic behaviour of the curvilinear corrugated-core sandwich structure panels subjected to low velocity projectile impact.

Keywords: Curvilinear corrugated-core sandwich structures, low velocity impact, finite element, perforation failure.

Introduction

Sandwich structures are considered as optimal designs for a wide range of applications such as insulated structures, marine construction, transportation and aerospace vehicles. A composite sandwich panel is usually made from a lightweight foam, honeycomb or corrugated core sandwiched between two composite face sheets. Such a combination offers exceptional specific strength-to-weight ratio or stiffness-to-weight ratio, buoyancy, dimensional stability, and thermal and acoustical insulation characteristics. The curvilinear corrugated-core sandwich structure is one of outstanding sandwich structures offering superior mechanical properties. Many researches have been study on various types of sandwich structures [Biancolini (2005) , Nyman and Gustafsson (2000) , Rejab and Cantwell (2013) , Herrmann, Zahlen (2005) , Kazemahvazi and Zenkert (2009) , Xiong, Ma (2011) , Lin, Liu (2007) , Zenkert (1995) , Zhang Y (2011) , Yokozeki, Takeda (2006)]. However, it was found that few of published worked involved in curvilinear corrugated-core sandwich structures in spite of a versatile applications.

In this paper, the curvilinear corrugated-core sandwich structures from Metawell® company, which is a patented lightweight construction aluminium panel made by bonding two cover sheets to the core material, consisting of wave formed sheet metal,

using a hot melt adhesive in a continuous, process were used and tested in order to study the influence of low velocity impact attached by the spherical indenter response to the rigid panels.

Experimental Work

The curvilinear corrugated-core sandwich structures in this study were based on EN AW-1582 H48 aluminium alloy sheets from fabricated by bonding two cover sheets into core material, which consists of wave formed sheet metal, using a hot melt adhesive in a continuous process. There were two panel configurations, which different fact sheet thicknesses and core sizes were tested. Fig.1 shows a design and dimension of both panels.
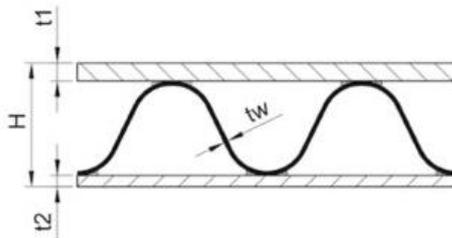
t_1 - thickness of top cover sheet

t_w - thickness of corrugation

t_2 - thickness of bottom cover sheet

H - panel height in mm.

Fig.1 shows a design and dimension of both panels.

Table 1. Panel dimensions

Type	t1 (mm)	t_w (mm)	t_2 (mm)	H (mm)	weight (kg/m^2)	Descriptions
Alu hl 05-02-05 hl/H6	0.5	0.2	0.5	6.0	3.8	lightweight panel (primer coated)
Alu cc 08-03-05 hl/H10	0.8	0.3	0.5	10.0	5.2	White coating on one side

Low velocity impact tests on the panels started from 1.93 m/s and increased gradually until 5.4m/s were conducted by using an Instron CEAST 9350 drop tower machine. A cylindrical impactor of 5.32 kg with 25.4 mm diameter spherical end was used. The test specimens had the dimension 155 mm. x 155 mm. The specimens were clamped by cylindrical ring with inside and outside diameter of 76 and 100 mm. respectively. The 200 N. of clamp force between both bottom and top rings was applied. Details about the test configuration are shown in Figure 2.

In order to get the materials properties for the input parameters used in finite element modelling, the top and bottom face sheets were tested by using Instron 4505 to conduct the uniaxial tensile test. The result from tensile test is shown as the graph in Fig. 3.
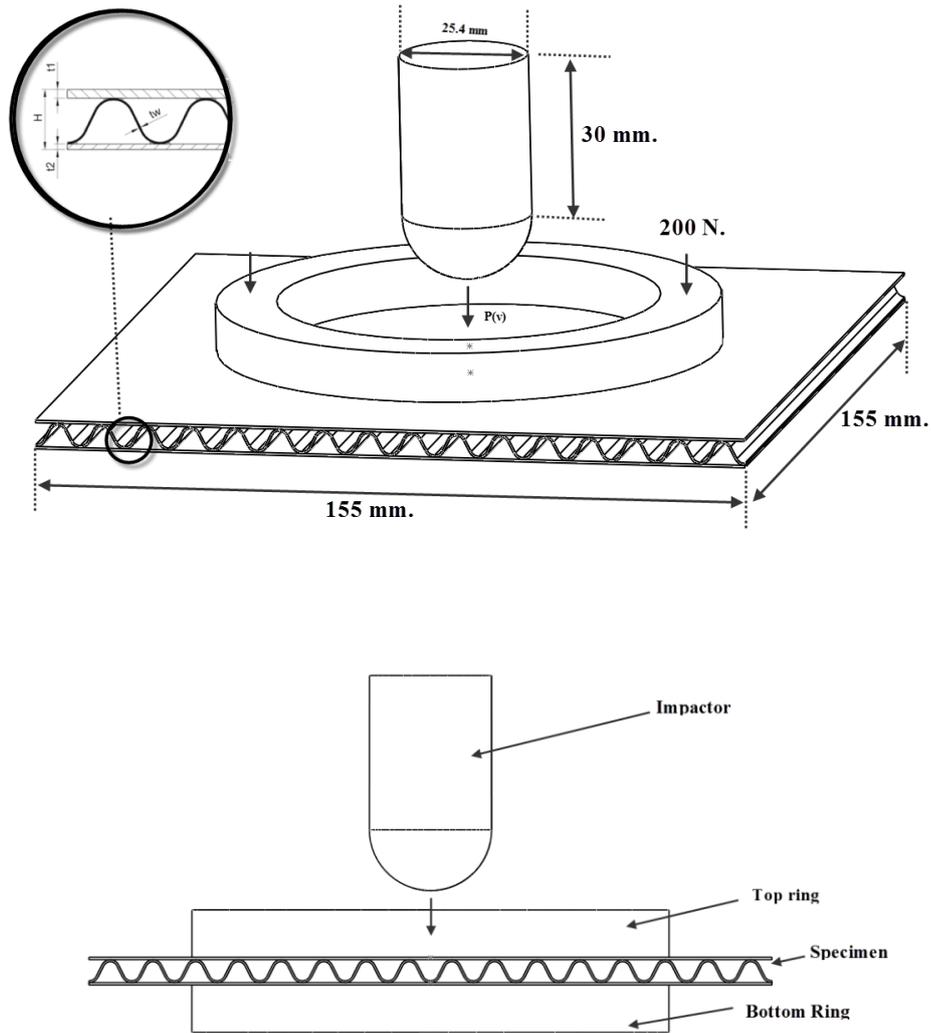
Fig.2 (a) Schematic of drop-weight apparatus, using spherical impactor (b) side view

Finite element modelling

ABAQUS/Explicit [Abaqus6.12-3 (2012)]was used to develop numerical simulations of the curvilinear corrugated-core sandwich structures under low velocity impact. The aluminium alloy was modelled as an elasto-plastic material with rate-dependent behaviour. For a rate-dependent material, the relationship follows the uniaxial flow rate definition as:

$$\dot{\bar{e}}^{pl} = h(q, \bar{e}^{pl}, \theta) \tag{1}$$
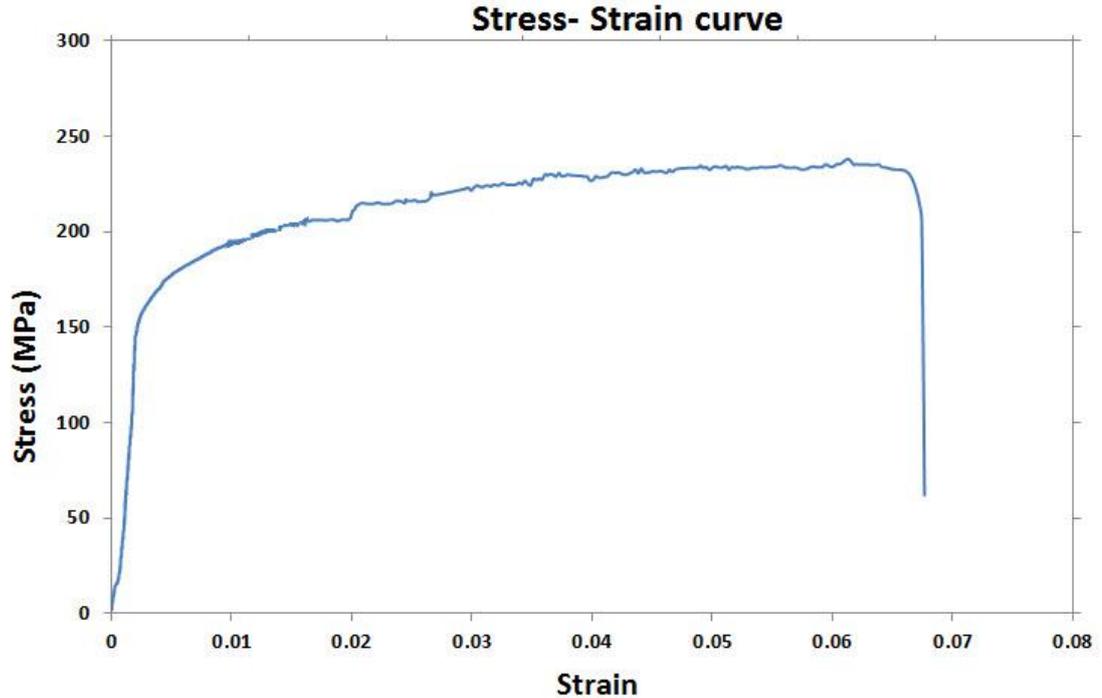
Fig.3 The stress- strain curve of EN AW-1582 H48 from tensile test

Where h is a known strain hardening function, q is the von-Mises equivalent stress, e^{-pl} is the equivalent plastic strain, and θ is the temperature. The isotropic hardening data for the EN AW-1582 H48 aluminium alloy are given in Table 2. The density of the aluminium was taken as ρ = 2690 kg/m³. The material properties of EN AW-1582 H48 can be found in table 3.

Table 2. Isotropic hardening data for the EN AW-1582 H48 aluminium alloy

Yield stress (MPa)	153	160	178	203	214	224	231	234	235	232
Plastic strain	0	4E-4	0.002	0.013	0.020	0.030	0.040	0.050	0.056	0.065

The rate-dependent hardening curves can be expressed as:

$$\bar{\sigma}\left(\bar{\varepsilon}_{pl}, \dot{\bar{\varepsilon}}_{pl}\right) = \delta_y\left(\bar{\varepsilon}_{pl}\right)R\left(\dot{\bar{\varepsilon}}_{pl}\right) \tag{2}$$

Where $\bar{\varepsilon}_{pl}$ and R are the equivalent plastic strain and stress ratio (= $\bar{\sigma} / \sigma_y$) respectively.

Damage initiation criteria
Ductile damage criterion is a phenomenological model for predicting the onset of damage due to nucleation, growth, and coalescence of voids. The model assumes that the equivalent plastic strain at the onset of damage, $\bar{\varepsilon}_D^{pl}$, is a function of stress triaxiality and strain rate:

$$\bar{\varepsilon}_D^{pl}\left(\eta, \dot{\bar{\varepsilon}}_{pl}\right) \tag{3}$$

Where $\eta = -p/q$ and η is the stress triaxiality, p is the pressure stress, q is the Misses equivalent stress, and $\dot{\bar{\varepsilon}}_{pl}$ is the equivalent plastic strain rate. The criterion for damage initiation is met when the following condition is satisfied:

$$\omega_D = \int \frac{d\,\bar{\varepsilon}_{pl}}{\bar{\varepsilon}_D^{pl}(\eta, \dot{\bar{\varepsilon}}_{pl})} = 1 \tag{4}$$

Where ω_D is a state variable that increases monotonically with plastic deformation. At each increment during the analysis the incremental increase is computed as:

$$\Delta\omega_D = \int \frac{\Delta\,\bar{\varepsilon}_{pl}}{\bar{\varepsilon}_D^{pl}(\eta, \dot{\bar{\varepsilon}}_{pl})} \geq 0 \tag{5}$$

Shear failure criterion
The shear failure model is based on the value of the equivalent plastic strain at element integration points; failure is assumed to occur when the damage parameter exceeds 1. The damage parameter, ω, is defined as :

$$\omega = \frac{\bar{\varepsilon}_0^{pl} + \sum \Delta\bar{\varepsilon}^{pl}}{\bar{\varepsilon}_f^{pl}} \tag{6}$$

where $\bar{\varepsilon}_0^{pl}$ is any initial value of the equivalent plastic strain, $\sum \Delta\bar{\varepsilon}^{pl}$ is an increment of the equivalent plastic strain, is the strain at failure, and the summation is performed over all increments in the analysis. The strain at failure, $\bar{\varepsilon}_f^{pl}$, is assumed to depend on the plastic strain rate, $\dot{\bar{\varepsilon}}_{pl}$; a dimensionless pressure-deviatoric stress ratio, p/q (where p is the pressure stress and q is the Mises stress); temperature; and predefined field variables. However, in this model, the temperature parameter would be ignored as a small effect to the results.

Element removal
When the shear failure criterion is met at an integration point, all the stress components will be set to zero and that material point fails. By default, if all of the material points at any one section of an element fail, the element is removed from the mesh; it is not necessary for all material points in the element to fail. For example, in a first-order reduced-integration solid element removal of the element takes place as soon as its only integration point fails. However, in a shell element all through-the-thickness integration points must fail before the element is removed from the mesh. In the case of second-order reduced-integration beam elements, failure of all integration points through the section at either of the two element integration locations along the beam axis leads, by default, to element removal[Abaqus6.12-3 (2012)].

Geometry and Mesh design

In order to reduce time of processing, only a quarter of modelling was generated. The Aluminium corrugated core and skin parts were meshed with a uniform mesh consisting primarily of 8-node linear brick, reduced integration, hourglass control elements (C3D8R). Core and skins were completely bonded with tie constrain around the interface areas. A 4-node 3-D bilinear rigid quadrilateral (R3D4) was used to contribute support rings and spherical end projectile.
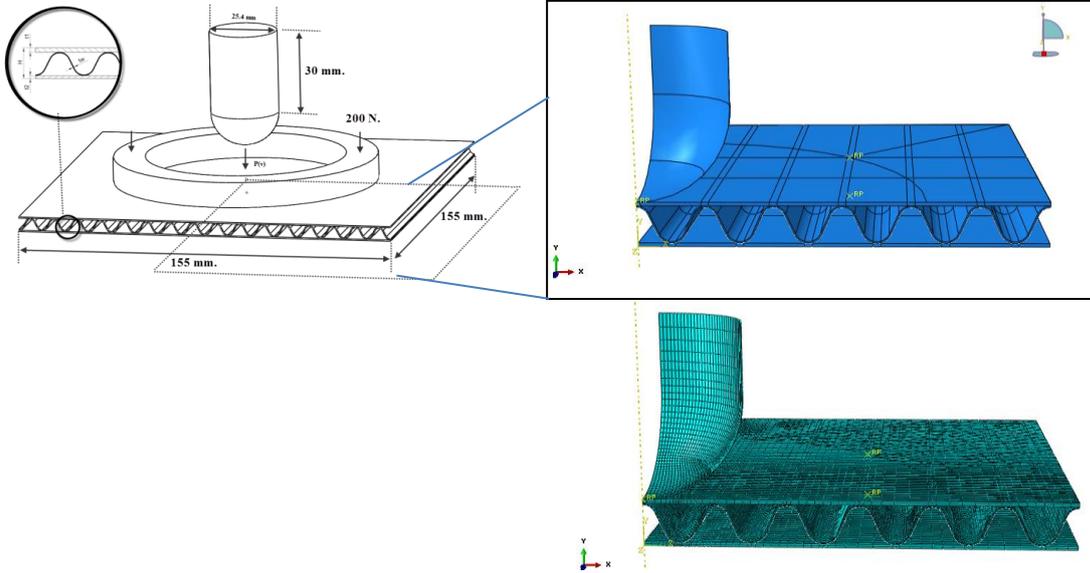
Fig.4 shows the quarter model assembly and mesh design.

Boundary conditions and loading

For the support bottom support ring, it was fixed all of degree of freedom and the -200 N. of uniform pressure was applied on the top support ring imitating as the experimental clamp condition. The projectile, which had the inertia of 5.321 kg, was allowed to translate only in y direction with the required predefined field of initial velocity.

The general contact, which had the contact domain included surface pairs by all with self-contact was applied for the whole model. The contact properties had frictionless tangential behaviour and hard contact for normal behaviour.

Table 3. Materials properties and parameters used in finite element modelling

Properties	Values
Young's modulus (Gpa.)	68
Density (kg/m^3)	2650
Strain rate	150
Fracture strain for ductile damage	0.065
Fracture strain for shear damage	0.050
Stress triaxiality	0.33
Fracture energy (kJ/m^2)	67

Results and discussions

Fig. 5 and 6 compare typical load-displacement plots for the impact energy from 10 J. up to 80 J. It could be indicated that the agreement between the experimental results
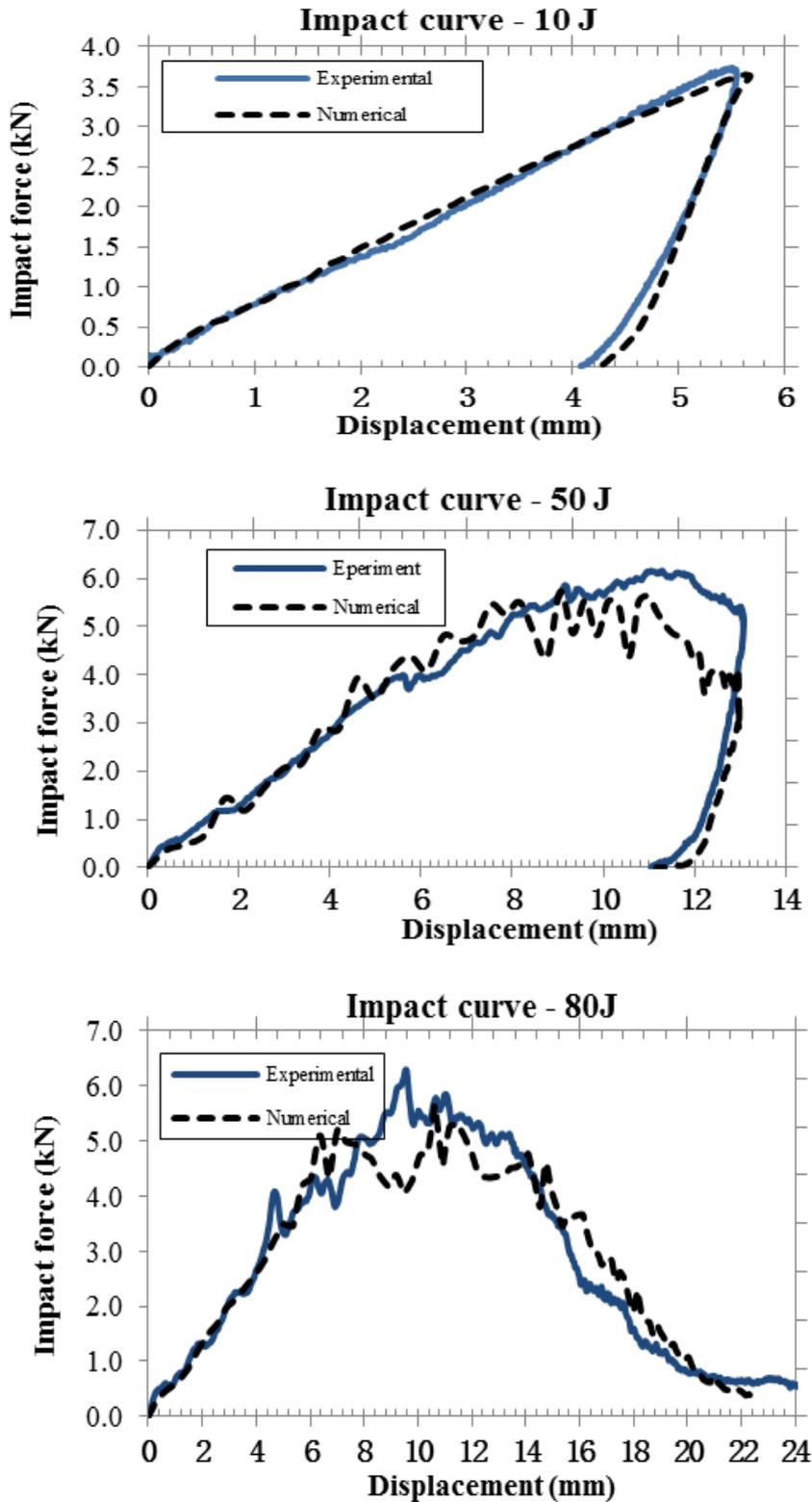
Fig. 5 Typical load-displacement plots from Alu hl 05-02-05 hl/H6 panels in ascending impact energy

and the numerical predictions is very good for both panels. For Alu hl 05-02-05 hl/H6, the prediction from numerical model slightly offered a higher impact displacement when 50J. was applied as shown in fig. 5. The results from numerical model seem be perforated slightly later than the experimental results according to the
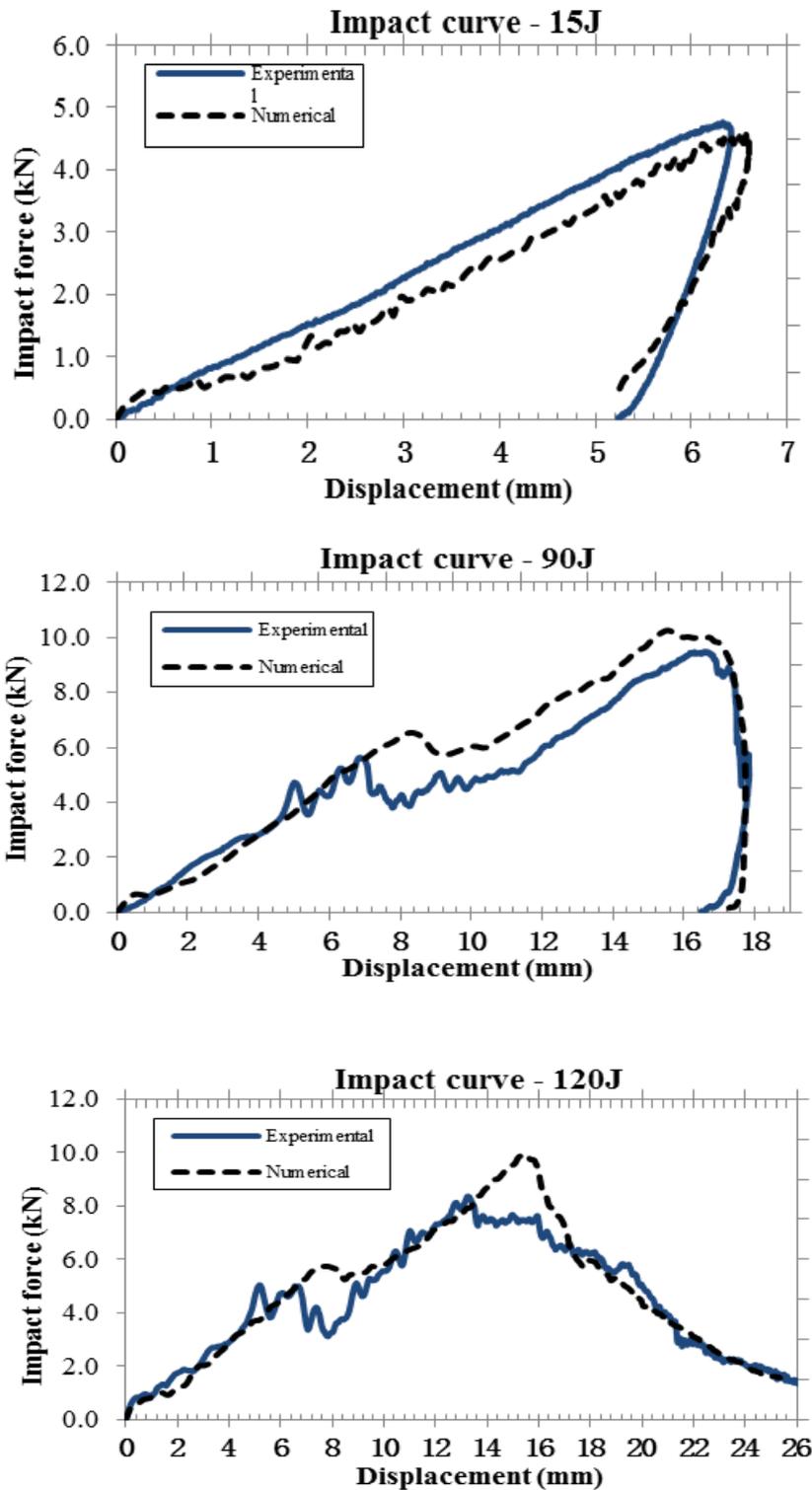
Fig. 6 Typical load-displacement plots from Alu cc 08-03-05 hl/H10 panels ascending impact energy

panel Alu cc 08-03-05 hl/H10 presented in fig. 6. Clearly, the peak load increases with the velocity. However, it was found that the panel Alu hl 05-02-05 hl/H6, which has less structures and bottom face sheet thickness, could offer a higher peak load in the range of velocity since 2.73 m/s to 3.8 m/s. It could be indicated that after 6 m/s, the bottom face sheet of Alu cc 08-03-05 hl/H10 obviously affected to the peak load as shown in fig. 7.
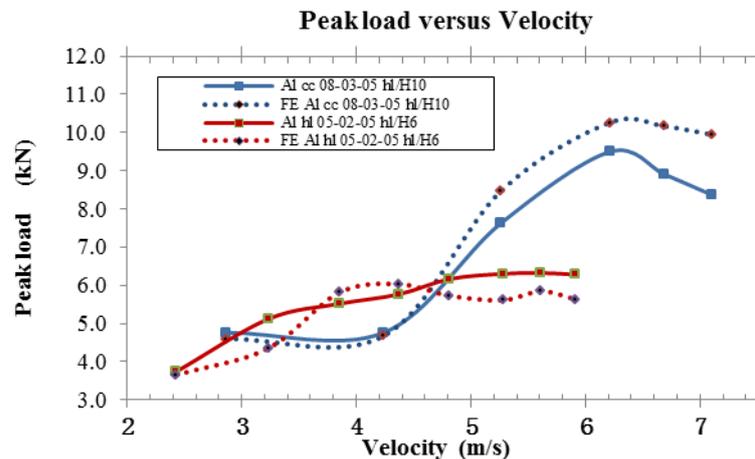
Fig. 7 compares peak load against velocity between panels Alu hl05-02-05 hl/H6 and Alu cc 08-03-05 hl/H10.

Apparently, the prediction offers correlation of peak load from Alu cc 08-03-05 hl/H10 in the initial state and it seem diverge when the velocity increased. Only in the range of 3.35 - 3.78 m/s from numerical results had slightly higher than the experimental results. It could be considered that the maximum perforation load is 9.4 kN. at 90 J. before dropping when increasing of velocity for Alu cc 08-03-05 hl/H10. Meanwhile, the trend of peak load seems to be constant while the impact velocity is increasing since 4.71 m/s.

From the finite element model results in fig. 8(c), it could gradually reveal the initial stress concentration and the propagation of failure on the panel since t = o millisecond until the panel was fully perforated at t = 6.00 milliseconds. It also could predict that the stress comes along the longitudinal corrugation direction (Z axis). The evidence revealed that it could not find the debonding failure mode between the corrugated-core and both top and bottom face sheets. Therefore, using the tie constrains between core and skins could be acceptable in the finite element model. It was found the buckling mode of failure mechanism occurred before the propagation of fracture would initiate. The initial crack did not propagate from the middle of impact, but started from the cavity inside the coalescent core then spread along z-direction as a crescent form.

The influence of projectile on the perforation resistance of the curvilinear corrugated-core sandwich structures are shown in fig. 8(a) and (b). Surprisingly, the diameter of penetration were investigated and found in double of the projectile diameter.

(a)

(b)

(c)
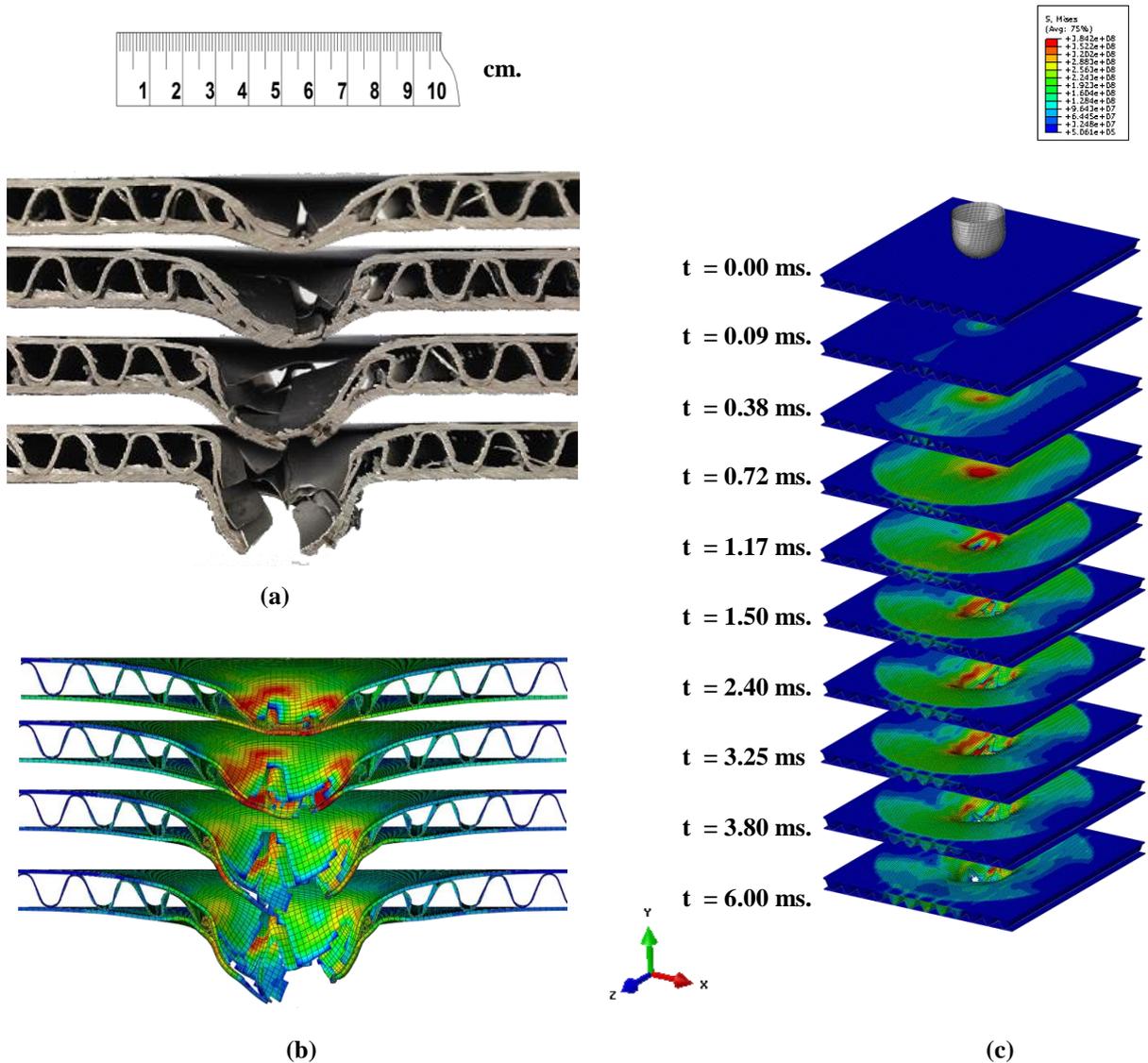
Fig. 8 (a) and (b) Compares central cross-section view of perforation between experimental and finite element modelling, using Alu hl 05-02-05 hl/H6, (c) Deformation of perforation since t = 0 millisecond until fully perforated at t = 6.00 millisecond.

Conclusions

Agreement between the experimental and predicted data is reasonably good, with the model tending to follow the experimental data. Only in some regions were observed not associated in particular the impact displacement, which seem offers slightly greater than measured data.

Increasing of the core and face sheets thickness enhances the stiffness and impact energy resistance quite in double of maximum peak load.

References

Biancolini, M.E. (2005), Evaluation of equivalent stiffness properties of corrugated board. Composite Structures, **69**(3): p. 322-28.

Nyman, U. and P.J. Gustafsson (2000), Material and structural failure criterion of corrugated board facings. Composite Structures, **50**(1): p. 79-83.

Rejab, M.R.M. and W.J. Cantwell (2013), The mechanical behaviour of corrugated-core sandwich panels. Composites Part B: Engineering, **47**(0): p. 267-77.

Herrmann, A., P. Zahlen, and I. Zuardy (2005), Sandwich Structures Technology in Commercial Aviation, in *Sandwich Structures 7: Advancing with Sandwich Structures and Materials*, O.T. Thomsen, E. Bozhevolnaya, and A. Lyckegaard, Editors., Springer Netherlands. p. 13-26.

Kazemahvazi, S. and D. Zenkert (2009), Corrugated all-composite sandwich structures. Part 1: Modeling. Composites Science and Technology, **69**(7–8): p. 913-19.

Xiong, J., et al. (2011), Mechanical behavior and failure of composite pyramidal truss core sandwich columns. Composites Part B: Engineering, **42**(4): p. 938-45.

Lin, Y.-K., et al. (2007), Fracture evolution in thick composites under compression. Polymer Composites, **28**(4): p. 425-36.

Zenkert, D. (1995) An introduction to sandwich construction / Dan Zenkert: Engineering Materials Advisory Services Ltd, 1995.

Zhang Y, Z.S., Wang Z. (2011), Crush behavior of corrugated cores sandwich

panels. Advanced Materials Research, **217-218**: p. 1584-89.

Yokozeki, T., et al. (2006), Mechanical properties of corrugated composites for candidate materials of flexible wing structures. Composites Part A: Applied Science and Manufacturing, **37**(10): p. 1578-86.

Abaqus6.12-3 (2012), Analysis User's Manual. Dassault Systemes Simulia Corp. .

Adaptive Precise Integration BEM for Solving Transient Heat Conduction Problems

Bo Yu, *Weian Yao, Qiang Gao

State Key Laboratory of Structural Analysis for Industrial Equipment,
Dalian University of Technology, China

*Corresponding author: ywa@dlut.edu.cn

Abstract

A combined approach of boundary element method (BEM) and precise integration method (PIM) is presented for solving transient heat conduction problems with variable thermal conductivity. The boundary integral equation is derived by means of the Green's function for the Laplace equation. As a result, three domain integrals are involved in the integral equation. The radial integration method is used to transform the domain integrals into the boundary integrals. After discretization the solved domain by the BEM, a system of ordinary differential equations (ODEs) can be obtained. Adaptive PIM can solve efficiently ODEs and improve greatly the computational efficiency. Numerical examples show that the present approach can obtain satisfactory performance even for very large time step size. In addition, the results are independent of the time step size when the integral of free term can be analytically integrated, here, the free term is formed by boundary conditions and heat sources.

Keywords: Adaptive precise integration method, Radial integration method, Boundary element method, Transient heat conduction

Introduction

It is generally known that the finite difference method (FDM) is used to solve the transient heat conduction problems. However, the result of FDM is unstable when change the time step size. The precise integration method (PIM) [Zhong (1994)] can obtain stable and accurate results for different time step sizes. Particularly, the results are independent of the time step size when the free term can be divided into the functions of space and time and the time-related integral can be integrated analytically. Up to now, the PIM in conjunction with the finite element method (FEM) has been applied to conduct the transient heat transfer analysis [Cheng et al. (2004)], the transient forced vibration analysis of beams [Tang (2008)] and the sensitivity analysis and optimization problems [Xu et al. (2011)]. In addition, the method combining the PIM with meshless local Petrov–Galerkin method has been applied to the transient heat conduction problems [Li et al. (2011)].

Compared with FDM, FEM and the meshless method, BEM is very robust for solving the linear and homogeneous heat conduction problems [Song and Li (2003)]. However, BEM is still a challenge for solving nonlinear problems such as variable thermal conductivity problems. The main reason is that the fundamental solution of the problem obtains extremely difficult. Fortunately, we can use the fundamental solution of the linear problem to solve the nonlinear problem, whereas domain integrals are involved in resulting integral equations.

Generally, there are mainly two methods which can transform the domain integrals into the boundary integrals. The first one is the dual reciprocity method (DRM) [Nardini and Brebbia (1983)]. The deficiency of the method is that the particular solutions may be difficult to obtain for some complicated problems. In addition, even for known heat sources term, the method still requires an approximation of the known function. The second one is the radial integration method (RIM) [Gao (2002)]. The RIM not only can transform any complicated domain integral into the

boundary without using particular solution, but also can remove various singularities appearing in domain integrals. The method combining the RIM with the BEM is called the radial integration boundary element method (RIBEM).

The RIBEM has been widely applied to many fields including the crack analysis in functionally graded materials [Zhang (2011)], the heat transfer problems [AL-Jawary and Wrobel (2012); Yu et al. (2014a; 2014b; 2014c;)] and the viscous flow problems [Peng (2013)]. The RIBEM still exists a problem, which solved results are sensitive for different time step sizes when the problems are transient. The RIBEM and the PIM have been combined to solve transient heat conduction problems [Yu et al. (2014c)].

In this paper, an adaptive technique is introduced in the present method to improve the computational efficiency without affecting accuracy. First of all, we discretize the space domain by using the RIBEM to obtain a system of ordinary differential equations (ODEs) with respect to time, and then solve the ODEs by the PIM. Finally, two numerical examples are presented to validate the proposed method.

Governing Equation

Considering a two-dimensional bounded domain Ω with heat source and a spatially variable heat conductivity, the governing equation for transient heat conduction problems in isotropic media can be expressed as

$$\frac{\partial}{\partial x_i}\left(k\left(\mathbf{x}\right)\frac{\partial T\left(\mathbf{x},t\right)}{\partial x_i}\right)+f\left(\mathbf{x},t\right)=\rho c\left(\frac{\partial T\left(\mathbf{x},t\right)}{\partial t}\right) \qquad \mathbf{x}\in\Omega \tag{1}$$

where $\mathbf{x}=\left(x_1,x_2\right)$, $T\left(\mathbf{x},t\right)$ is the temperature at point $\mathbf{x}\in\Omega$ and at time t, $k(\mathbf{x})$ is the thermal conductivity, $f\left(\mathbf{x},t\right)$ is a known heat source, ρ is the density and c is the specific heat. The repeated subscript i denotes the summation through its range which is 2 for two-dimensional problem.

The initial condition is $T\left(\mathbf{x},0\right)=T_0$, where T_0 is a prescribed function. The boundary conditions are

$$T\left(\mathbf{x},t\right)=\overline{T}\left(\mathbf{x},t\right) \quad \mathbf{x}\in\Gamma_1 \tag{2}$$

$$-k\frac{\partial T}{\partial x_i}n_i=\overline{q}\left(\mathbf{x},t\right) \quad \mathbf{x}\in\Gamma_2 \tag{3}$$

where $\Gamma=\partial\Omega$, $\Gamma_1\cup\Gamma_2=\Gamma$, $\Gamma_1\cap\Gamma_2=\varnothing$, n_i is the i-th component of the outward normal vector $\mathbf{n}$ to the boundary Γ, $\overline{T}$ and $\overline{q}$ are prescribed temperature history and heat flux on the boundary, respectively.

Implementation of RIBEM

Boundary-domain Integral Equation

To derive the boundary integral equation, a weight function G is introduced to Eq. (1) and the following domain integrals can be written as

$$\int_\Omega G\frac{\partial}{\partial x_i}\left(k(\mathbf{x})\frac{\partial T}{\partial x_i}\right)\mathrm{d}\Omega+\int_\Omega Gf\mathrm{d}\Omega=\rho c\int_\Omega G\frac{\partial T}{\partial t}\mathrm{d}\Omega \tag{4}$$

Using Gauss' divergence theorem, the first domain integral can be manipulated as

$$\int_\Omega G(\mathbf{x},\mathbf{y})\frac{\partial}{\partial x_i}\left(k(\mathbf{x})\frac{\partial T}{\partial x_i}\right)d\Omega = \int_\Gamma G(\mathbf{x},\mathbf{y})k(\mathbf{x})\frac{\partial T}{\partial x_i}n_i d\Gamma - \int_\Gamma k(\mathbf{x})T\frac{\partial G(\mathbf{x},\mathbf{y})}{\partial x_i}n_i d\Gamma +$$

$$\int_\Omega T\frac{\partial G(\mathbf{x},\mathbf{y})}{\partial x_i}\frac{\partial k(\mathbf{x})}{\partial x_i}d\Omega + \int_\Omega k(\mathbf{x})T\frac{\partial}{\partial x_i}\left(\frac{\partial G(\mathbf{x},\mathbf{y})}{\partial x_i}\right)d\Omega$$

(5)

If Green's function $(-\ln r/2\pi)$ is acted as the weight function G, the last domain integral in Eq. (5) can be written as

$$\int_\Omega k(\mathbf{x})T\frac{\partial}{\partial x_i}\left(\frac{\partial G(\mathbf{x},\mathbf{y})}{\partial x_i}\right)d\Omega = -k(\mathbf{y})T(\mathbf{y})$$

(6)

where $r(\mathbf{x},\mathbf{y})$ is the distance between the source point $\mathbf{y}$ and the field point $\mathbf{x}$. Substituting the equation into Eqs. (4) and (5), it follows that

$$\tilde{T}(\mathbf{y},t)=-\int_\Gamma G(\mathbf{x},\mathbf{y})q(\mathbf{x},t)d\Gamma - \int_\Gamma\frac{\partial G(\mathbf{x},\mathbf{y})}{\partial\mathbf{n}}\tilde{T}(\mathbf{x},t)d\Gamma+\int_\Omega G(\mathbf{x},\mathbf{y})f(\mathbf{x},t)d\Omega +$$

$$\int_\Omega V(\mathbf{x},\mathbf{y})\tilde{T}(\mathbf{x},t)d\Omega -\rho c\int_\Omega\frac{G(\mathbf{x},\mathbf{y})}{k(\mathbf{x})}\frac{\partial\tilde{T}(\mathbf{x},t)}{\partial t}d\Omega$$

(7)

where $q(\mathbf{x},t)=-k(\mathbf{x})\partial T(\mathbf{x},t)/\partial\mathbf{n}$, $\tilde{k}(\mathbf{x})=\ln k(\mathbf{x})$, $\tilde{T}(\mathbf{x},t)=k(\mathbf{x})T(\mathbf{x},t)$, $V(\mathbf{x},\mathbf{y})=(\partial G(\mathbf{x},\mathbf{y})/\partial x_i)(\partial\tilde{k}(\mathbf{x})/\partial x_i)$ in which $q(\mathbf{x},t)$ is the heat flux, $\tilde{T}(\mathbf{x},t)$ and $\tilde{k}(\mathbf{x})$ are the normalized temperature and thermal conductivity, respectively. Eq. (7) is valid only for internal points. For boundary points, a similar integral equation can be obtained by letting $\mathbf{y}\to\Gamma$ as is done in the conventional BEM such as

$$c(\mathbf{y})\tilde{T}(\mathbf{y},t)=-\int_\Gamma G(\mathbf{x},\mathbf{y})q(\mathbf{x},t)d\Gamma - \int_\Gamma\frac{\partial G(\mathbf{x},\mathbf{y})}{\partial\mathbf{n}}\tilde{T}(\mathbf{x},t)d\Gamma+\int_\Omega G(\mathbf{x},\mathbf{y})f(\mathbf{x},t)d\Omega +$$

$$\int_\Omega V(\mathbf{x},\mathbf{y})\tilde{T}(\mathbf{x},t)d\Omega -\rho c\int_\Omega\frac{G(\mathbf{x},\mathbf{y})}{k(\mathbf{x})}\frac{\partial\tilde{T}(\mathbf{x},t)}{\partial t}d\Omega$$

(8)

where

$$c(\mathbf{y})=\begin{cases}1 & , \quad \mathbf{y}\in\Omega \\ \dfrac{\varphi(\mathbf{y})}{2\pi} & , \quad \mathbf{y}\in\Gamma\end{cases}$$

(9)

$\varphi(\mathbf{y})$ is the interior angle at a point $\mathbf{y}$ of the boundary Γ. Particularly, $c(\mathbf{y})=0.5$ if $\mathbf{y}$ is a smooth point on the boundary.

Transformation of Domain Integrals to the Boundary by RIM

In general, the heat source $f(\mathbf{x},t)$ is a known function. In this circumstances, RIM [Gao (2002)] can be directly used to transform the first domain integral in Eq. (8) into the boundary as follows:

$$\int_\Omega G(\mathbf{x},\mathbf{y})f(\mathbf{x},t)d\Omega(\mathbf{x})=\int_\Gamma\frac{1}{r(\mathbf{z},\mathbf{y})}\frac{\partial r}{\partial n}F^A(\mathbf{z},\mathbf{y},t)d\Gamma(\mathbf{z})$$

(10)

where the radial integral F^A can be expressed as $F^A(\mathbf{z},\mathbf{y},t)=\int_0^{r(\mathbf{z},\mathbf{y})}G(\mathbf{x},\mathbf{y})f(\mathbf{x},t)\xi\,\mathrm{d}\xi$.

For the last two domain integrals in Eq. (8), the RIM formulation cannot be directly used because $\tilde{T}$ and $\partial\tilde{T}/\partial t$ are unknown. To solve this problem, $\tilde{T}$ and $\partial\tilde{T}/\partial t$ are approximated by the combination of the radial basis functions (RBFs) and the polynomials in terms of global coordinates [Zhang (2011)]. Thus, $\tilde{T}$ and $\partial\tilde{T}/\partial t$ are respectively expressed as

$$\tilde{T}=\sum_{i=1}^{N}\alpha_i\phi_i(R)+a_1x_1+a_2x_2+a_3 \tag{11}$$

$$\frac{\partial\tilde{T}}{\partial t}=\sum_{i=1}^{N}\beta_i\phi_i(R)+b_1x_1+b_2x_2+b_3 \tag{12}$$

and the following equilibrium conditions have to be satisfied:

$$\sum_{i=1}^{N}\alpha_i=\sum_{i=1}^{N}\alpha_ix_{1,i}=\sum_{i=1}^{N}\alpha_ix_{2,i}=0 \tag{13}$$

$$\sum_{i=1}^{N}\beta_i=\sum_{i=1}^{N}\beta_ix_{1,i}=\sum_{i=1}^{N}\beta_ix_{2,i}=0 \tag{14}$$

where N is the total number of boundary and interior nodes, $R=r(\mathbf{x},\mathbf{x}_i)$ is the distance from the i-th application point $\mathbf{x}_i=(x_{1,i},x_{2,i})$ to the field point $\mathbf{x}$ and $\phi(R)$ is the RBF. In this paper, the compactly supported fourth-order spline RBF is adopted, i.e.,

$$\phi_i(R)=\begin{cases}1-6\left(\dfrac{R}{d_i}\right)^2+8\left(\dfrac{R}{d_i}\right)^3-3\left(\dfrac{R}{d_i}\right)^4 & 0\le R<d_i\\[2mm]0 & d_i\le R\end{cases} \tag{15}$$

in which d_i is radius of the supported region at the i-th point.

The coefficients α_i, a_1, a_2 and a_3 in Eq. (11) can be determined by collocating the application point $\mathbf{x}_i$ in Eq. (11) at all nodes. A set of algebraic equations can be written in the matrix form as $\tilde{\mathbf{T}}_\alpha=\boldsymbol{\phi}\boldsymbol{\alpha}$, where $\boldsymbol{\alpha}=\{\alpha_1,\alpha_2,\cdots,\alpha_N,a_1,a_2,a_3\}^{\mathrm{T}}$, $\tilde{\mathbf{T}}_\alpha=\{\tilde{T}_1,\tilde{T}_2,\cdots,\tilde{T}_N,0,0,0\}^{\mathrm{T}}=\{\{\tilde{\mathbf{T}}\}^{\mathrm{T}},\mathbf{0}\}^{\mathrm{T}}$. If no two nodes share the same coordinates, the matrix $\boldsymbol{\phi}$ is invertible and thereby $\boldsymbol{\alpha}=\boldsymbol{\phi}^{-1}\tilde{\mathbf{T}}_\alpha$. According to $\tilde{\mathbf{T}}_\alpha=\{\{\tilde{\mathbf{T}}\}^{\mathrm{T}},\mathbf{0}\}^{\mathrm{T}}$, the matrix $\boldsymbol{\phi}^{-1}$ can be expressed in the block form as $\left[\left(\tilde{\boldsymbol{\phi}}_1\right)_{(N+3)\times N},\left(\tilde{\boldsymbol{\phi}}_2\right)_{(N+3)\times 3}\right]$. Then $\boldsymbol{\alpha}$ can be rewritten as $\boldsymbol{\alpha}=\tilde{\boldsymbol{\phi}}_1\tilde{\mathbf{T}}$. Similarly, the coefficients in Eq. (12) can also be simply expressed as $\boldsymbol{\beta}=\tilde{\boldsymbol{\phi}}_1\dot{\tilde{\mathbf{T}}}$, where $\boldsymbol{\beta}=\{\beta_1,\beta_2,\cdots,\beta_N,b_1,b_2,b_3\}^{\mathrm{T}}$, $\dot{\tilde{\mathbf{T}}}=\{\partial\tilde{T}_1/\partial t,\partial\tilde{T}_2/\partial t,\cdots,\partial\tilde{T}_N/\partial t\}$.

Substituting Eqs. (11) and (12) into the last two domain integrals in Eq. (8), then transforming it into the boundary integrals by RIM, a pure boundary integral equation can be obtained as follows [Yu et al. (2014b)]:

$$c(\mathbf{y})\tilde{T}(\mathbf{y})=-\int_\Gamma G(\mathbf{x},\mathbf{y})q(\mathbf{x})\mathrm{d}\Gamma-\int_\Gamma\frac{\partial G(\mathbf{x},\mathbf{y})}{\partial\mathbf{n}}\tilde{T}(\mathbf{x})\mathrm{d}\Gamma+\int_\Gamma\frac{1}{r}\frac{\partial r}{\partial\mathbf{n}}F^A\mathrm{d}\Gamma+\mathbf{V}_\mathbf{y}\tilde{\mathbf{T}}-\mathbf{C}_\mathbf{y}\dot{\tilde{\mathbf{T}}} \tag{16}$$

where $\mathbf{V_y}$ and $\mathbf{C_y}$ are the boundary integral terms corresponding to the last two domain integrals in Eq. (8).

System of Differential Equations

Assuming that the boundary Γ is discretized into N_b linear elements and the region is distributed N_I internal nodes, the total number of nodes is $N=N_b+N_I$. Eq. (16) can be conveniently expressed in the following matrix form:

$$\mathbf{C}_\alpha \tilde{\mathbf{T}}_b = \mathbf{G}_b \mathbf{Q}_b - \mathbf{H}_b \tilde{\mathbf{T}}_b + \mathbf{f}_b + \mathbf{V}_b \tilde{\mathbf{T}} - \mathbf{C}_b \dot{\tilde{\mathbf{T}}} \qquad on \ \Gamma \qquad (17)$$

$$\tilde{\mathbf{T}}_I = \mathbf{G}_I \mathbf{Q}_b - \mathbf{H}_I \tilde{\mathbf{T}}_b + \mathbf{f}_I + \mathbf{V}_I \tilde{\mathbf{T}} - \mathbf{C}_I \dot{\tilde{\mathbf{T}}} \qquad in \ \Omega \qquad (18)$$

where

$$\mathbf{C}_\alpha = k\mathbf{diag}\left\{c(\mathbf{y}_1), c(\mathbf{y}_2), \cdots, c\left(\mathbf{y}_{N_b}\right)\right\}, \mathbf{Q}_b = \left\{-k\partial T_1/\partial n, \ -k\partial T_2/\partial n, \cdots, -k\partial T_{N_b}/\partial n\right\}^{\mathrm{T}}, \tilde{\mathbf{T}}_b = \left\{\tilde{T}_1, \tilde{T}_2, \cdots, \tilde{T}_{N_b}\right\}^{\mathrm{T}},$$

$\tilde{\mathbf{T}}_I = \left\{\tilde{T}_{I1}, \tilde{T}_{I2}, \cdots, \tilde{T}_{N_I}\right\}^{\mathrm{T}}, \mathbf{f}_b = \left\{f_1, f_2, \cdots, f_{N_b}\right\}^{\mathrm{T}}, \mathbf{f}_I = \left\{f_{I1}, f_{I2}, \cdots, f_{N_I}\right\}^{\mathrm{T}}$. The matrices $\mathbf{G}_b$, $\mathbf{H}_b$, $\mathbf{G}_I$ and $\mathbf{H}_I$ correspond to the coefficients of boundary integrals and $\mathbf{f}_b$, $\mathbf{V}_b$, $\mathbf{C}_b$, $\mathbf{f}_I$, $\mathbf{V}_I$ and $\mathbf{C}_I$ refer to the coefficients of domain integrals term.

After the application of boundary conditions and elimination the unknown heat flux quantity, a system of ordinary differential equations is obtained only relation to temperature as follows [Yu et al. (2014b)]:

$$\dot{\tilde{\mathbf{T}}}_u(t) = \mathbf{B}_u \tilde{\mathbf{T}}_u(t) + \mathbf{F}(t) \qquad (19)$$

Adaptive Precise Integration Method

The general solution of Eq. (19) can be written as

$$\tilde{\mathbf{T}}_u(t_{k+1}) = \mathbf{E}\tilde{\mathbf{T}}_u(t_k) + \int_0^{\Delta t} \exp(\mathbf{B}_u(\Delta t - \xi))\mathbf{F}(t_k + \xi)\mathrm{d}\xi \qquad (20)$$

where $\mathbf{E} = \exp(\mathbf{B}_u \Delta t)$ and $t_k = k\Delta t$. The matrix $\mathbf{E}$ can be rewritten as $\mathbf{E} = [\exp(\mathbf{B}_u \Delta t / m)]^m$, where m is an integer. Now, $m = 2^M$ is selected, where M is an integer. The following truncated Taylor series expansion can be used:

$$\exp(\mathbf{B}_u \eta) \approx \mathbf{I} + \mathbf{B}_u \eta + (\mathbf{B}_u \eta)^2 / 2! + \cdots + (\mathbf{B}_u \eta)^p / p! = \mathbf{I} + \mathbf{E}_a \qquad (21)$$

where $\eta = \Delta t / m$, $\mathbf{I}$ is the identity matrix. How to compute the matrix $\mathbf{E}$ has been detailedly shown in literature [Zhong (1994)].

The main factor of influence computation efficiency is how to select a optimal M and p. Because the most of the computational cost of PTI is the times of the matrix multiplications (TMM), where $\mathrm{TMM} = M + p - 1$. The optimal selection of TMM is shown in literature [Chen et al. (2004)] for different prescribed error tolerance. In addition, in Eq. (20), the function $\mathbf{F}(t_k + \xi)$ is formed by the known temperature boundary conditions, heat flux boundary conditions or heat sources. In this article, the term $\int_0^{\Delta t} \exp(\mathbf{B}_u(\Delta t - \xi))\mathbf{F}(t_k + \xi)\mathrm{d}\xi$ in Eq. (20) is analytically integrated for all numerical example.

Finally, true temperature $T(\mathbf{x}, t)$ can be computed by using $T(\mathbf{x}, t) = \tilde{T}(\mathbf{x}, t) / k(\mathbf{x})$.

Numerical Examples

To check the convergence of the proposed method, the root mean square (RMS) error is given by

$$RMS = \sqrt{\sum_{i=1}^{N}(T_{numerical,i} - T_{exact,i})^2 / \sum_{i=1}^{N}T_{exact,i}^2} \qquad (22)$$

where $T_{numerical,i}$ and $T_{exact,i}$ are the numerical solution and the exact solution of the *i*-th node, respectively. For comparison, two examples are also computed by using the RIBEM, which use the finite difference technique to simulate the derivative of temperature with respect to time (it will be abbreviated to RIBEM-FD) [Yu et al. (2014b)].

Example 1: In this example, a square plate $\Omega = [1,2]^2$ is considered with $k(\mathbf{x}) = x_1 + x_2$, $\rho = 1$ and $c = 1$. The initial condition and the heat source are $T_0 = x_1^2 + x_2^2$ and $f(\mathbf{x},t) = -6(x_1 + x_2) + 10\cos(10t)$, respectively. The boundary conditions are given by $T(x_1,1,t) = x_1^2 + 1 + \sin(10t)$, $T(2,x_2,t) = 4 + x_2^2 + \sin(10t)$, $T(x_1,2,t) = x_1^2 + 4 + \sin(10t)$, $T(1,x_2,t) = 1 + x_2^2 + \sin(10t)$. The exact solution of the problem is $T(\mathbf{x},t) = x_1^2 + x_2^2 + \sin(10t)$. The plate is discretized into 20 equally space linear boundary elements and distributed uniformly 16 internal nodes.

Table 1. The value of TMM **for different** ε_p

Δt	TMM $\varepsilon_p = 10^{-5}$	TMM $\varepsilon_p = 10^{-6}$	TMM $\varepsilon_p = 10^{-7}$	TMM $\varepsilon_p = 10^{-8}$	TMM $\varepsilon_p = 10^{-9}$	TMM $\varepsilon_p = 10^{-10}$	TMM $\varepsilon_p = 10^{-11}$	TMM $\varepsilon_p = 10^{-12}$
0.2	15	15	16	16	16	16	17	18
5	19	19	20	20	20	20	21	22

Table 1 shows the optimal value of TMM for different time step sizes and computational error tolerance. Comparison with the general selection TMM=23, the adaptive PTI improves the computational efficiency greatly. For different time step size, it can be seen from Figure 1 that the RMS errors of the PIBEM are highly coincident, but the errors of the RIBEM-FD emerge a big fluctuation.
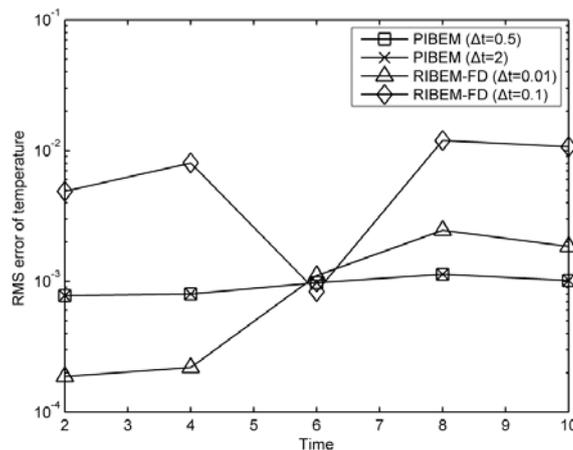
Figure 1. RMS error of temperature with $\varepsilon_p = 10^{-5}$ **for example 1.**

Example 2: In this example, we consider a concave geometry with $k(\mathbf{x}) = \exp(x_1)$, $\rho = c = 1$ and $0 < t \leq 1$. The initial temperature and the heat source are $T_0 = 0$ and $f(x,t) = 10$, respectively. The time-dependent temperature condition is $T(0, y, t) = 60t$ for the left boundary and the other boundaries are insulated. The geometry and computational model of the BEM can be seen in Figure 2 with 36 boundary elements and 13 internal nodes. The problem is also computed using the FEM software ANSYS, which the results are considered as the reference solutions T_{exact} in Eq. (22). The solved domain is uniformly discretized into 832 4-noded elements. Table 2 shows the optimal value of TMM for different time step sizes and computational error tolerance. It can be seen from Figure 3 that the solutions of PIBEM are very stable and accurate than the solutions of RIBEM-FD for the different time step size.

Table 2. The value of TMM for different ε_p

Δt	TMM $\varepsilon_p = 10^{-5}$	TMM $\varepsilon_p = 10^{-6}$	TMM $\varepsilon_p = 10^{-7}$	TMM $\varepsilon_p = 10^{-8}$	TMM $\varepsilon_p = 10^{-9}$	TMM $\varepsilon_p = 10^{-10}$	TMM $\varepsilon_p = 10^{-11}$	TMM $\varepsilon_p = 10^{-12}$
0.001	9	9	10	10	10	10	11	12
0.2	17	17	18	18	18	18	19	20
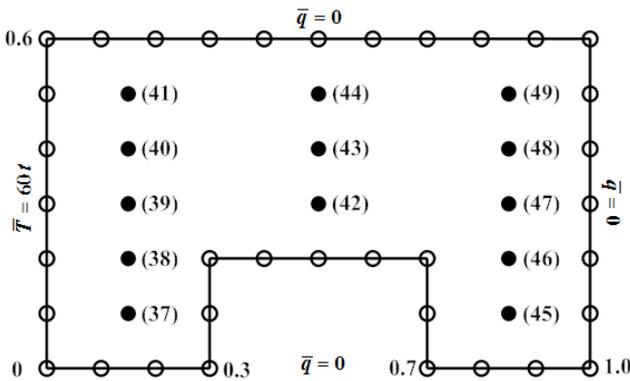
Figure 2. Computational model of the BEM for example 2.
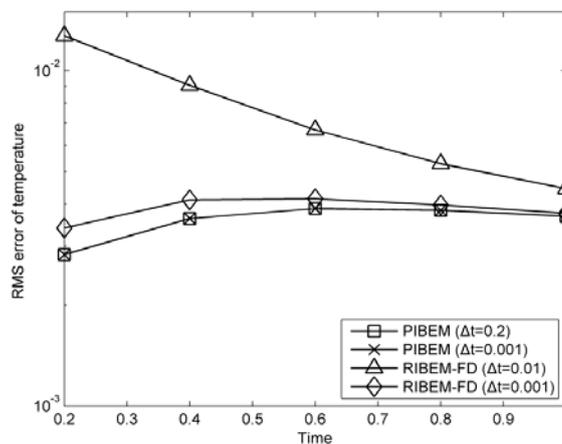
Figure 3. RMS error of temperature with $\varepsilon_p = 10^{-5}$ for example 2.

Conclusions

In this paper, the adaptive PIM is introduced into the RIBEM for solving the transient heat conduction problems with variable thermal conductivity. For the RIBEM-FD, the sensitive results are caused by the finite difference method to solve the derivative of temperature with respect to time. The PIBEM can perfectly solve the problem. Numerical examples show the PIBEM with adaptive technique can obtain the stable and accurate results for a big time step size and improve efficiency, whereas only in the case of a small time step the RIBEM-FD can obtain accurate results.

References

AL-Jawary, M. A.,and Wrobel, L. C., (2012) Radial Integration Boundary Integral and Integro-differential Equation Methods for Two-dimensional Heat Conduction Problems with Variable Coefficients, *Engineering Analysis with Boundary Elements*, **36**, 685-695.

Chen, B. S., Tong, L. Y., Gu, Y. X., Zhang, H. W.,and Ochoa, O., (2004) Transient Heat Transfer Analysis of Functionally Graded Materials Using Adaptive Precise Time Integration and Graded Finite Elements, *Numerical Heat Transfer, Part B*, **45**, 181-200.

Gao, X. W., (2002) The Radial Integration Method for Evaluation of Domain Integrals with Boundary-only Discretization, *Engineering Analysis with Boundary Elements*, **26**, 905-916.

Li, Q. H., Chen, S. S., and Kou, G. X., (2011) Transient heat conduction analysis using the MLPG method and modified precise time step integration method, *Journal of Computational Physics*, **230**, 2736-2750.

Nardini, D.,and Brebbia, C. A., (1983) A New Approach for Free Vibration Analysis Using Boundary Elements, *Applied Mathematical Modelling*, **7**, 157-162.

Peng, H. F., Cui, M. and Gao, X. W. (2013) A Boundary Element Method without Internal Cells for Solving Viscous Fow Problems, *Engineering Analysis with Boundary Elements*, **37**, 293-300.

Song, S. P., and Li, B. Q., (2003) Boundary Integral Solution of Thermal Radiation Exchanges in Axisymmetric Furnaces, *Numerical Heat Transfer, Part B*, **44**, 489-507.

Tang, B., (2008) Combined Dynamic Stiffness Matrix and Precise Time Integration Method for Transient Forced Vibration Response Analysis of Beams, *Journal of Sound and Vibration*, **309**, 868–876.

Xu, W. T., Zhang, Y. H., Lin, J. H., Kennedy, D. and Williams, F. W., (2011) Sensitivity Analysis and Optimization of Vehicle–bridge Systems Based on Combined PEM–PIM Strategy, *Computers & structures*, **89**, 339-345.

Yu, B., Yao, W. A., Gao, X. W., ans Zhang S., (2014a) Radial Integration BEM for One-Phase Solidification Problems, *Engineering Analysis with Boundary Elements*, **39**, 36–43.

Yu, B., Yao, W. A., Gao, X. W., and Gao Q., (2014b) A combined approach of RIBEM and precise time integration algorithm for solving transient heat conduction problems, *Numerical Heat Transfer, Part B*, **65**,155-173.

Yu, B., Yao, W. A., and Gao Q., (2014c) A precise integration boundary element method for solving transient heat conduction problems with variable thermal conductivity, *Numerical Heat Transfer, Part B*, **65,** 472–493.

Zhong, W. X., (1994) On precise time-integration method for structural dynamics, *Journal of Dalian University of Technology* **34**, 131-136. (in Chinese).

Zhang, C., Cui, M., Wang, J., Gao, X. W., Sladek, J., and Sladek, V., (2011) 3D Crack Analysis in Functionally Graded Materials, *Engineering Fracture Mechanics*, **78**, 585-604.

Validation DualSPHysics code for liquid sloshing phenomena

*Sergei.K. Buruchenko[1], Alejandro J.C. Crespo[2]

[1]South Ural State University, Snezhinsk, Russia.
[2]Universidade de Vigo, Spain.

*Corresponding author: askskb@gmail.com

Abstract

The DualSPHysics code is proposed as a numerical tool for the simulation of liquid sloshing phenomena. A particular type of sloshing motion can occur during the core meltdown of a liquid metal cooled reactor (LMR) and can lead to a compaction of the fuel in the center of the core possibly resulting in energetic nuclear power excursions. This phenomenon was studied in series of "centralized sloshing" experiments with a central water column collapsing inside the surrounding cylindrical tank. These experiments provide data for a benchmark exercise for accident analysis codes. To simulate "centralized sloshing" phenomena, a numerical method should be capable to predict the motion of the free surface of a liquid, wave propagation and reflection from the walls. The DualSPHysics code based on the smoothed particle hydrodynamics method was applied to the simulation of "centralized sloshing" experiments. Simulation results are compared with the experimental results. In a series of numerical calculations it is shown that overall motion of the liquid is in a good agreement with experimental observations. Dependence on the initial and geometrical symmetry is studied and compared with experimental data.

Keywords: ICCM2014, Computational method, Sloshing Experiment, Smoothed Particle Hydrodynamics

Introduction

The problem of safety in nuclear reactors has been intensively studied from the time of the development of the first reactor designs. Over time, several severe accidents occurred at nuclear reactors, but without dangerous consequences for the environment, until the accident at Chernobyl Nuclear Power Plant (NPP) occurred in 1986 and massive severe accident at Fukushima NPP (2011), where four units were seriously damaged by a tsunami wave. Today it is clear that further successful development of the nuclear energy industry is impossible without deeper knowledge of severe accidents and without the provision of safety guarantees to the public, based on comprehensive analyses of nuclear reactor safety.

One of the current problems in severe accident analysis is the problem of molten corium motion, which could possibly result in a recriticality event. The movement of the corium during an accident involving melting of the reactor core may be initiator of a recriticality event with dangerous high power excursions.

Simulation with Eulerian methods is difficult, since special treatments are required for capturing the indicated phenomena. To be more specific, the treatments required are the Volume Of Fluid (VOF) method, combined with mesh refinement, for tracking the free-surface, and sliding meshes, for the connection between the moving and stationary meshes. The above treatments increase the computational cost and requirements of the simulation considerably. An alternative way of simulating the flow is the Smoothed Particle Hydrodynamics (SPH) method which will be used in the present work for the simulations.

The SPH method was initially developed by [Lucy (1977)], [Gingold & Monaghan (1977)] and has been used for modeling astrophysical problems. The application of SPH to a wide range of scientific areas has led to significant extensions and improvements of the original method [Monaghan (2005)], [Liu(2003)]. SPH is a Lagrangian, particle, mesh-less method and has the

advantages of tracing and resolving the free-surface without any special treatment and describing moving/deforming boundaries easily.

The DualSPHysics code [Gomez-Gesteira et al. (2012a, 2012b)] is proposed as a numerical tool for the simulation of liquid sloshing phenomena. A particular type of sloshing motion can occur during the core meltdown of a liquid metal cooled reactor (LMR) and can lead to a compaction of the fuel in the center of the core possibly resulting in energetic nuclear power excursions. This phenomenon was studied in series of "centralized sloshing" experiments [Maschek et al. (1992a, 1992b)] with a central water column collapsing inside the surrounding cylindrical tank. These experiments provide data for a benchmark exercise for accident analysis codes. To simulate "centralized sloshing" phenomena, a numerical method should be capable to predict the motion of the free surface of a liquid, wave propagation and reflection from the walls.

Standart SPH formalism

The SPH formalism relies on the use of kernel approximation of field functions for the calculation of the operators appearing in the discretization of the flow equations, instead of using a computational grid. In this way it is able to approximate derivatives or functions from unconnected and randomly scattered computation points. The basis of the SPH approximations originates from the following identity:

$$f(x) = \int f(x')\delta(x - x')dx' \tag{1}$$

where $f(x)$ is a function of three dimensional position vector **x**, $\delta(x - x')$ is the Dirac delta distribution and Ω is the volume of the integral that contains **x**. The above relation can be approximated using a smoothing kernel function $W(x - x', h)$:

$$\langle f(x) \rangle = -\int_{\Omega} f(x')W(x - x', h)dx' \tag{2}$$

A similar equation can be derived for the gradient of a function:

$$\langle \nabla f(x) \rangle = -\int_{\Omega} f(x')\nabla W(x - x', h)dx' \tag{3}$$

In order the above approximations to be valid, the kernel function $W(x - x', h)$ has to fulfill certain requirements, such
as:
• Unity or normalization condition : $\int W(x - x', h)dx = 1$
• Dirac distribution property : $\lim_{h \to 0} W(x - x', h) = \delta(x - x')$
• Compact condition : $W(x - x', h) = 0$, for $|x - x'| > k \cdot h$, where $k \cdot h$ is the kernel's support domain
• Also the kernel function has to be even, positive and monotonically decreasing function.

There are many types of kernel functions. In the present work the quintic kernel is used [Monaghan (2005)]

$$W(q) = \frac{21}{16\pi h^3}\left(1 - \frac{q}{2}\right)^4 (2q + 1) \quad 0 \leq q \leq 2 \tag{4}$$

where $q = |r|/h$, with $|r|$ the distance between two computational points and h a characteristic smoothing length.

In the SPH method the entire system is represented with a finite number of particles that carry individual mass, occupy individual space and the characteristic quantities of the flow (e.g. velocity, density, pressure etc.). Thus the continuous integral relations can be written in the following form of discretized particle approximation:

$$\langle f(x_i) \rangle = \sum_{j=1} \frac{m_j}{\rho_j} f(x)W_{ij} \tag{5}$$

$$\langle \nabla f(x_i) \rangle = \sum_{j=1} \frac{m_j}{\rho_j} f(x) \nabla W_{ij} \tag{6}$$

In the above equations $W_{ij} = W(x_i - x_j, h)$, m_j is the j particle's mass and p_j is the j particle's density.

Using the above approximations for a function and the derivative of a function, one can derive the SPH flow equations [Monaghan (2005)]:

Momentum equation:

$$\frac{dV_i^{\alpha}}{dt} = -\sum_j m_j \left(\frac{P_i}{\rho_i^2} + \frac{P_j}{\rho_j^2} + \Pi_{ij} \right) W_{ij,\alpha} + g \tag{7}$$

Continuity equation:

$$\frac{d\rho_i}{dt} = \rho_i \sum_j \frac{m_j}{\rho_j} \left(V_i^{\alpha} - V_j^{\alpha} \right) W_{ij,\alpha} \tag{8}$$

where g - gravity acceleration, Π_{ij} is the viscosity term suggested by [Monaghan (2005)], V_i^{α} is velocity, P_i is pressure, *latin* indexes denotes particles number, *greek* index denote coordinate direction.

$$\frac{\partial W_{ij}}{\partial x_i^{\alpha}} = W_{ij,\alpha} \tag{9}$$

Pressure is calculated from an equation of state, thus the method is weakly compressible. The Tait equation of state is commonly used .

$$p = \frac{\rho_0 c_0^2}{\gamma} \left[\left(\frac{\rho}{\rho_0} \right)^{\gamma} - 1 \right] \tag{10}$$

In the above equation $\gamma = 7$, ρ_0 is the reference density and c_0 is an artificial speed of sound, since the real speed of sound would require a very small time step. In order to keep density variations less than 1%, the value of c_0 is chosen ~10 V_{max}, according to [Monaghan (2005)].

The dynamic boundary conditions described in [Crespo et al., (2007)] are used in this work. The boundary particles satisfy the same continuity equation as the fluid particles, therefore, their density and pressure also evolve. Hence, when a fluid particle approaches a boundary particle, and they are at the interaction distance defined by the kernel range, the density of the boundary particles increases giving rise to an increase distance of the pressure and the force exerted on the fluid particle also increases due to the pressure term in the momentum equation creating a repulsive mechanism between fluid and boundary.

Implennmention details

The SPH scheme presented in the previous section is implemented in the DualSPHysics code. The code is implemented using both the C++ and CUDA programming languages. The code can then be executed either on the CPU or on the GPU since all computations have been implemented both in C++ for CPU simulations and in CUDA for the GPU simulations. The philosophy underlying the development of DualSPHysics is that most of the source code is common to CPU and GPU which makes debugging straightforward as well as the code maintenance and new extensions. This allows the code to be run on workstations without a CUDA-enabled GPU, using only the CPU implementation. On the other hand, the resulting codes should be necessarily different since code developers have considered efficient approaches for every processing unit.

Computational runtime increases dramatically with the number of particles in the SPH simulations. Hence, parallelisation methods are essential to run simulations with a huge number of particles in a reasonable execution time. GPUs constitute a suitable hardware for scientific tasks where mathematical calculations are carried out using large sets of data.

Optimization strategies for CPU and GPU implementations of a smoothed particle hydrodynamic method [Gomez-Gesteira et al. (2012a, 2012b)] introduced the framework to implement SPH codes using the best techniques and performance optimizations on GPU. That work focused on identifying suitable algorithms for efficient parallelization since a proper and full use of all the capabilities of the GPU architecture is not straightforward. As an initial step, the implementation focused on solving the particle interactions on a GPU using CUDA and the next step was the implementation of the neighbour list and the time integration in order to develop an entire GPU-SPH model.

Experiment description

The DualSPHysics code described in the previous sections, has been applied to the numerical simulation of the three-dimensional sloshing liquid motion problems. These problems has been experimentally studied in KfK (presently KIT) in the framework of the safety analysis of fast nuclear reactors [Maschek (1992a)].

The sloshing experiments had two main objectives. The first was to obtain a better understanding of centralized sloshing phenomenon. In a hypothetical severe accident of a fast nuclear reactor, a possible recriticality may occur following core melting and relocation of the fissile materials.

The second purpose of the sloshing liquid motion experiments was to provide data for a benchmark exercise for reactor accident analysis codes [Maschek(1992b)]. These data were subsequently used to verify and validate the SIMMER-III/IV reactor safety analysis code [Shirakawa (2008)], [Yamano et. Al (2008)].

Experiments were performed with water under normal conditions. The experimental installation consists of a cylindrical container separated by a membrane into two coaxial parts. The container was opened, so that the environment is air under atmospheric pressure. At the initial moment, the membrane is quickly moved up, resulting in the water column collapsing under the force of gravity.
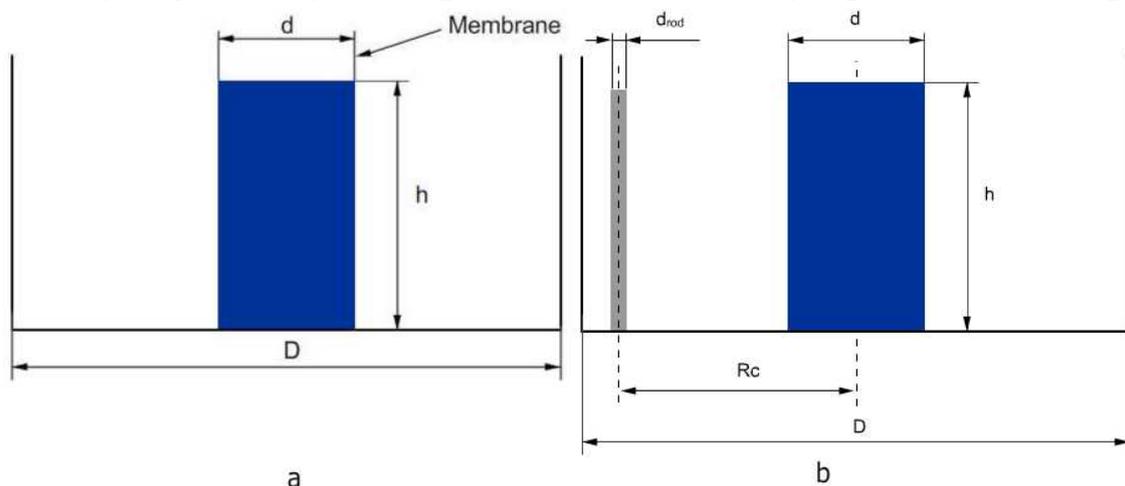
Fig. 1 Central sloshing experiment geometry.

In Fig.1 (a) and Fig.1 (b), the different experimental configurations at the beginning of the experiments are presented:

(a) a fully symmetrical configuration with no obstacles in the flow

(b) an asymmetrical configuration with no obstacles

(c) a symmetrical configuration with a rod bank installed around the liquid column.

For the asymmetrical case, the same experimental and computational domain geometry is used, but the position of the water column is shifted by an offset of 8.25 cm from the container center.

The parameters of the numerical model used in the calculations for the symmetrical geometry are: container diameter D = 0.44 m, diameter of water column d = 0.11 m, height of water column h = 0.2 m, water density 1000 kg/m^3.

The geometry of the numerical model for the test cases with rod imitators is the same as for the experimental series without rods. The difference is the presence of twelve vertical rods equidistantly positioned around the water column. Their distance from the center is R_c = 17.6 cm . The rod diameter (d_{rod}) in the experiments was 2 cm, to simulate a blockage ratio similar to that in a real reactor pool. The same value for the rod diameter has been used in the numerical model. An overview and sketch of the experimental setup, with geometrical sizes, for these test series are given in Fig. 1 (b).

The initial velocity field in water is zero. The initial pressure field is hydrostatic:

$$p = \rho g h \tag{11}$$

Experiment and numerical results

In this section the results of the simulation of the centralized sloshing experiment in the symmetrical geometry, the asymmetrical geometry and experiment geometry with 12 rod bank are presented.

The main quantitative parameters for the symmetrical case are the arrival time of the liquid at the wall, the time and height of the maximal wave at the wall, and the time and height of the central peak. The central peak height is the most important of these for the recriticality analysis, and, as has been found, the peak height is difficult to reproduce in the numerical simulation.
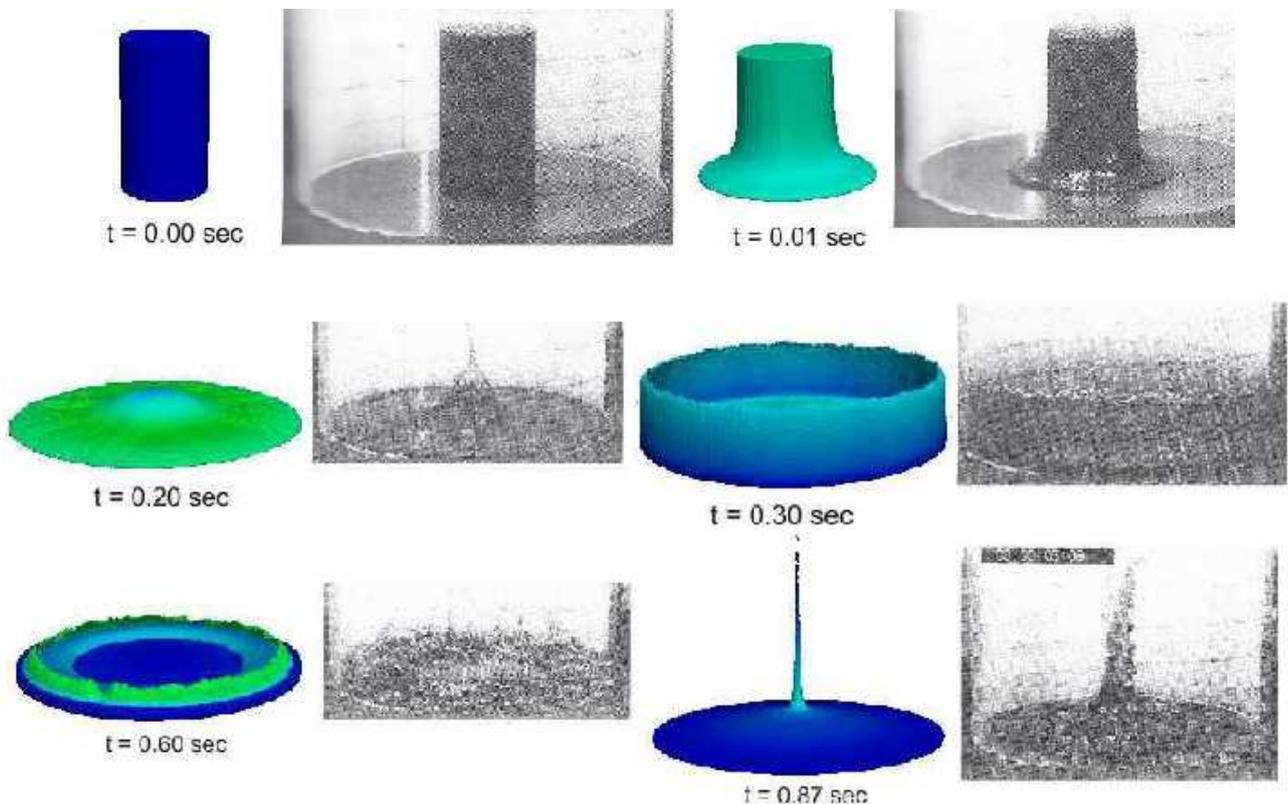
Figure 2. Experiment and simulation result for symmetrical case.

Furthermore, a correct definition and measurement of the central peak height is not very obvious. In the applied experimental technique, large drops on top of the peak moving with the same velocity as the bulk flow were included in the height measurement (see the more detailed discussion of the definition of the central peak value in the following subsection).
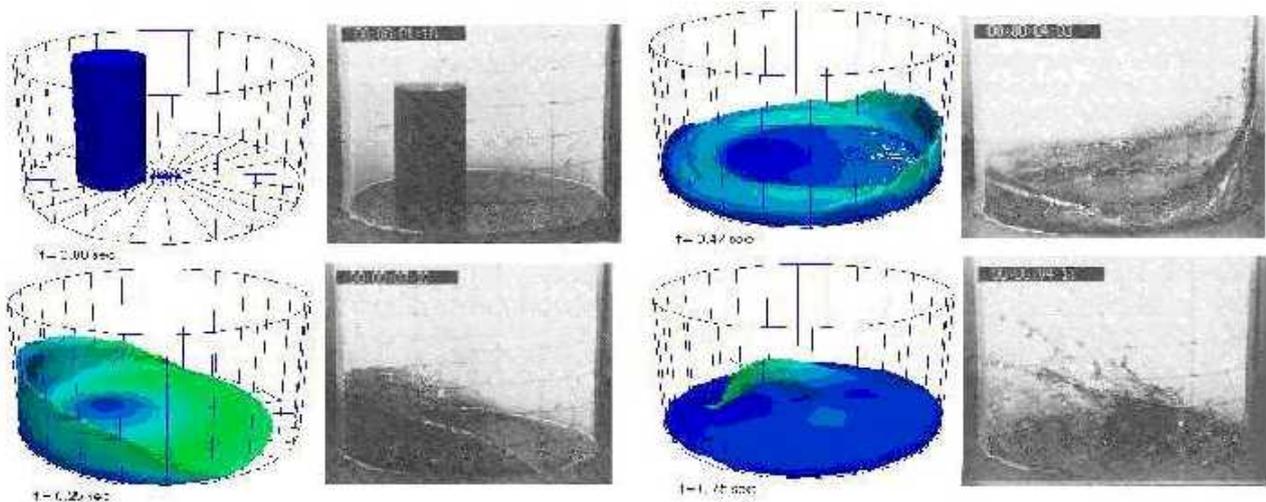
Figure 3. Experiment and simulation result for asymmetrical case.

T=0 sec, T=0.25 sec, T=0,47 sec, T=0.75 sec

For the asymmetrical case, only the timing of the maximum height at the walls and the height of the maxima were measured in the experiments.

Fig. 2-4 shows a visualization of the results of the simulation in comparison with the experimental observations of the liquid sloshing motion.
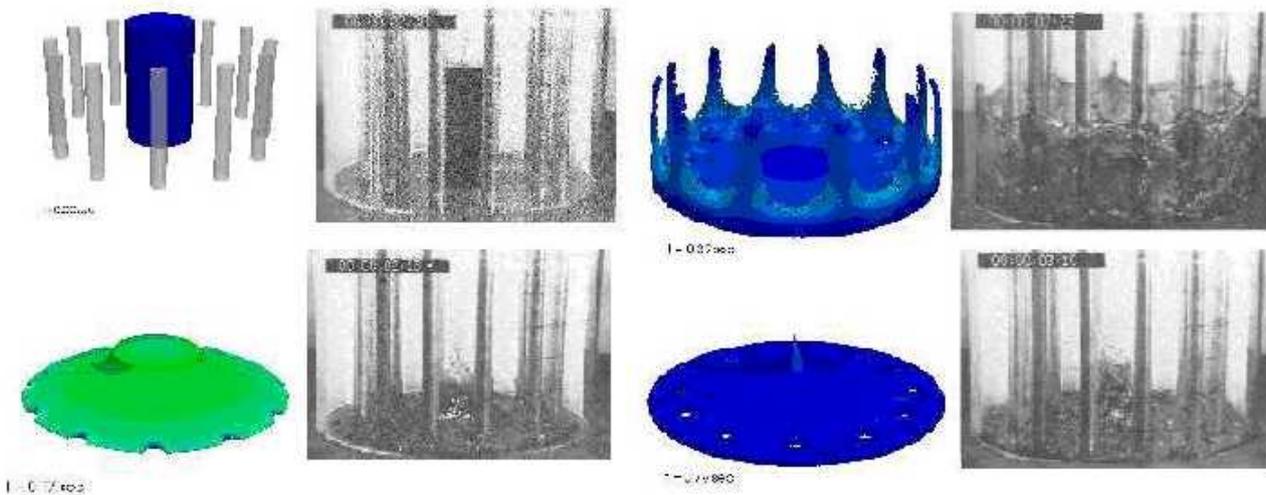
Figure 4. Experiment and simulation results with 12 rod bank.

T=0 sec, T=0.21 sec, T=0.36 sec, T=0.86 sec.

The results are summarized in Table 1-3. Most of the quantitative simulation results are in a good agreement with both experimental data and the numerical results predicted by the reactor safety analysis code SIMMER-IV, although some deviations in the central peak value are observed. Also, for the asymmetrical case, a lower value of the height of the right-hand slosh is predicted. A similar value was obtained using the SIMMER-IV code. For the asymmetrical geometry, the lower values may be due to the relatively low resolution of the numerical model.

Table 1. 3D Central Sloshing: Symmetrical Case

Symmetrical case	Slosh at outer container wall			Slosh at pool center	
	Arrival time at wall [s]	Time of maximum heights [s]	Maximum heights [cm]	Time of maximum height [s]	Maximum height [cm]
Experiment	0.20±0.02	0.42±0.02	16±1.0	0.88±0.04	40±5
SPH result 2.8 M particles	0.21	0.39	14	0.88	0.33
SIMMER-IV (coarse mesh - 44×44×100)	0.20	0.40	17.25	0.88	36
SIMMER-IV (fine mesh - 92×92×100)	0.20	0.38	18.75	-	>50 (overestimated)

Table 2. 3D Central Sloshing: Asymmetrical Case

Asymmetrical Case	Slosh at pool center		Slosh at right wall	
	Time of maximum height [s]	Maximum height [cm]	Time of maximum height [s]	Maximum height [cm]
Experiment	0.36±0.02	14.0±2.0	0.48±0.02	24±2.0
SPH result 2.8 M particles	0.36	14.5	0.48	21.5
SIMMER-IV	0.36	17.25	0.48	21

Table 3. 3D Central Sloshing: Symmetrical Case. Vertical Rod Bank

Vertical Rod Bank	Slosh at outer wall			Slosh at pool center	
	Arrival time at wall [s]	Time of maximum heights [s]	Maximum height [cm]	Time of maximum height [s]	Maximum height [cm]
Experiment	0.20±0.02	0.42±0.02	15±1.0	0.88±0.04	15±3
SPH result 2.8 M particles	0.21	0.39	16	0.86	0.12

The comparative snapshots from the experiment and numerical simulation for the test cases with 12 rod bank are presented in Fig. 4.

Conclusion

The DualSPHysics computer code based on the SPH method has been applied to the numerical simulation of the three-dimensional sloshing liquid motion problem. A number of numerical models have been created to reflect different configurations of the experimental installations. These are as follows:

- Fully symmetrical configuration: the liquid column is symmetrically located at the container center.

- Asymmetrical configuration: the liquid column is located with an offset from the container center.

- Symmetrical configuration with obstacles: rod imitators are installed around the liquid column.

The quantitative parameters of the flows predicted by the numerical algorithm have been compared with the available results of the simulations performed with the SIMMER-III/IV reactor safety analysis code and with experimental data. These measured flow quantities, such as the heights of the wall sloshes and the central peak, and the timings of these events, are accurately predicted with high resolution simulations. At the same time, the present algorithm based on the SPH method is capable of resolving the high central peak in the fully symetrical case, which was an issue for the SIMMER code.

A sensitivity study for the value of the central peak height in the symmetrical configuration has also been performed. The study showed the convergence of the central peak height value with an increase in the number of particles used for modeling.

In analyzing for a possible recriticality event, the height values of the central peak calculated for the different experimental configurations and different resolutions of the numerical model were compared. The highest peak, corresponding to the maximal volume of the fissile materials compacted in the center of the pool, is observed in calculations of the fully symmetrical configuration with the fine numerical resolution. Thus conclusion demonstrates the experimentally observed sensitivity of the liquid flow to the geometrical asymmetries of the vessel.

References

Lucy L., (1977) "A numerical approach to the testing of the fission hypothesis", *Astron. J. Vol.* 82, pp. 1013–1024.
Gingold R.A., Monaghan J.J. (1977) "Smoothed particle hydrodynamics: theory and application to non- spherical stars", *Mon. Not. R. Astr. Soc.*, Vol. 181, pp. 375–389.
Monaghan J.J., (2005) "Smoothed Particle Hydrodynamics", *Rep. Prog. Phys.* Vol. 68, pp. 1703–1759.
Liu G.R., Liu M.B., (2003), "Smoothed particle hydrodynamics: a meshfree particle method", World Scientific Publishing Company, December 2003.
Gómez-Gesteira M., Rogers B.D., Crespo A.J.C., Dalrymple RA, Narayanaswamy M and Domínguez JM. (2012a). SPHysics - development of a free-surface fluid solver- Part 1: Theory and Formulations. Computers & Geosciences, 48: 289-299.
Gómez-Gesteira M., Crespo A.J.C., Rogers B.D., Dalrymple R.A., Domínguez J.M. and Barreiro A., (2012b). "SPHysics - development of a free-surface fluid solver- Part 2: Efficiency and test cases*", Computers & Geosciences*, Vol. 48, pp. 300-307.
Maschek W., Munz C.D., Meyer L., (1992a) "Investigations of sloshing fluid motions in pools related to recriticalities in liquid-metal fast breeder reactor core meltdown accidents", *Nucl. Techn.*, 1992. Vol. 98(1), p. 27.
Maschek W., Roth A., Kirstahler M., Meyer L., (1992b) "Simulation experiments for centralized liquid sloshing motions", Karlsruhe Research Centre Report KfK 5090, Dez. 1992.
Crespo, A.J.C., Gómez-Gesteira, M., Dalrymple, R.A. (2007) "Boundary conditions generated by dynamic particles in SPH methods", *CMC: Computers, Materials, & Continua*, Vol. 5(3), pp. 173–184.
Shirakawa N., (2008)"R&D of the next generation safety analysis methods for fast reactors with new computational science and technology", *Proc. of 14th JAEA-FZK/CEA/IRSN/ENEA SIMMER-III/IV Review Meeting*. Karlsruhe, Germany, September 9–12, 2008.
Yamano H., Hosono S., Sugaya M., (2008)"Analysis of sloshing experiments", *Proc. of 14*[th] *JAEA-FZK/CEA/IRSN/ENEA SIMMER-III/IV Review Meeting*, Karlsruhe, Germany, September 9–12, 2008.

Topology Optimization with Shape Preserving Design

* ZHU Ji-Hong*, LI Yu, ZHANG Wei-Hong

[1] Engineering Simulation and Aerospace Computing (ESAC), The Key Laboratory of Contemporary Design and Integrated Manufacturing Technology, Northwestern Polytechnical University, Xi'an, P.R. China

*Corresponding author: jh.zhu@nwpu.edu.cn

Abstract

This paper is to present an extended shape preserving topology optimization formulation aiming at preserving specific local structural domain configuration. By introducing Artificial Week Elements (AWE) established with respect to shape preserving control points, we constrain its elastic strain energy to suppress the warping deformation. Compared with the existing global compliance topology optimization, this formulation acts as a control of local compliance of the structure. Numerical results have shown how the strain energy constraint related to AWE influences the optimized solution, especially the effect of the upper limit of the constraint. Comparative studies have evidently shown that the effect of shape preserving can be successfully achieved. Possible structural distortions are also illustrated in order to have an in-depth understanding of the design mechanism.

Keywords: Topology Optimization, Shape Preserving, Artificial Weak Elements, Warping Deformation, Local Strain Energy

Introduction

Topology optimization method has been developed as one of the most effective techniques in saving structural weight and improving multidisciplinary performances. Recent advances of topology optimization techniques have been summarized by excellent literature surveys such as Guo and Cheng (2010), Sigmund and Maute (2013), Deaton and Grandhi (2014).

Meanwhile, different topology optimization formulations were also presented to obtain required structural deformation patterns. In these literatures, constraints on a single or multiple nodal displacements were normally issued. For example, in the works of Liu et al. (2008), warping deformation of beam cross-section was considered in a new anisotropic beam theory as well as in topology optimization. Rong and Yi (2010) designed the multi-points displacements using a newly developed phase transferring method. Typically, in the works of Qiao and Liu (2012), a geometric average displacement function integrating the deformation field, which was similar to a P-norm scheme, was proposed to minimize the structural maximum deformation. In this way, the magnitudes of different nodal displacements were controlled to form a better deformation. Other displacement designs can be found mostly in topology optimization of compliant mechanisms (see e.g. Wang et al. 2005, Stanford et al. 2012 and 2013).

However, constraints on the magnitudes of nodal displacements might not appropriate in many complicated engineering cases searching better structural deformation behaviors. For example, challenges of suppressing structural local warping deformation to maintain structural coordinative displacements are always faced during the aircraft structure design, manufacturing and assembling (Niu 1988, Barrett 1992, Wang 2000), which are considered as shape preserving design. Key difficulties lies in that the popularly used global compliance and nodal displacements in topology optimization cannot effectively describe and suppress the local warping deformation.

Therefore, this paper proposes to implement multi-point shape preserving constraints in an extended topology optimization formulations by introducing strain energy based quantitative approach describing warping deformation magnitudes in shape preserving domain.
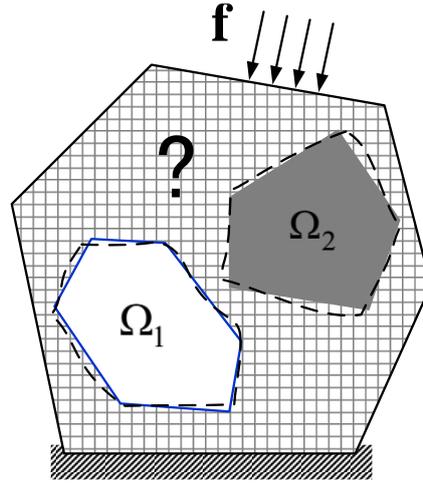
Figure 1. An illustrative structure system for shape-preserving design problem (Dashed lines indicates probable deformation for the loaded structure)

Multi-Point Shape Preserving Design

Local domains are concerned for shape preserving as shown in **Figure** 1. They may be void (e.g. a structural opening for process, feature or maintenance) or solid (e.g. a structural branch or a component), or even hybrid (e.g. parts or equipment). When it comes to situations like structural installing, connecting and assembling problems mostly based on the point locations, it is essential to have proper design of multi-point shape preserving i.e. coordinative displacement of control points. Therefore, we propose to define Artificial Week Elements (AWE) established with respect to the above mentioned control points. The local strain energies related to AWE are considered as additional constraints to suppressing the warping deformation.

Structural Deformation

The nodal displacement vector $\mathbf{u}_\Omega$ of the local domain Ω is composed by two components of rigid displacement vector $\mathbf{u}_{\Omega R}$ and warping deformation vector $\mathbf{u}_{\Omega W}$, i.e.

$$\mathbf{u}_\Omega = \mathbf{u}_{\Omega R} + \mathbf{u}_{\Omega W} \tag{1}$$

To achieve the structural shape preserving design necessitates suppression of the warping deformation. As a result, local strain energy is used to describe and constrain warping deformation quantitatively here. It is expressed as

$$C_\Omega = \frac{1}{2}\mathbf{u}_\Omega^{\mathrm{T}}\mathbf{K}_\Omega\mathbf{u}_\Omega \tag{2}$$

where $\mathbf{K}_\Omega$ is the local domain stiffness matrix.

Since no strain energy produced by rigid displacement, the above expression can be written as

$$C_\Omega = \frac{1}{2}\mathbf{u}_{\Omega W}^{\mathrm{T}}\mathbf{K}_\Omega\mathbf{u}_{\Omega W} \tag{3}$$

Theoretically, there would be no elastic warping deformation but only rigid body movement under a perfect shape preserving design where the local strain energy is 0. But practically the perfect effect is unobtainable. The constraint is given by a minor upper bound above zero, i.e. ε. The shape preserving design achieves a fairly well effect in permissible tolerance when the strain energy value satisfies

$$C_\Omega \leq \varepsilon \tag{4}$$

Artificial Week Elements

However, the shape preserving design will degenerate into an all-domain shape preserving when the elastic strain energy of the local domain is directly defined as a constraint function, which is an over constraint issue compared with multi-point shape preserving design. In this paper, Artificial Weak Elements (AWE) is proposed and established with respect to the shape preserving control points. The AWE nodal Degrees of Freedom (DOFs) are coupled to those of the control points. By calculating the AWE strain energy, the warping deformation of these multiple points can be measured.

Besides, to ensure the precision of structural analysis, the stiffness of additional AWE should be weak enough not to influence the structural mechanical properties. In this paper, the Poisson's ratio is set to a general value 0.3, and the elastic modulus is set to 1 Pa, which is much smaller than regular material.
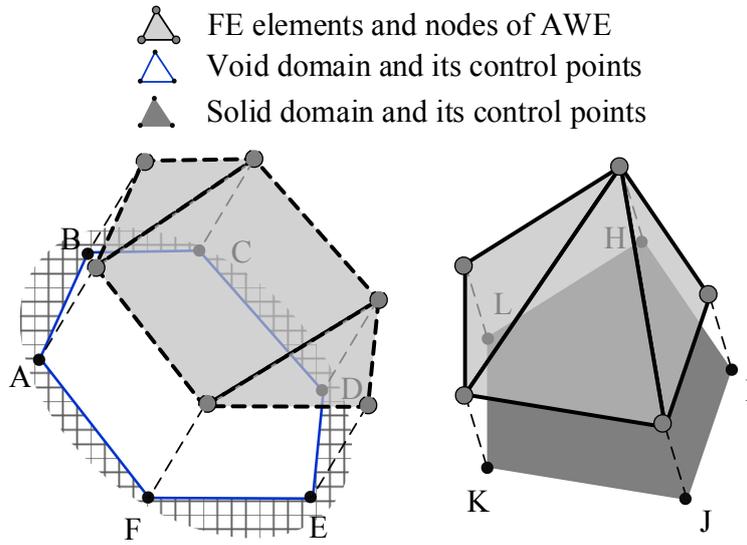
Figure 2. The definition of AWE

For the shape preserving design illustrated in **Figure** 1, AWE can be established as shown in **Figure** 2. The outline boundaries contain 11 control points, i.e. points A to L. Then 6 additional weak elements are created with the 11 points respectively. When the total structure is loaded, the AWE deform along with the control points. At this point, the shape preserving constraint can be defined as AWE strain energy constraint, i.e.

$$C_{\mathrm{AWE}} \leq \varepsilon \tag{5}$$

Therefore, the topology optimization with shape preserving design is formulated as

$$\text{find:} \quad \boldsymbol{\eta} = \left(\eta_1, \eta_2, ..., \eta_i, ..., \eta_n \right)$$

$$\text{min:} \quad C = \frac{1}{2} \mathbf{u}^{\mathrm{T}} \mathbf{K} \mathbf{u} \tag{6}$$

$$\text{s.t.:} \quad \mathbf{f} = \mathbf{K} \mathbf{u} \, ; \, V \leq V_0 \, ; \, C_{\mathrm{AWE}} \leq \varepsilon$$

In the above formulations, $\boldsymbol{\eta}$ is the vector of pseudo-density design variables, whose items' values vary from 0 to 1 describing material distribution in design domain. SIMP interpolation model (see Bendsøe and Sigmund 1999, Rozvany 2001) is used here with the penalty factor equals to 3. The global strain energy C is minimized as the object function. $\mathbf{K}$ is the global stiffness matrix. V is

the material volume and V_0 is its upper bound. ε is a given minor upper bound, whose value is relevant to specific structure and problem.

Sensitivity Analyses on Shape Preserving Constraint

The design sensitivity of the object function, i.e. the global strain energy with respect to the pseudo-densities is easily obtained and can be found in many references dealing with the topology optimization problems (e.g. Sigmund 2001), which will not be provided here.

We mainly concern the sensitivity of the constrained AWE strain energy. It can be expressed as

$$C_{\mathrm{AWE}} = \frac{1}{2}\mathbf{u}_{\mathrm{AWE}}^{\mathrm{T}}\mathbf{K}_{\mathrm{AWE}}\mathbf{u}_{\mathrm{AWE}} \tag{7}$$

$\mathbf{u}_{\mathrm{AWE}}$ is the displacement vector of control points, i.e. nodes of AWE. $\mathbf{K}_{\mathrm{AWE}}$ is the stiffness matrix of AWE

Derivative of the AWE strain energy is written as

$$\frac{\partial C_{\mathrm{AWE}}}{\partial \eta_i} = \frac{1}{2}\mathbf{u}_{\mathrm{AWE}}^{\mathrm{T}}\frac{\partial \mathbf{K}_{\mathrm{AWE}}}{\partial \eta_i}\mathbf{u}_{\mathrm{AWE}} + \mathbf{u}_{\mathrm{AWE}}^{\mathrm{T}}\mathbf{K}_{\mathrm{AWE}}\frac{\partial \mathbf{u}_{\mathrm{AWE}}}{\partial \eta_i}$$
$$= \mathbf{u}_{\mathrm{AWE}}^{\mathrm{T}}\mathbf{K}_{\mathrm{AWE}}\frac{\partial \mathbf{u}_{\mathrm{AWE}}}{\partial \eta_i} \tag{8}$$

where the stiffness matrix of AWE is independent from topology design variables η_i.

Here we define $\mathbf{u}_{\mathrm{AWE}} = \mathbf{T}_{\mathrm{AWE}}\mathbf{u}$, where $\mathbf{T}_{\mathrm{AWE}}$ is a constant matrix which converts the global displacement vector $\mathbf{u}$ to the local one $\mathbf{u}_{\mathrm{AWE}}$. Following the derivative of the equilibrium equation, we further have

$$\frac{\partial \mathbf{u}_{\mathrm{AWE}}}{\partial \eta_i} = \mathbf{T}_{\mathrm{AWE}}\frac{\partial \mathbf{u}}{\partial \eta_i} = \mathbf{T}_{\mathrm{AWE}}\mathbf{K}^{-1}\left(\frac{\partial \mathbf{f}}{\partial \eta_i} - \frac{\partial \mathbf{K}}{\partial \eta_i}\mathbf{u}\right) \tag{9}$$

Substituting the above equation into equation (8), it turns into

$$\frac{\partial C_{\mathrm{AWE}}}{\partial \eta_i} = \mathbf{u}_{\mathrm{AWE}}^{\mathrm{T}}\mathbf{K}_{\mathrm{AWE}}\mathbf{T}_{\mathrm{AWE}}\mathbf{K}^{-1}\left(\frac{\partial \mathbf{f}}{\partial \eta_i} - \frac{\partial \mathbf{K}}{\partial \eta_i}\mathbf{u}\right)$$
$$= \left(\boldsymbol{\lambda}^*\right)^{\mathrm{T}}\mathbf{K}^{-1}\left(\frac{\partial \mathbf{f}}{\partial \eta_i} - \frac{\partial \mathbf{K}}{\partial \eta_i}\mathbf{u}\right) \tag{10}$$

where we formulated a new vector $\boldsymbol{\lambda}^*$ calculated from the AWE displacements vector, stiffness matrix and the constant matrix, i.e. $\boldsymbol{\lambda}^* = \mathbf{u}_{\mathrm{AWE}}^{\mathrm{T}}\mathbf{K}_{\mathrm{AWE}}\mathbf{T}_{\mathrm{AWE}}$.

It is informed that $\boldsymbol{\lambda}^*$ is a column vector whose dimension is equal to total DOFs. After one additional finite element analysis by applying $\boldsymbol{\lambda}^*$ as an artificial load vector on the structure, we have

$$\boldsymbol{\lambda}^* = \mathbf{K}\mathbf{u}^*$$
$$\left(\boldsymbol{\lambda}^*\right)^{\mathrm{T}}\mathbf{K}^{-1} = \left(\mathbf{u}^*\right)^{\mathrm{T}} \tag{11}$$

Then the derivative of local elastic strain energy can be expressed as

$$\frac{\partial C_{\mathrm{AWE}}}{\partial \eta_i} = \left(\mathbf{u}^*\right)^{\mathrm{T}}\left(\frac{\partial \mathbf{f}}{\partial \eta_i} - \frac{\partial \mathbf{K}}{\partial \eta_i}\mathbf{u}\right) \tag{12}$$

The derivatives of the load vector and the stiffness matrix with respect to the pseudo-densities are easily obtained according to the SIMP interpolation model used in this paper. Typically, in the case of design independent loads, the derivative of the load vector will be zero, i.e. $\frac{\partial \mathbf{f}}{\partial \eta_i} = 0$.

Numerical Examples of Shape Preserving Design

L-shape Beam

Here we optimize an L-shape beam aimed at preserving the cutout configuration as shown in **Figure** 3. The top boundary is fixed and a single-point force of 100N is applied on the right corner. A frame with a particular non-design width is assigned around the cutout. Shape preserving control points are the four corners of the frame and the corresponding AWE is one quadrangle weak element linked to the control points A to D. Under the constraint of 40% material volume fraction, standard topology optimization design merely maximizing the overall structural stiffness is shown in **Figure** 4(a). Afterwards, without any other conditions changed, shape preserving design is shown in **Figure** 4(b), where ε equal to 2×10^{-15}J. The optimized strain energies of global structure, shape preserving frame domain and the AWE are listed in Table 1. The strain energy of AWE is decreased from 8.58×10^{-15}J to 2.00×10^{-15}J under the effect of shape preserving constraint, while the loss of global structure stiffness is less than 6%.

To have an obvious view of the shape preserving effect, a comparison of enlarged deformation of the frame is presented in **Figure** 5. The standard design generates a large warping deformation. On the contrary, the shape preserving design achieves a better deformation behavior where the frame corners' displacements was coordinated.
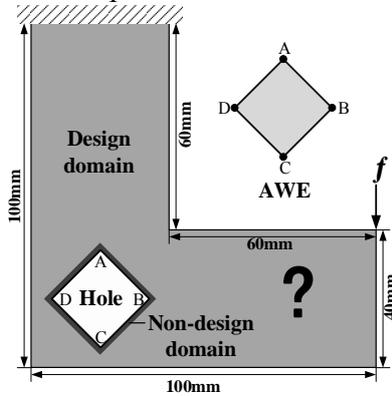
Figure 3. An L-shape beam with a quadrate cutout and its AWE

(a)   (b)
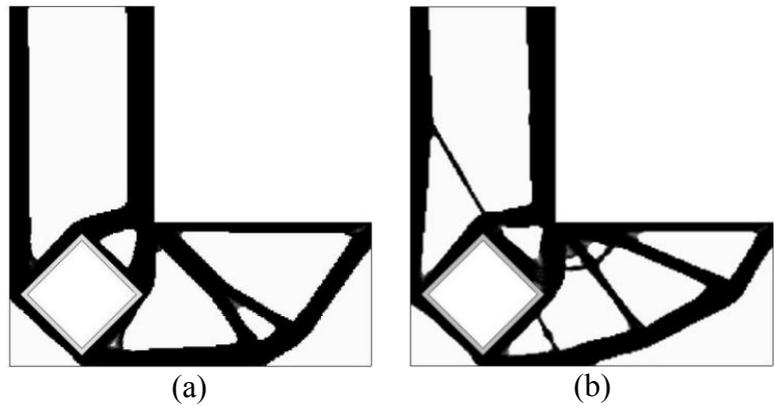
(a) Standard topology optimization
(b) Shape preserving design
Figure 4. Comparison of the L-shape beam designs

Table 1. Comparisons of strain energies of the optimized L-shape beam

Strain Energy	Global structure	Frame around the cutout	AWE
Standard topology optimization	1.17×10^{-4}J	8.57×10^{-6}J	8.58×10^{-15}J
Shape preserving design	1.24×10^{-4}J	4.21×10^{-6}J	2.00×10^{-15}J
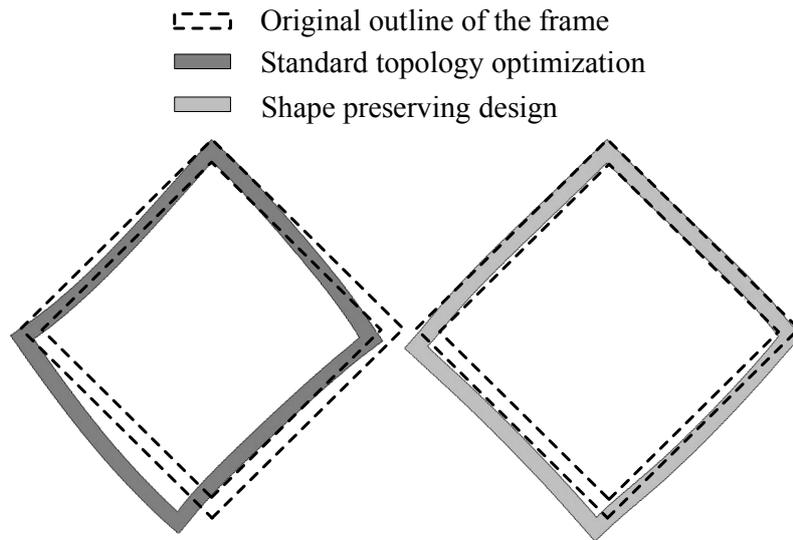
Figure 5. Comparison of enlarged deformations of the frame around the hole in optimized designs (Amplification factor 1.5×10^7)

Furthermore, the optimized results with different material volume are presented in **Figure** 6. C and C_{AWE} represent the optimized strain energy of global structure and AWE in standard topology design. C^* and C^*_{AWE} represent the optimized strain energy of global structure and AWE in shape preserving design.

For the standard topology optimization, the structure material is always distributed on the optimal load carrying path as a result of seeking maximum stiffness of global structure. Consequently, the standard design results always have smaller global strain energies. In shape preserving design, the local strain energy of AWE is much lower than the standard one with a little sacrifice on its global stiffness to satisfy local shape preserving constraint. This paradox between shape preserving constraint and global strain energy indicates that the final optimized design will be a compromise between global stiffness and local deformation.
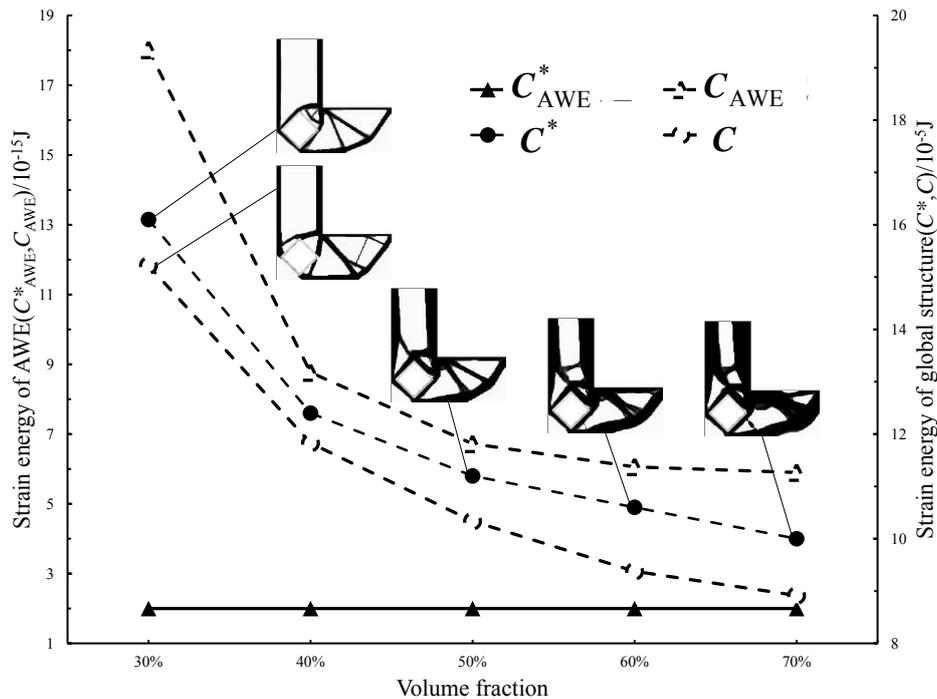
Figure 6. Optimized designs versus different volume fraction and their strain energy

Distortion of Load Carrying Path

For in-depth understanding of the paradox, further discussions on the upper bound of shape preserving constraint and its influence on the structural optimization design are discussed here.

Taking the L-shape beam design for example, we obtain optimized design in **Figure** 7 in turn via changing the value of ε with the rest conditions keeping identical. The optimized configurations change gradually as the value of ε decreases. When the shape preserving constraint is so strong, the structural load carrying path will be distortional (e.g. 11th and 12th result) with unsatisfied large sacrifice of global stiffness. In these cases, regular structural design cannot meet the requirement of shape preserving constraints. The topology optimization is forced to separate the shape preserving domain from the load carrying path to obtain an approximate rigid deformation. Such result is mathematically reasonable but loses actual physical significance and engineering value in optimization design.
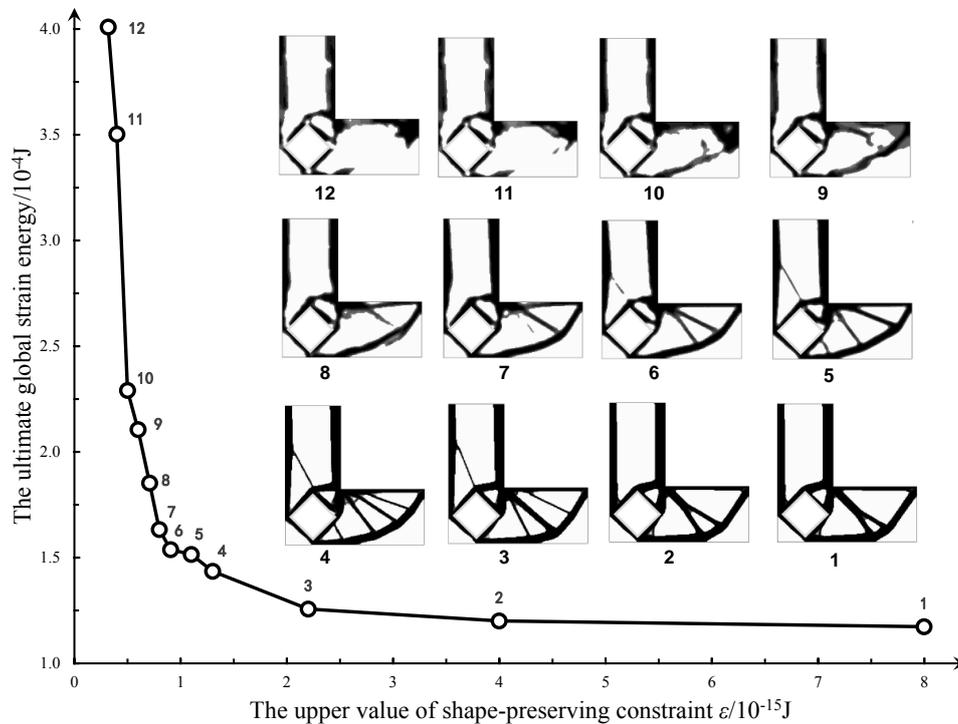
Figure 7. The global strain energy and corresponding optimized design results versus different constraint values of ε

Accordingly, the upper bound of shape preserving constraint should be appropriately chosen to avoid phenomena of load carrying path distortion. Meanwhile, researchers are not only to solve a mathematical model but also to account for more practical problems into consideration, which is one of the key difficulties in optimization design for engineering structures.

Shape Preserving Design for Windshields

Consider now an airframe shown in **Figure** 8. The front fuselage is connected to the center one at its rear side. The whole fuselage bears aerodynamic loads. Warping deformations of windshields need to be avoided not to cause the glasses fracture. Here, AWE is defined as illustrated in **Figure** 8. The control points of each windshield contain four corners and four midpoints of the boundaries as well. With the airframe's layout as topology optimization design subject, two material distribution results of skin reinforcement from standard design and shape preserving design are presented in **Figure** 9. The value of the shape preserving constraint ε is set as 0.02J.
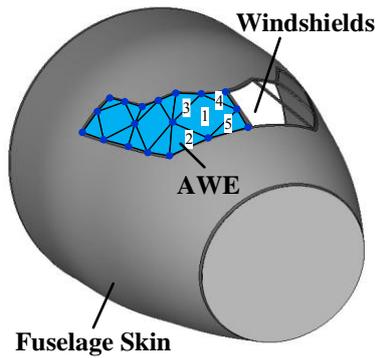
Figure 8. Illustration of shape preserving design of windshields
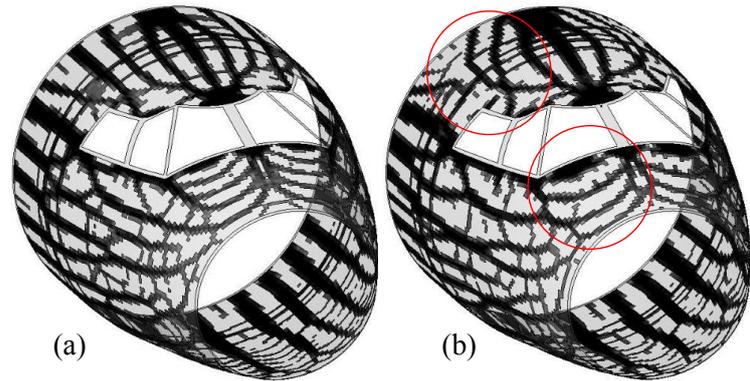
(a) Standard topology optimization design
(b) Shape preserving design
Figure 9. Comparison of material distributions of skin reinforcement

Referring to the optimized designs with the same volume fractions in **Figure** 9, we can distinguish that the material distributed around the windshields increases in the shape preserving design. Therefore, the prescribed local domain is strengthened and the warping deformation is suppressed. Additionally in the weak loaded area between windshields and center fuselage, the shape preserving design modifies the load carrying path to offset the warping deformation in the windshields. The detailed data of shape preserving design and standard stiffness design is listed and compared in Table 2. Although there is a 5% sacrifice on the stiffness of global structure, the shape preserving design has improved the effect of shape preserving for 4 times better than the standard one. Thus, the effectiveness of shape preserving topology optimization design is further demonstrated, which possesses a good perspective in practical structure design applications.

Table 2. Comparisons of strain energies of optimized designs

Strain energy	The whole fuselage	The AWE of windshields
Standard optimization design	6893 J	0.092 J
Shape-preserving optimization design	7263 J	0.020 J

Conclusions

We proposed an extended structural topology optimization method with multi-point shape preserving constraint in this paper. The shape preserving constraint of local domain is constructed by the strain energy of Artificial Weak Elements (AWE). Compared with the standard topology optimization design maximizing structural stiffness, this formulation have evidently shown that the coordination of multi-point displacements and the effect of shape preserving can be successfully achieved. Further numerical results are analyzed to show the influence of shape preserving constraint on the optimized design pattern and the entire performance of structure. Besides, the design distortion due to improper definition of the shape preserving constraint is revealed and studied in this paper.

Acknowledgment

References

Barrett R. T. (1992) Fastener Design Manual. *NASA Reference Publication* 1228

Bendsøe M. P. Sigmund O. (1999) Material interpolation schemes in topology optimization. *Archive of Applied Mechanics*, **69(9-10)**:635~654.

Deaton J. D., Grandhi R. V. (2014) A survey of structural and multidisciplinary continuum topology optimization: post 2000. *Structural and Multidisciplinary Optimization* **49**:1-38

Guo X., Cheng G. D. (2010) Recent development in structural design and optimization. *Acta Mechanica Sinica* **26**:807-823

Liu S., An X., Jia H. (2008) Topology optimization of beam cross-section considering warping deformation. *Structural and Multidisciplinary Optimization* **35**:403-411

Niu M. C. Y. (1988) Airframe Structural Design: Practical Design Information and Data on Aircraft Structures, *Technical Book Co*

Qiao H. T., Liu S. T. (2013) Topology optimization by minimizing the geometric average displacement. *Engineering Optimization* **45**:1-18

Rong J., Yi J. (2010) A structural topological optimization method for multi-displacement constraints and any initial topology configuration. *Acta Mechanica Sinica* **26**:735-744

Rozvany G. I. N. (2001) Aims, scope, methods, history and unified terminology of computer-aided topology optimization in structural mechanics. *Structural and Multidisciplinary Optimization*, **21(2)**: 90-108.

Sigmund O., Maute K. (2013) Topology optimization approaches. *Structural and Multidisciplinary Optimization* **48**:1031-1055

Sigmund O. (2001) A 99 line topology optimization code written in Matlab[J]. *Structural and Multidisciplinary Optimization*, **21(2)**: 120-127.

Stanford B., Beran P. (2012) Optimal Compliant Flapping Mechanism Topologies With Multiple Load Cases. *Journal Of Mechanical Design* 134

Stanford B., Beran P., Kobayashi M. (2013) Simultaneous topology optimization of membrane wings and their compliant flapping mechanisms. *Aiaa J* **51**:1431-1441

Wang B. Z. (2000) Aircraft Design Manual 10: Structural Design, *Aviation Industry Press*, Beijing

Wang M. Y. , Chen S. K. , Wang X. M., Mel Y. L. (2005) Design of multimaterial compliant mechanisms using level-set methods. *J Mech Design* **127**:941-956

Interfacial fracture of polymer foam-metal composites at micro-scale using finite element analysis

*R. Das[1], N. Navaranjan[2], P. Browne[1], G. Riley[1], and C. Nguyen[3]

[1]Centre for Advanced Composite Materials, Department of Mechanical Engineering, University of Auckland, Auckland, New Zealand.
[2]Scion Research, Rotorua, New Zealand
[3]Product Accelerator, University of Auckland, Auckland, New Zealand.

*Corresponding author r.das@auckland.ac.nz

Abstract

Foam-metal composites are being increasingly used in a variety of applications. One important aspect in the structural integrity of foam-metal interface is the ability to resist failure around the interface whilst ensuring required load bearing capacity. This study investigated the mechanical and failure behaviour at the interface region at micro scale. The foam-metal composite consisted of polyurethane foam directly adhered to a galvanised steel face sheet. Optical, scanning electron and atomic force microscopy were used to examine the interface geometry and to obtain a realistic surface profile for use in a finite element (FE) model. Finite element analysis (FEA) was used to study the effects of different interfacial roughness profiles on mechanical interlocking and modes of failure, which are directly related to interfacial strength. A set of finite element models of idealised surface pairs of different geometries and dimensions were developed based on the microscopic observations at the foam-metal interface. The finite element modelling results show that the micro-scale roughness profile at the foam-metal interface causes mechanical interlocking and affects the stress field at the scale of the interface surface roughness, which consequently governs the specific failure mode and the relative proportion of the cohesive to adhesive failure in the interface region for a given foam-metal interface. It was found that the aspect ratio (relative width and height) and width ratio (relative spacing) of roughness elements have a significant effect on the stresses and deformations produced at the interface and consequently control the modes (cohesive or adhesive) of failure.

Keywords: Numerical modelling, foam, interface, fracture, cohesive fracture, adhesive fracture.

Introduction

Sandwich composites consisting of polymer foam and metal face sheets have many advantages for structural applications. The notable benefits are light weight, high bending stiffness and strength [Grujicic et al. (2008)]. The bond between the foam and metal must remain intact to ensure the structural integrity of the composite. So optimising the adhesion strength of the interface is crucial to the performance of the composite. To effectively achieve this, it is essential to understand the mechanisms of adhesion and the effects of these on the strength of the adhesive bond [Kim et al. (2010)].

Characteristics of the interface of a solid polymer and metal have been extensively investigated. The main factors affecting the interfacial strength are chemical/physico-chemical (e.g. ionic or covalent bonds or van der Walls force) and physical (e.g. mechanical interlocking) interactions [Buehler (2008)]. Chemical interactions are related to the primary and secondary bond formations [Ho (1989); Grujicic et al. (2009)]. Mechanical interlocking is the interaction between the two material surfaces due to geometric effects [Noijen et al. (2009)] and plays a dominant role in interface bonding. A typical interlocking feature originates from the surface roughness of the interface and generally occurs at the microscopic scale. Specifically in the case of a polymer adhered directly to a metal surface, mechanical interlocking and absorption are the most significant mechanisms that contribute to the strength of the interface [Kim (2003); Grujicic et al. (2009); Kim et al. (2010); Ochoa-Putman and Vaidya (2011)]. The friction at the polymer-metal interface along with the polymer stiffness also contributes to the mechanical behaviour of the interface [Ochoa-Putman and Vaidya (2011)].

The effect of altering the surface roughness of steel at a micro-scale along with chemical treatment of the steel surface on the behaviour and strength of a steel-polymer interface [Ochoa-Putman and

Vaidya (2011)] showed that the strength of the steel-polymer interface increased as the surface roughness increased. It is suggested that the polymer-metal interfaces usually fail due to a combination of both interfacial adhesive failure between the polymer and the metal and cohesive failure of the polymer due to cracking [Yao and Qu (2002); Kim et al. (2010); Ochoa-Putman and Vaidya (2011)]. The relative strengths against cohesive and adhesive failures determine the resultant interfacial strength. For some material systems, cohesive failure typically requires more energy than that of adhesive failure; therefore interfacial strength could be improved by increasing the proportion of cohesive failure compared to adhesive one, as studied [Yao and Qu (2002)]. Altering surface roughness has been shown to result in an increase in cohesive failure and a reduction in adhesive failure [Yao and Qu (2002)], thus can increase the interfacial strength. A relationship exists between adhesive failure and the non-dimensional roughness $\bar{R}/\lambda$, where $\bar{R}$ is the mean half-depth of the roughness and λ is the mean distance from peak to trough of the roughness (Figure 1).
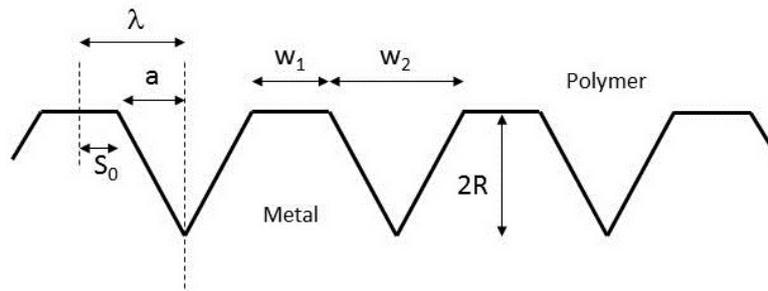
Figure 1. Idealised profile representing surface roughness of polymer-metal interface (adapted from [Yao and Qu (2002)])

Kim et al. [Kim et al. (2010)] also showed via a similar study that increasing the ratio of cohesive to adhesive failure along the interface is an effective way to improve the interfacial strength of a metal-polymer composite. They roughened the surface of steel with micro-line patterns and investigated the effect of varying the roughness dimensions; depth R, widths w_1 and w_2 and width ratio w_1/w_2 (shown in Figure 1). However, changing the roughness depth R was found to have no effect on the interfacial fracture toughness. Increasing the ratio w_1/w_2 resulted in a smaller fraction of area failing purely due to adhesive failure along the interface and a larger fraction of area failing due to cohesive failure of the polymer. Hence, increasing the ratio of cohesive to adhesive failure can increase the interfacial strength of several metal-polymer composites.

Numerical modelling has been successfully used to characterise polymer-metal interfaces [Yao and Qu (2002); Noijen et al. (2009)]. The finite element method (FEM) was used by Yao and Qu [Yao and Qu (2002)] to predict the energy release rate (ERR) of adhesive and cohesive cracks at different positions along a typical surface roughness profile of a metal-polymer interface. They observed that a crack along the interface propagated into the polyurethane (PU) foam when the ratio of adhesive to cohesive energy release rates (G_R) reached a critical value.

In the numerical study by Noijen et al. [Noijen et al. (2009)], the crack was assumed to be first formed on the flat surface of the metal due to low adhesive strength, and propagated along the interface until the ERR condition along the interface and through the polymer (affected by the roughness geometry at the interface) were met, which changed the direction of crack propagation deflecting into the polymer. This was however unable to account for the relative dominance between the adhesive and cohesive failure, so an improved numerical model was deemed necessary. This work indicated that the numerical modelling can predict the location at the polymer-metal interface where the crack propagation will deviate from the interface into the polymer, which in turn enables determination of the ratio of cohesive to adhesive failure [Noijen et al. (2009)].

In this paper, we extend the previous studies on polymer-metal interface to the material system of foam-metal sandwich composite interface, namely galvanised (zinc coated) steel and polyurethane foam composite interface. It is known that polymers can adhere well to zinc coatings [Kim (2003)]. However, the key difference of foam-metal interface from that of a homogenous polymer-metal

interface is the porous and cellular structure of the foam, which affects the localised adhesive and cohesive failure mechanisms and the resulting fracture behaviour. Moreover, variability, such as temperature, polymer distribution, and poor wetting during the casting of polyurethane foam onto zinc coated steel can result in different adhesive strengths in different regions of the foam-metal interface [Kim (2003)].

Most importantly, the bulk failure behaviour at the interfacial region is manifested by micro-scale deformation and fracture patterns. To address this, the present work investigated the interface of the specific foam-metal sandwich composite made of polyurethane foam and galvanised steel face sheets at microscopic scale using experimental observations and finite element (FE) analysis. The focus was to understand how the geometry of micro-scale surface roughness influences the nature of failure at the interfacial region of the foam-metal composite. The stress-strain distributions in the micro-scale roughness profile were analysed. In the FE analyses, idealised geometries of the interface representing different surface roughness profiles were created, and the failure modes were predicted under the tensile and shear loading conditions. In the experimental study, the foam-metal sandwich samples were fabricated without using any adhesive for bonding so as to avoid any chemical effect of external adhesives. The metal-foam interface was imaged using optical and scanning electron microscopes. The surface roughness was characterised using an atomic force microscope. Subsequently, the effect of changing surface roughness parameters (i.e. aspect ratio and width ratio) on the crack propagation patterns and failure modes, and the resulting interaction between adhesive and cohesive failures were investigated using finite element analysis.

Materials and Methods

Material Systems

Polyurethane foam is porous, and has a cellular structure with voids. The properties of the foam can vary greatly depending on the type and proportion of the reactants used. The polyurethane foam used in this study was produced by mixing Endurathane GP38 polyol blend with Endurathane 5005 isocyanate, supplied by New Zealand Polymer Group Ltd. The stress-strain curves of the foam in tension and compression are shown in Figure 2.
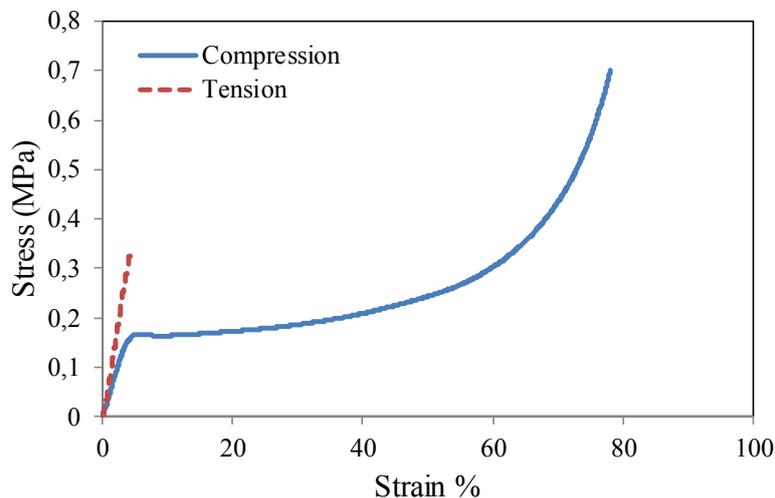
Figure 2. Polyurethane foam stress strain curves showing the difference in behaviour when subjected to tension and compression

A commercially available G550 galvanised high strength structural steel was used for the composite. It had a thickness of 0.75 mm. The material properties are given in Table 1. The steel surface is covered by a zinc rich layer as a result of the galvanisation. The foam contact is therefore with this zinc layer, and not with the steel. The zinc layer is significantly stiffer and stronger than the polyurethane foam (Table 1).

Table 1. Material properties of the constituents of the foam-metal composite

Material	Modulus of Elasticity	Yield Strength	Poisson's Ratio
Polyurethane foam [Randall and Lee (2002)]	7.5 GPa	0.34 MPa	0.40
Zinc [Davis (2014)]	83 GPa	160 MPa	0.25
G550 steel [Davis (2014)]	200 GPa	550 MPa	0.3

Sample Preparation

Polyurethane foam-G550 steel sandwich composite was fabricated in a controlled laboratory condition. Composite samples that were subjected to compression tests were failed by local buckling, and foam and metal were delaminated along the interface. The foam-metal specimens were cut into approximately 10 × 10 mm sections from the failed region of the samples to observe the interface failure. These specimens were used for surface roughness analysis and imaging foam cell sections. The foam-metal specimens were then moulded in an epoxy resin without causing any damage to the delaminated interface and foam (see Figure 3a). The section of each specimen was then ground and polished to obtain a smooth surface for microscopic examination. Sections of 10 × 10 mm specimens were also mounted flat using double-sided carbon tape (Figure 3b) to observe the top and bottom surfaces of the failed composite.

(a)   (b)
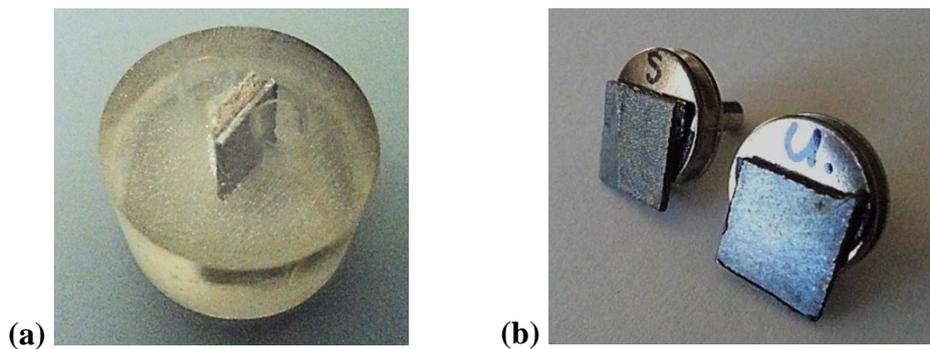
Figure 3. (a) Specimen cross-section of foam-metal interface prepared for microscopic examination, (b) Specimens showing the top and bottom surfaces of foam-metal composite where it has failed along the interface (top specimen shows failure surface, and the bottom specimen is used as the control)

Microscopic Examination

An Olympus MX6B microscope was used to examine the prepared specimens. Dark field lighting was used, as it provides a clearer image than that of bright field lighting. The entire length of the interface was examined at 100, 200 and 500 magnification levels. Images were taken at regular intervals along the interface. Surface roughness of zinc coating was analysed using an Atomic Force Microscope.

The cross-section of the foam-metal interface, the failed surface, and the surface of steel (without foam) were observed using a Scanning Electron Microscope (SEM) at a range of magnification from 100× to 24000×. The specimens were coated with a thin platinum layer to avoid surface charge accumulation and to improve the image quality. Both the back scatter electron detector (for improved materials distinction) and secondary electron detector (for improved topography) were used to observe the specimens. Figure 4 shows a typical optical microscope image of the foam-metal interface.

Figure 4. A typical optical microscope image of the foam-metal interface (using Dark field and 100× total magnification) showing steel (black), zinc (silver), and polyurethane foam
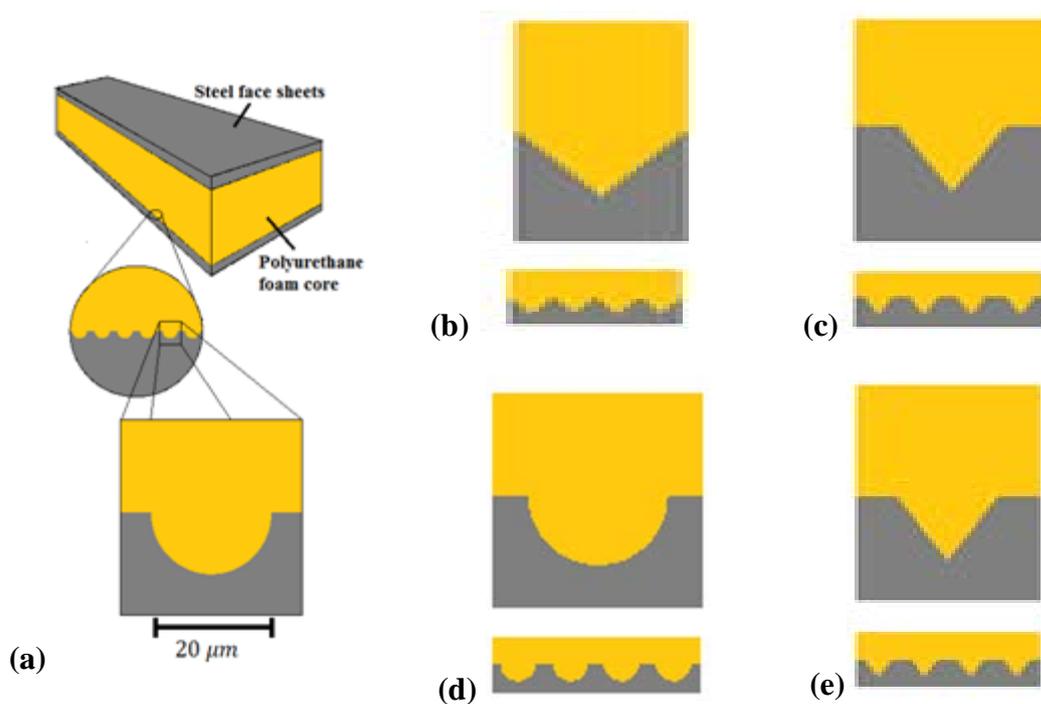
Figure 5. Division of the foam-metal composite into individual roughness elements: (a) Foam-metal sandwich composite, micro-scale view of the foam-metal interface, and zoomed view of a typical roughness profile, (b) jagged profile, (c) triangular profile, (d) semi-circular profile, (e) Filleted triangular profile

Finite Element Modelling of the Interface

The foam-metal interface was modelled using FEM to understand the failure mechanisms and investigate how different roughness parameters affect the deformation behaviour and failure modes. An idealised surface roughness profile is an accurate assumption for modelling an induced roughness, especially if the technique used to induce the roughness is accurate and consistent. Even for modelling a random surface roughness (i.e. a naturally rough surface) an idealised surface roughness profile can be an accurate assumption provided it is based upon average dimensions and

a profile similar to that of the actual roughness. Idealised micro-scale surface roughness profile was modelled by dividing the foam-metal interface into individual repeating elements (RE), as shown in Figure 5. Each RE represented a roughness profile with a part of the metal and foam.

Materials Modelling

A two dimensional FE model of RE was developed and analysed using the software ABAQUS [ABAQUS (2014)]. The polyurethane foam was modelled as a homogenous and isotropic solid material, as it was identified through optical microscopy that the material at the interface region was solid polyurethane rather than cellular foam (Figure 5a). The thickness of the solid film varied between 5-50 μm. The failed interface showed that the thin solid film perfectly adhered to the metal surface (see Figure 4). Four different idealised surface roughness profiles were modelled. These are referred to as the jagged, triangular, filleted triangular and semi-circular profiles, as shown in Figure 5b-e, respectively.

Both zinc and steel are considerably stiffer and stronger than polyurethane so that it is unlikely that either will fail before polyurethane does. Hence the metal component of the interface has been modelled as analytically rigid. To simplify the analysis we have chosen to model polyurethane as behaving the same in compression as it does in tension. This is a reasonable assumption as at the micro scale solid polyurethane attached to the zinc coating is unaffected by voids that are responsible for the difference in behaviour in tension and compression at the macro scale. Figure 6 shows stress-strain curve of the polyurethane foam used in the FEA analysis.
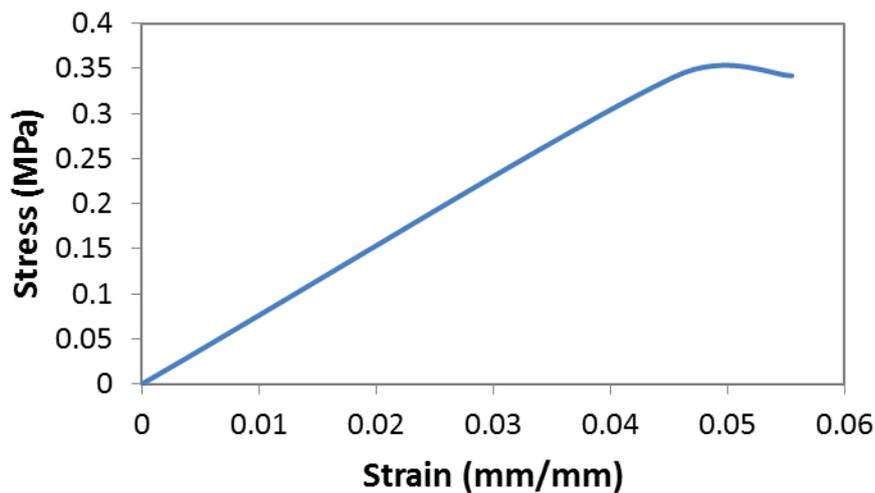
Figure 6. Polyurethane stress-strain curve used in FEA

Loading and Boundary Conditions

The foam metal interface was analysed under two loading conditions, tension and shear loadings. So each surface roughness profile was modelled under both tensile and shear loading. The boundary conditions imposed for the tensile and shear load cases are shown in Figure 7. Both the boundary conditions included specified displacement and symmetry conditions. Each roughness element was symmetric about the vertical axis (Figure 7a). Hence, a half of the single element was modelled applying the symmetric boundary condition about the vertical axis (red lines in Figure 7b) for tensile load case. Since the foam was perfectly adhered to the zinc layer of the galvanised steel, all degrees of freedom were constrained at the foam-zinc interface in the finite element model for both load cases. This is highlighted in yellow in Figure 7b.

(a)

(b)
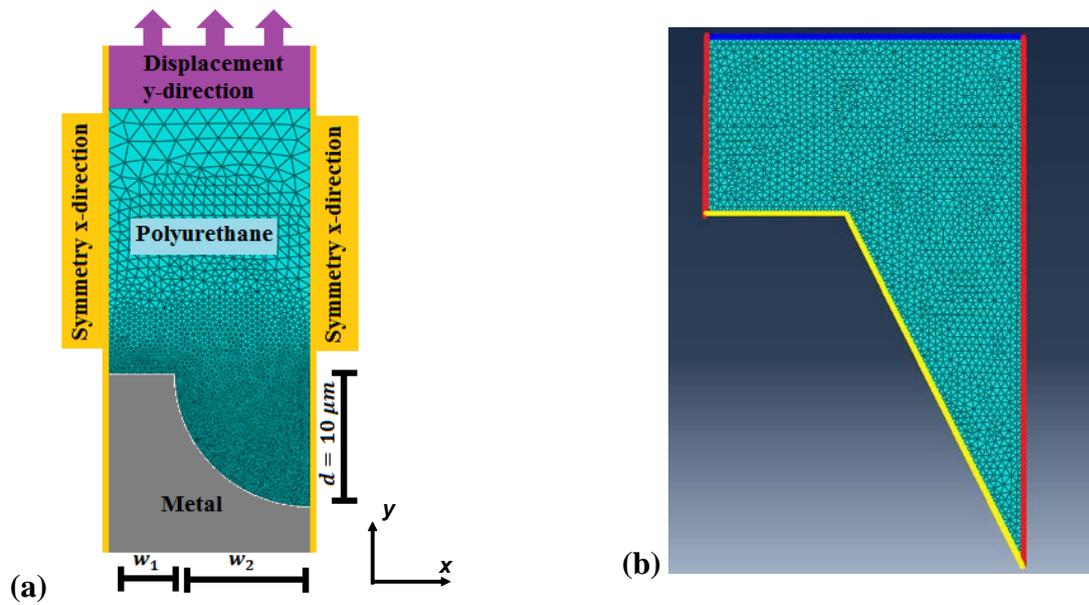
Figure 7. Tensile boundary conditions: (a) A typical foam-metal interface model showing loads and boundary conditions for tensile loading, (b) Different boundary conditions: yellow - fixed, red - vertical (y) symmetry, blue - specified vertical displacement

For the tensile loading, a specified displacement boundary condition of 0.5 μm was applied to the top surface of the roughness element in the vertical direction (y-direction), which is highlighted by the blue line in Figure 7b. The direction of shear load was parallel to the horizontal interface, and as a result asymmetric loading was generated across the roughness element, implying that the symmetry boundary condition could no longer be used. Moreover, the stresses and strains generated due to shear loading were not symmetrical between adjacent roughness elements. So symmetry boundary conditions could not be used at the edge of the roughness element to compensate for the edge of the roughness element not being at a free end of the interface (see Figure 8). To account for this, the geometry used in the FE model for shear loading was a section of the metal foam interface consisting of three roughness elements, and the behaviour of the middle roughness element was evaluated. The top surface of the roughness element is displaced in the horizontal (x) direction by 1.5 μm.

(a)

(b)
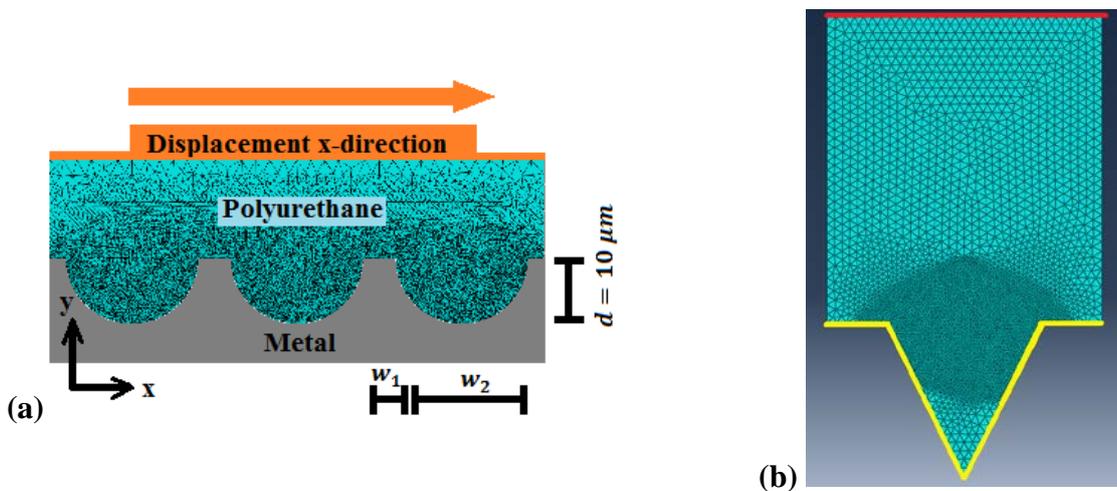
Figure 8. Shear boundary conditions: (a) A typical foam-metal interface model consisting of three roughness elements showing loads and boundary conditions for shear loading, (b) Different boundary conditions: yellow - fixed, red - specified horizontal (x) displacement

Meshing and Failure Modelling

Linear 3-node triangular plain strain elements were used for all the models. The mesh was biased and was locally refined to generate fine meshes at the metal-foam interface and around the corners of the profiles in order to capture the sharp stress variation along the interface and around the corners. A finite element mesh convergence study was performed to ensure that the mesh used in the models was of sufficient resolution so as to accurately predict the stress, strain and damage evolution in the interface region.

A continuum damage model was used to model crack propagation, adopted from [Neilsen et al. (1995)]. Each element had a scalar 'damage' parameter D that measures its load carrying capacity. The scalar damage parameter D is used to characterise the volume-averaged micro-fracture of the volume of material represented by each element. The damage parameter is calculated based on the principal stresses of the elements. It is used to inhibit the transmission of tensile stress between elements. The damage parameter lies between 0 and 1. Material with $D = 0$ is undamaged and is able to transmit the full tensile load, whereas material with $D = 1$ is fully damaged and cannot transmit any tensile load, thus creating a partial macro crack. Connected macro-cracks or contiguous cracked material across a body leads to fracture. A failure criterion was used, and a material stiffness degradation model was implemented. If an element met the failure criterion, its stiffness was reduced by scaling with (1-D).

Microscopic Characterisation of the Interface

Figure 9 shows a surface roughness image obtained from the surface area of zinc using AFM. Figure 10 shows the cellular structure of polyurethane foam obtained from optical microscopy. At a magnification of 500, the features, such as foam pores, contact between polyurethane and zinc at the interface, distinct regions of steel, zinc and polyurethane as well as the irregularities in the zinc galvanising layer, can be observed. Figure 11 shows the solid polyurethane film (top grey) that covers the majority of the metal surface at a magnification of 8000 from SEM. The film thickness generally varies between 5-50 μm although there are a few locations along the surface which are not covered by the film.
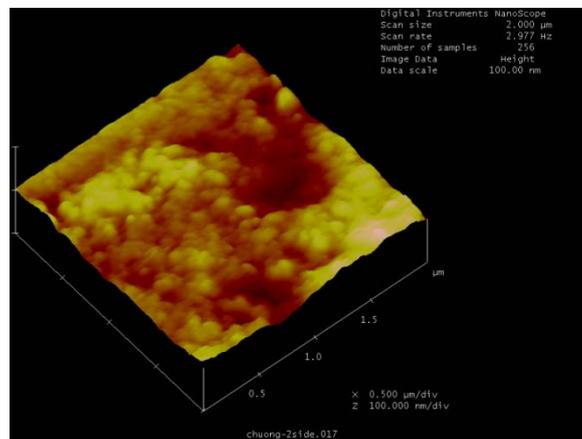
Figure 9. Roughness profile of zinc surface of the galvanised steel
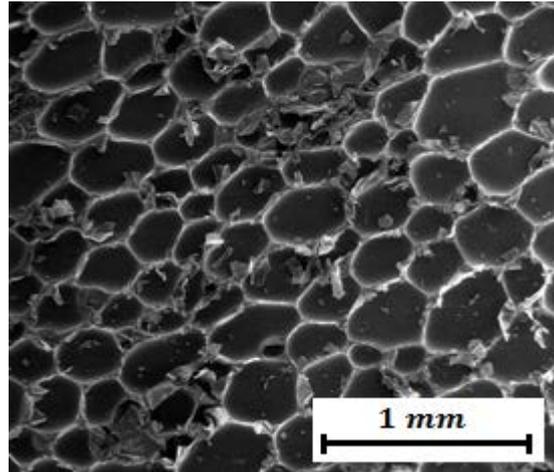
Figure 10. Cellular structure of polyurethane foam obtained from optical microscope
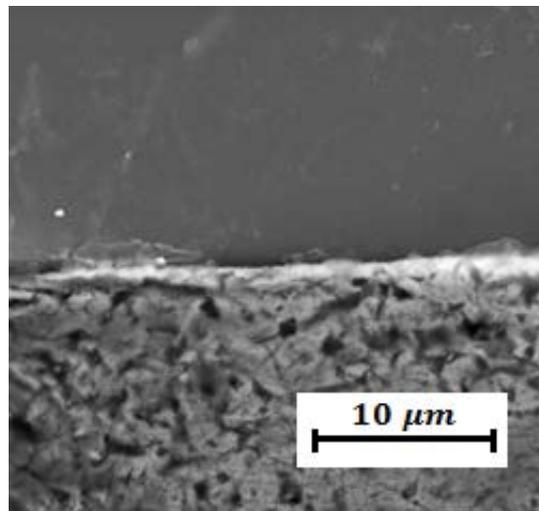
Figure 11. SEM image showing solid polyurethane (without pores/voids) layer covering the majority of the metal surface (observed using 8000 × total magnification)

The observations suggested that the foam-metal composite did not predominantly fail along the interface adhesively, and rather it failed near the interface by cohesive manner. The top surface specimens failed much closer to the interface than the cross-sectional specimens, i.e. in the cell walls of the cells in the vicinity of the metal surface. The stresses in the cellular area of the polyurethane foam where voids are present are considerably higher than those in the solid foam present at the foam-metal interface. This leads to the trend in cohesive failure predominantly in the cell wall as opposed to solid polyurethane .As a result of this, crack propagation due to cohesive failure would not be in the immediate vicinity of the interface. A transition from cohesive failure to adhesive failure would be unlikely due to the small area of the cell walls compared to the solid layer of polyurethane attached to the metal layer. This explains the large areas of continuous cohesive and adhesive failure zones. As a result, when the adhesive strength of the interface exceeds the cohesive strength of the polyurethane foam at a macroscopic scale (including the effect of cellular structure), the mode of failure would be exclusively cohesive in the cell walls immediately next to the interface. This limits the strength of the interface to the cohesive strength of the macroscopic polyurethane foam. If the cohesive strength of the foam on a macroscopic scale exceeds the adhesive strength of the interface, the strength of the interface would then be limited to the strength of adhesion. This study focuses on the failure that is essentially confined to the 'interface region' – the region from the metal surface up to the top of the solid polyurethane foam layer. Hence for the purpose of studying the failure modes within this interface zone, the cellular structure of the foam can be ignored.

Stress Analysis using FEM

As we focused on the solid polyurethane foam near the interface, the bulk properties of the foam was used for finite element analysis. Increasing the aspect ratio (ratio of width to height) roughness elements decreases stress concentration and reduces the likelihood of cohesive failure. Shear of the interface generally results in more adhesive failure than compared to tension (this may vary depending on material properties and adhesive strength). Typical maximum principle stress and strain distributions in a roughness element are provided in Figure 12.

(a)   (b)
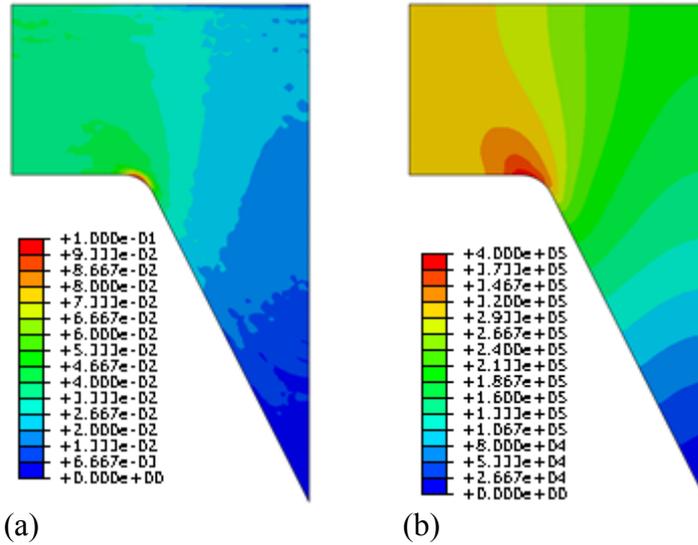
Figure 12. (a) Maximum principle strain distribution in polyurethane foam at point of yield, (b) distribution of maximum in plane stress in polyurethane at point of yield (MPa)

Failure is initiated at points of stress concentration. The stress distribution determines the location, where polyurethane foam will first yield initiating cohesive failure. The distribution and concentration of strains at the interface between polyurethane and metal determines where adhesive failure will occur. The characteristic shape of a roughness element determines the distribution and concentration of stress and strain at the interface between the polyurethane foam and the metal. Figure 13 shows that stress distribution for roughness profiles of different shapes and how it influences the stress field and concentrations. We discuss below how a typical roughness element behaves under tensile and shear loading prior to damage or crack initiation.

(a)   (b)
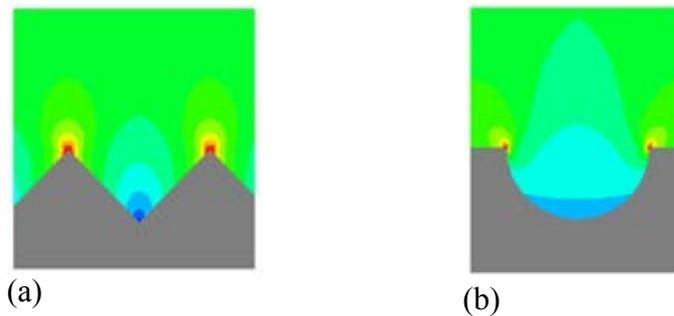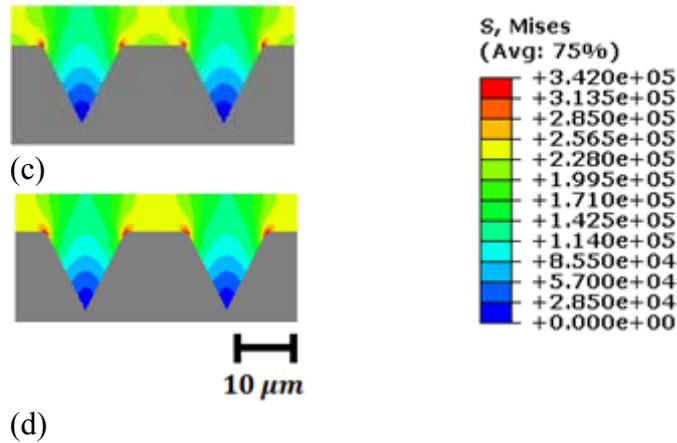
Figure 13. von Mises Stress distribution (MPa) in the roughness profiles at point of failure for various shapes of the roughness elements; (a) Jagged, (b) Semi-circular, (c) Triangular, (d) Filleted triangular

Effect of Roughness Parameters on the Failure Mode

Failure is initiated at points of stress concentration. Failures which propagate along the polyurethane-steel interface will be adhesive failure and failures which propagate through the polyurethane foam will be cohesive failure. The relative magnitudes of the cohesive strength of foams and the adhesive strength of an interface govern the failure mode for a given geometry of a roughness profile and loading conditions. However, these strength values depend on the foam type, structure, processing conditions, and method of adherence to the substrate; so the exact relative proportion of cohesive and adhesive failures will depend on the specific material system. For many cases, when the adhesive strength of the interface is generally weaker than the cohesive strengths of the materials which comprise that interface, adhesive failures typically may occur before cohesive failures for foam-metal composites as cohesive failures require more energy than adhesive failures. In this case the purpose of inducing surface roughness at the foam-metal interface will be to cause cohesive failures, as a greater ratio of cohesive failure to adhesive failure will result in a stronger interface. In this study, the adhesive strength of the polyurethane steel interface and the cohesive strength of polyurethane foam were assumed to be equal, as the objective was to analysis how the geometry of the roughness profile affects the failure mechanics under different loading conditions for an interfacial region with equal strengths.

It has been identified that modifying certain roughness parameters (e.g. aspect ratio and width ratio) can improve interfacial strength by increasing the ratio of cohesive to adhesive failure around the interface ([Kim (2003); Kim et al. (2010)]). In this study we explore the effect of aspect ratio, width ratio and shape of the roughness profile on the strength and failure mechanism of foam-metal interfaces. The aspect ratio is defined as the ratio of roughness width (w) to roughness depth (d) as shown Figure 14a, and the width ratio is defined as the ratio of widths w_1 to w_2 as shown in Figure 14b. Both of these parameters have been shown to affect the strength of polymer-metal interfaces. Next the failure mechanics of roughness profiles for a range of different aspect ratios (1:1 to 4:1) and width ratios (1:1 to 8:1) will be evaluated using FEA in order to determine how they affect the interfacial failure mode.
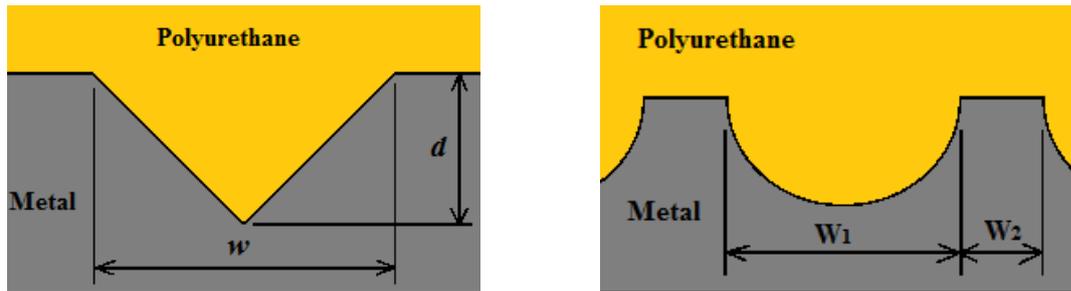
(a) Aspect ratio, w:d (b) Width ratio, w_1:w_2

Figure 14. Definition of aspect ratio (left) and width ratio (right)

Effect of Aspect Ratio

A key parameter of micro-scale roughness is the ratio of the width to the depth of the roughness elements. Increasing the aspect ratio in general reduces stress concentrations within the foam making cohesive failure less likely. Path of crack propagation is indicated by the completely damaged material, as shown by red lines in Figure 15, Figure 16, and Figure 17. Elements in red indicate where the failure criterion was satisfied, and the stiffness of corresponding elements was degraded.

Tensile loading

We considered 'jagged' roughness profiles of three different aspect ratios, viz: 1:1, 2:1, and 4:1, and analysed the crack propagation and failure modes in each case under tension (Figure 15). The crack initiated at the corners of the base of the roughness element for all the aspect ratios considered. For the 1:1 aspect ratio profile, the crack propagated exclusively through the polyurethane material between the two base corners, indicating a pure cohesive failure as shown in Figure 15a.

An increased aspect ratio of 2:1 resulted in a combination of mixed adhesive and cohesive failures (Figure 15b). The crack propagated adhesively along the interface before a cohesive failure is initiated which propagated into the polyurethane foam. For this 2:1 aspect ratio roughness profile, the cracks propagated approximately half-way down the left side of the interface of the roughness element, and then traversed across it horizontally through the polyurethane foam up to the centre, where the cracks from both the sides met. This crack pathway resulted in a failure mode which was about 50% cohesive.

For the 4:1 aspect ratio profile, the crack propagated a small distance down the interface of the roughness element, and then traversed horizontally across it through the polyurethane foam cohesively (Figure 15c). Some secondary cracking was also noticed. A second crack formed from the corner of the base of the roughness element and extended vertically into the polyurethane foam for a short distance. Two similar short cracks originated from the middle of the base and propagated upwards following slightly inclined paths. So a higher aspect ratio again leads to predominant cohesive failure with localised secondary fracture.

(a)

(b)

(c)
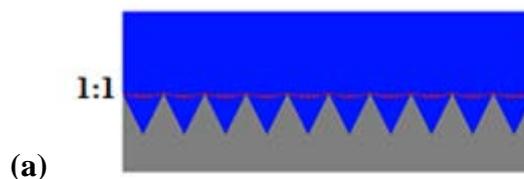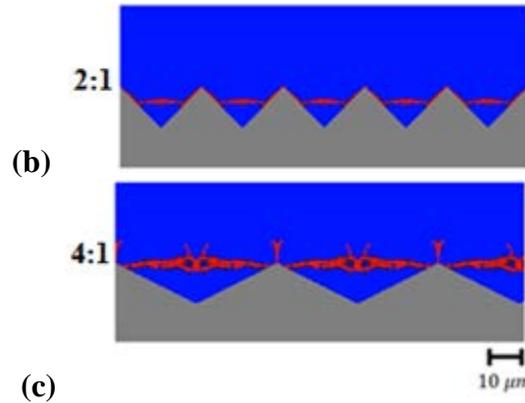
Figure 15. Jagged roughness profile in tension, showing fracture paths for different aspect ratios

Shear loading

We considered roughness profiles of the same aspect ratios as before, and analysed the crack propagation and failure mode in each case under shear. The cracks paths are shown in Figure 16. Under shear loading for all of the aspect ratios considered, the crack propagation initiated from the left corner of the roughness element. For the 1:1 aspect ratio case, two crack paths were formed (Figure 16a). The main crack propagated from the left side of the roughness element to the other side causing fracture. This caused a nearly pure cohesive failure. A secondary crack was also created from the left corner and propagated a short distance down the roughness profile nearly parallel to the interface of the roughness element. For the case with an aspect ratio of 2:1, the failure mode changed to partly adhesive and partly cohesive (Figure 16b). Only one crack path was present in this case. This crack path extended nearly three quarters of the way down the left face of the roughness element interface causing adhesive failure. Then the crack traversed across the roughness element, moving diagonally up to the top of the roughness element and across to the right corner causing cohesive fracture. The crack path for the 4:1 aspect ratio profile was similar to that of the 2:1 aspect ratio profile with a larger adhesive failure component (Figure 16c). The crack propagated downwards along the left surface of the roughness element up to approximately 90 percent of the left interface. Then it turned upwards and traversed through the polyurethane foam, but remained close to the right interface, before reaching the upper right corner. This shows a combination of adhesive and cohesive failures in the left and right interfaces for the roughness profile with a high aspect ratio. Overall, as the aspect ratio of the roughness profile increases, the proportion of adhesive failure increases under shear. This study was repeated for the triangular profile and similar trends in the transition in behaviour from cohesive to adhesive failure was noticed, as shown in Figure 17.

(a)

(b)

(c)
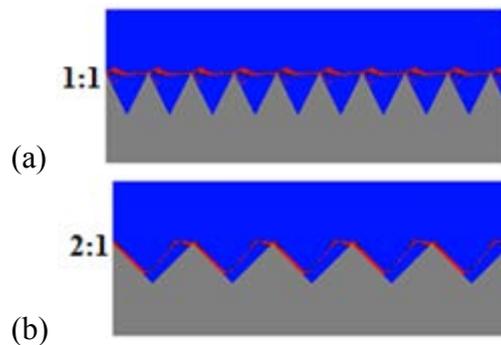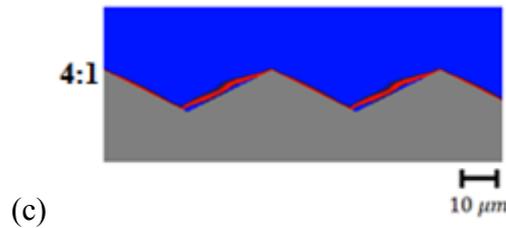
Figure 16. Jagged roughness profile in shear, showing fracture paths for different aspect ratios

(a)

(b)
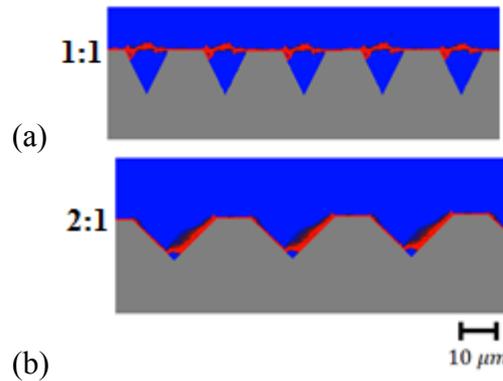
Figure 17. Triangular roughness profile in shear, showing fracture paths for different aspect ratios

Effect of Width Ratio

One characteristic feature of surface roughness is the relative spacing between adjacent roughness elements. We studied the effect of this by investigating the effect of width ratio on the fracture pattern. In this case, we considered a semi-circular roughness profile so as to be different from that considered in the aspect ratio case.

Tensile loading

For this study, we considered three width ratios, 1:1, 3:1, and 8:1, for the semi-circular roughness profile. The fracture paths for the primary (major) cracks of the semi-circular roughness profile in tension are nearly the same for all width ratios (Figure 18). For all three width ratios modelled under tensile load, the primary cracks are initiated from the two corners of the base of each roughness element, where the stress concentrations occur. From each corner, two primary cracks are initiated and propagated in the reverse direction. One of them propagated along the flat part of the interface first adhesively, and another one propagated cohesively into the polyurethane foam. The two cohesive cracks met at the centre of each roughness element.

However, secondary cracks and local level of damage for the three width ratios are somewhat different. For the 1:1 width ratio roughness profile, the two cohesive cracks met at the centre, and then extended upward into the body of the polyurethane (Figure 18a). The cracks generated in the 3:1 width ratio roughness profile caused larger damage as it propagated across the base of the roughness element (Figure 18b). The localised cracks from the centre are smaller in length and damage at the middle of the base is large. The 8:1 width ratio roughness profile showed a crack path similar to that of the 3:1 width ratio model (Figure 18c); however, the crack paths was thinner, and it produced less damage as it propagated across the base of the semi-circular roughness element.

A roughness profile of 1:1 width ratio results in approximately 50% cohesive failure, a 3:1 width ratio in 75% cohesive failure, and a 8:1 width ratio in 89% cohesive failure. From the width ratios considered, the 8:1 width ratio produces the largest cohesive failure due to it having the smallest width 'w_2'.

Broadly, within a roughness element, the width ratio does not significantly affect the mode of failure. Adhesive failure in the flat portion and cohesive failure within the roughness elements are observed in general with some difference in crack paths, due to the loading direction and the presence of stress raisers (corners). However, width ratio has a direct effect on the strength of an interface by altering the proportion of cohesive and adhesive failure. As noticed in the semi-circular surface profile under tension, when the width ratio decreases the proportion of the horizontal flat surface at the interface decreases, given the length of the roughness element base remains constant. The crack propagated across the base of the roughness elements for all of the width ratios; so decreasing the width ratio reduced the area where adhesive failure could occur, thus increasing the proportion of the cohesive failure zone.
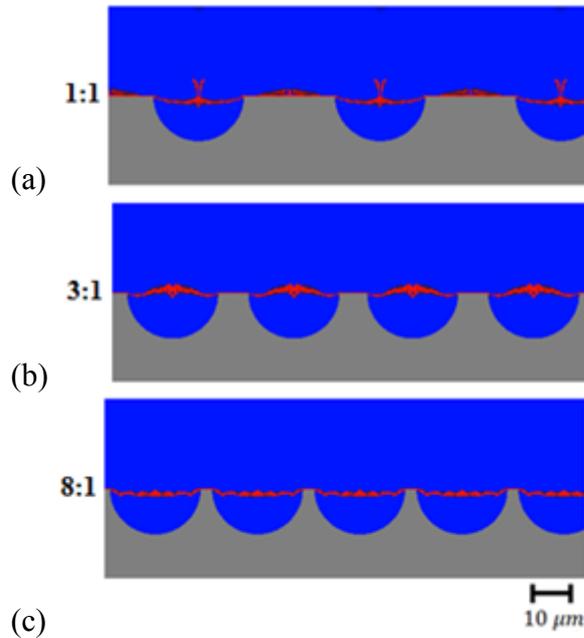
Figure 18. Semi-circular profile in tension showing failure paths for width ratios of 1:1, 3:1, and 8:1.

Shear loading

Next we study the effect of the width ratio on the failure mode under shear by considering a semi-circular roughness profile. Three different width ratios from low to high values, 1:1, 4:1 and 8:1, were adopted. Figure 19 shows the fracture pattern under shear loading for various width ratios of the roughness profile. For all of the width ratios modelled, the crack initiated from the left side and first propagated along the flat interface, thus causing adhesive failure. Upon reaching the left corner where the semi-circular profile met the flat face, the crack behaviour depended on the width ratio of the roughness profile.

Figure 19 shows the fracture pattern for the 1:1 width ratio profile. When the tip of the crack reached the left corner of the roughness element, the crack divided into branches to produce two cracks. Both the cracks subsequently traversed cohesively within the foam. The first crack propagated downwards across the roughness element in an inclined direction. Before reaching the interface, it turned about 45° to move in the upward direction to reach the right corner of the roughness element. The second crack initiated from the left corner, propagated into the body of the polyurethane foam in an inclined upward direction (approximately 45°) up to about 10 μm, and then bent towards right to traverse horizontally for a short length.

The crack path for the 4:1 width ratio profile had some features similar to that of the 1:1 width ratio profile (Figure 19b). The key differences are that the first crack propagated along the interface following the circular boundary, causing adhesive failure of the interface that continued up to nearly the bottom point of the roughness element. The crack then diverted upwards into the foam, and propagated across the roughness element in a curved path to reach the right corner. Another crack

branched out from this 'first' crack just before reaching the right corner. A second crack also originated from the left corner for this case. This crack propagated upwards in a curved path, causing cohesive failure of the polyurethane foam, similar to the 1:1 width ratio case. For the 8:1 width ratio roughness profile, the crack path followed the left hand semi-circular interface of the roughness element up to the bottom point (Figure 19c). This crack produced a fracture that primarily occurred along the interface in the roughness element, signifying a pure adhesive failure.

Hence for the 1:1 width ratio profile, the polyurethane foam fails cohesively within the semi-circular roughness elements, although the failure follows a path closer to the interface. For the roughness profiles with 4:1 and 8:1 width ratios, the failure progressively changes from cohesive to adhesive, with partial adhesive failure for the 4:1 width ratio profile and complete adhesive failure for the 8:1 width ratio profile.
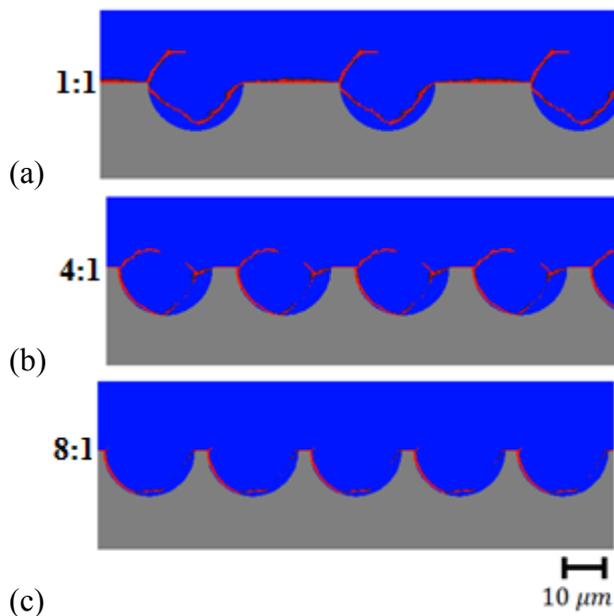
Figure 19. Semi-circular profile in shear showing failure paths for width ratios of 1:1, 4:1, and 8:1

Conclusions

In this study, the interface of a polyurethane foam and galvanised steel composite was examined using optical microscopy, SEM, and AFM. A micro-scale finite element model was subsequently developed to analyse the effect of interface roughness on the mode of failure under tension and shear. It was found that whilst the macroscopic properties of polyurethane foam are affected by voids, the microscopic properties of polyurethane at the foam-metal interface are not affected considerably, because a thin film of solid (void-less) polyurethane effectively covers the majority of the surface of the galvanised steel. Thus, the macroscopic properties of polyurethane foam including the effect of voids are not good representation of the micro-scale properties of polyurethane at the foam-metal interface.

Both the relative width to depth of roughness elements and the spacing between roughness elements, as characterised by the aspect ratio and width ratio, respectively, were found to have a profound influence on the mode of failure and interfacial strength. Cohesive failure was the dominant mode of failure in the roughness profiles under tensile load irrespective of the aspect ratio as anticipated. It was found that under tension the roughness profiles with a lower aspect ratio would fail with a large proportion in a cohesive manner when compared to the higher aspect ratios of the same roughness profile. Based on the results of the present analysis, the jagged roughness profile with a 1:1 aspect ratio leads to almost pure cohesive failure. With a moderate aspect ratio of 2:1, the failure mode changes to a combination of partly adhesive and partly cohesive modes, whilst a further increase in the aspect ratio (4:1) leads to predominant cohesive failure mode again. Failure mode under shear strongly depends on the aspect ratio of the roughness profile. For low aspect ratio

roughness, the failure is primarily cohesive, which then transitions to a combination of adhesive and cohesive failures with an increase in the aspect ratio.

Under tensile loading, adhesive failure in the horizontal flat portion and cohesive failure within the roughness elements are observed, in general, for all width ratios with some local differences in the crack paths. This is due to the direction of loading combined with the stress concentration generated at the corner of the roughness element.

When a roughness profile is subjected to shear (for the semi-circular shape in this case), the cohesive failure mainly occurs when the width ratio is low. However, for higher width ratios, the failure no longer propagates cohesively in the polyurethane, but instead propagates partially or completely adhesively along the interface. For example, for a width ratio of 8:1, a pure adhesive failure of the interface occurs.

Controlling interface roughness can change the relative proportion of cohesive and adhesive failures at a foam-metal interface, which in turn can improve the interfacial strength of a foam-metal composite based on the difference between the cohesive strength of the foam and the adhesive strength of the interface. For example, if the cohesive strength is greater than the adhesive strength, then the jagged surface roughness profile with a 1:1 aspect ratio is the optimum surface roughness profile as it results in approximately pure cohesive failure when loaded either in tension or in shear.

The approach adopted in this paper to evaluate the effect of roughness parameters on the failure mode around the interfacial region can be extended to other polymer-metal composites. Future work will be undertaken to determine micro-scale adhesive strength of polyurethane and galvanised steel for specified foam-metal composites to provide accurate model input. It appears that the key aspect to improve the strength of a foam-metal composite interfacial region is to enable preferred failure modes (cohesive/adhesive) by inducing controlled micro-scale roughness profiles accurately and consistently on the surface of the metal substrate [Kim et al. (2010)], which can be accomplished using mechanical (micro machining) and/or chemical (photochemical machining) methods.

Acknowledgement

References

ABAQUS (2014) ABAQUS Documentation, Dassault Systems, Providence, RI, USA.

Buehler, M. J. (2008) Atomistic Modeling of Materials Failure, Springer, New York.

Davis, J. R. (2014) Metals Handbook Desk Edition, ASM International, Materials Park, OH, USA.

Grujicic, M., Sellappan, V., Kotrika, S., Arakere, G., Obieglo, A., Erdmann, M. and Holzleitner, J. (2009) Suitability analysis of a polymer–metal hybrid technology based on high-strength steels and direct polymer-to-metal adhesion for use in load-bearing automotive body-in-white applications, *Journal of Materials Processing Technology* **209**, 1877-1890.

Grujicic, M., Sellappan, V., Omar, M. A., Seyr, N., Obieglo, A., Erdmann, M. and Holzleitner, J. (2008) An overview of the polymer-to-metal direct-adhesion hybrid technologies for load-bearing automotive components, *Journal of Materials Processing Technology* **197**, 363-373.

Ho, P. S. (1989) Chemistry and adhesion of metal-polymer interfaces, *Applied Surface Science* **41–42**, 559-566.

Kim, J. R., E. Miller, J. Bai, J. (2003) Adhesion of a rigid polyurethane foam to zinc phosphated steel, *Journal of Adhesion Science and Technology*.

Kim, W.-S., Yun, I.-H., Lee, J.-J. and Jung, H.-T. (2010) Evaluation of mechanical interlock effect on adhesion strength of polymer–metal interfaces using micro-patterned surface topography, *International Journal of Adhesion and Adhesives* **30**, 408-417.

Neilsen, M., Krieg, R. and Schreyer, H. (1995) A constitutive theory for rigid polyurethane foam, *Polymer Engineering and Science* **35**, 387-394.

Noijen, S. P. M., van der Sluis, O., Timmermans, P. H. M. and Zhang, G. Q. (2009) Numerical prediction of failure paths at a roughened metal/polymer interface, *Microelectronics Reliability* **49**, 1315-1318.

Ochoa-Putman, C. and Vaidya, U. K. (2011) Mechanisms of interfacial adhesion in metal–polymer composites – Effect of chemical treatment, *Composites Part A: Applied Science and Manufacturing* **42**, 906-915.

Randall, D. and Lee, S. (2002) The Polyurethanes Book, Wiley, New York.

Yao, Q. and Qu, J. (2002) Interfacial versus cohesive failure on polymer-metal interfaces in electronic packaging - Effects of interface roughness, *Journal of Electronic Packaging* **124**, 127-134.

Form-Finding and Stability Analysis of Tensegrity Structures using Nonlinear Programming and Fictitious Material Properties

*M. Ohsaki[1], J.Y. Zhang[2], and T. Taguchi[1]

[1]Department of Architecture, Hiroshima University, Kagamiyama 1-4-1, Higashi-Hiroshima 739-8527, Japan.
[2]Department of Architecture and Urban Design, Nagoya City University, Japan

*Corresponding author: ohsaki@hiroshima-u.ac.jp

Abstract

An optimization approach is presented for form-finding of tensegrity structures. It is shown that various equilibrium shapes can be easily found by solving a forced-deformation analysis problem formulated as a minimization problem of the total strain energy. The self-equilibrium forces can be found from the optimality conditions of the nonlinear programming problem, and the stability is always guaranteed owing to local convexity of the strain energy. The equilibrium shape and self-equilibrium forces can be modified by assigning fictitious material properties of cables. The proposed approach is successfully applied to form-finding of a tensegrity tower.

Keywords: Tensegrity, Form-finding, Optimization, Stability

Introduction

Tensegrity structure consists of cables and struts that carry tensile and compressive forces, respectively. Self-equilibrium forces, or prestresses, are introduced to stabilize the structure. Since the shape of the structure defined by nodal coordinates at self-equilibrium state depends on the member forces, it is difficult to obtain a desired shape. Therefore, several analytical and numerical approaches have been developed for form-finding of tensegrity structures (Zhang and Ohsaki, 2006).

Miki and Kawaguchi (2010) proposed an approach to form-finding by solving an optimization problem. Gaspani et al. (2011) carried out form-finding analysis using nonlinear programming approach. Chen et al. (2012) used an ant-colony method for form-finding.

In this study, we present a method for form-finding of tensegrity structures using a nonlinear programming approach. Various equilibrium shapes are found by utilizing fictitious material properties. Stability of the self-equilibrium state is also discussed.

Basic Equations

Let N_i ($i = 1, \ldots, m$) denote the axial force of member i of a tensegrity structure consisting of m members in the 3-dimensional space. The vector consisting of coordinates of all n nodes is denoted by $X \in \mathbb{R}^{3n}$. The unstressed length L_i^0 of member i is given. Then, the length $L_i(X)$ of i th member satisfying compatibility (connectivity) conditions at nodes is a function of X, and its gradient $\nabla L_i(X)$ consists of directional cosines of members. If we neglect the self-weight, the equilibrium equation is written as

$$\sum_{i=1}^{m} N_i \nabla L_i(X) = \boldsymbol{0} \tag{1}$$

Although the material of tensegrity structure is usually linear elastic, we use a fictitious material in

the process of form-finding. For the given unstressed member lengths, the strain energy of member i is regarded as a function of $L_i(\boldsymbol{X})$, which is denoted by $S_i(L_i(\boldsymbol{X}))$. Then the total strain energy $S(\boldsymbol{X})$ is obtained as

$$S(\boldsymbol{X}) = \sum_{i=1}^{m} S_i(L_i(\boldsymbol{X})) \tag{2}$$

The self-equilibrium shape is found by solving an optimization problem. The variables are nodal coordinates $\boldsymbol{X}$, and the objective function is the total strain energy $S(\boldsymbol{X})$. When no constraint is given, the stationary condition of $S(\boldsymbol{X})$ is given as

$$\frac{\partial S(\boldsymbol{X})}{\partial X_i} = \sum_{i=1}^{m} \frac{\partial S_i(L_i(\boldsymbol{X}))}{\partial L_i} \nabla L_i(\boldsymbol{X}) = \boldsymbol{0}, \quad (i=1,\ldots,3n) \tag{3}$$

At the optimal solution satisfying Eq. (3), the equilibrium equation (1) is satisfied by regarding $\partial S_i / \partial L_i$ as the axial force N_i of member i.

This optimization problem is a standard analysis problem with forced deformation for satisfying the compatibility at nodes for specified unstressed member lengths. Furthermore, the total potential energy is equal to the total strain energy, because no external load is applied at the self-equilibrium state. Therefore, the principle of minimum total potential energy ensures stability of the equilibrium shape obtained by minimizing the strain energy; however, we use a fictitious material, rather than the true material, in this process of form-finding.

After obtaining $\boldsymbol{X}$ as the solution of the optimization problem, we assign the properties of the true material, and compute the true axial force $N_i^*(\boldsymbol{X})$ from the member lengths $L_i(\boldsymbol{X})$ at equilibrium and the unstressed length L_i^0. Then, the tangent stiffness matrix $\boldsymbol{K}^* \in \mathbb{R}^{3n\times 3n}$ using the true material is defined as the sum of the linear stiffness matrix $\boldsymbol{K}_{\mathrm{E}}^* \in \mathbb{R}^{3n\times 3n}$ and the geometrical stiffness matrix $\boldsymbol{K}_{\mathrm{G}}^* \in \mathbb{R}^{3n\times 3n}$ as

$$\boldsymbol{K}^* = \boldsymbol{K}_{\mathrm{E}}^* + \boldsymbol{K}_{\mathrm{G}}^* \tag{4}$$

The tangent stiffness matrix using fictitious material is denoted by $\boldsymbol{K} \in \mathbb{R}^{3n\times 3n}$. Let $\lambda_{\min}$ denote the lowest (7th) eigenvalue of $\boldsymbol{K}$ excluding six zero eigenvalues corresponding to rigid-body motions. The principle of minimum total potential energy ensures that $\lambda_{\min} > 0$ at the equilibrium state. Let $\hat{\boldsymbol{K}} \in \mathbb{R}^{3n\times 3n}$ denote the increment of $\boldsymbol{K}^*$ from $\boldsymbol{K}$; i.e.,

$$\boldsymbol{K}^* = \boldsymbol{K} + \hat{\boldsymbol{K}} \tag{5}$$

Define the nodal displacement vector $\boldsymbol{d} \in \mathbb{R}^{3n}$ as a linear combination of the eigenvectors $\boldsymbol{\Phi}_i \in \mathbb{R}^{3n}$ $(i=7,\ldots,3n)$ excluding rigid-body motions as

$$\boldsymbol{d} = \sum_{i=7}^{3n} \alpha_i \boldsymbol{\Phi}_i \tag{6}$$

where α_i $(i=7,\ldots,3n)$ are arbitrary coefficients that are not equal to 0 simultaneously. Since the equilibrium state using the fictitious material is stable, $\boldsymbol{d}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{d} > 0$ holds. Therefore, the equilibrium state using the true material is stable if the following condition is satisfied:

$$\boldsymbol{d}^{\mathrm{T}}\hat{\boldsymbol{K}}\boldsymbol{d} \geq 0 \tag{7}$$

When the fictitious material is defined using a bilinear stress-strain relation with degrading stiffness, and the true material has constant stiffness that is equal to the initial stiffness of the fictitious material, then $\hat{\boldsymbol{K}}$ is positive semi-definite, and the condition (7) is satisfied. Note that this condition is a sufficient but not a necessary condition as demonstrated in the numerical examples.

We can also formulate a constrained optimization problem with upper bound $L_{J_i}^{\mathrm{U}}$ for cable J_i $(i = 1, \ldots, p)$ as

$$L_{J_i}(\boldsymbol{X}) - L_{J_i}^{\mathrm{U}} \leq 0, \quad (i = 1, \ldots, p) \tag{8}$$

The optimality condition for the minimization problem of $S(\boldsymbol{X})$ under constraint (8) is written as

$$\frac{\partial S(\boldsymbol{X})}{\partial X_i} = \sum_{i=1}^{m} \frac{\partial S_i(L_i(\boldsymbol{X}))}{\partial L_i} \nabla L_i(\boldsymbol{X}) + \sum_{i=1}^{p} \lambda_i \nabla L_{J_i}(\boldsymbol{X}) = \boldsymbol{0}, \quad (i = 1, \ldots, 3n) \tag{9}$$

Hence, the axial force of cable J_i should be equal to $\partial S_i / \partial L_i + \lambda_i$ to satisfy the equilibrium equation (1). Since $L_i(\boldsymbol{X})$ is not a convex function of $\boldsymbol{X}$, stability of the equilibrium shape using the fictitious material is not guaranteed, when constraints on member length are given.

Optimization is carried out using SNOPT Ver.7 (Gill et al., 2002) that is based on sequential quadratic programming (SQP). The sensitivity coefficients are computed analytically. When the approximate Hessian of Lagrangian is singular at a step of SQP, SNOPT stabilizes the QP subproblem by assigning small positive values on the diagonals of the Hessian, which leads to a penalty term of the quadratic norm of the increment of variables. Therefore, for the analysis problem of a free-standing tensegrity structure, the rigid-body motions are successfully excluded, and the nearest solution from the initial solution is obtained.

The algorithm of form-finding is summarized as follows:
1. Assign initial shape, unstressed lengths of members, and properties of fictitious material.
2. Solve the optimization problem to obtain the nodal coordinates at equilibrium.
3. Assign the properties of true material, and compute the axial forces at equilibrium and unstressed length using the true material.
4. Evaluate stability of the equilibrium shape.

Example of Tensegrity Tower

The proposed approach is applied to form-finding of a tensegrity tower that consists of struts, vertical cables, saddle cables, diagonal cables, and horizontal cables (Zhang and Ohsaki, 2008). An example of three-layer tower is shown in Fig. 1. Form-finding is carried out for a 20-layer tensegrity tower as shown in Fig. 2(a). The tower has three struts in each layer, and the radius and height of each layer are 1.0 m and 2.25 m, respectively. The units are omitted, in the following, for simple presentation of the results.
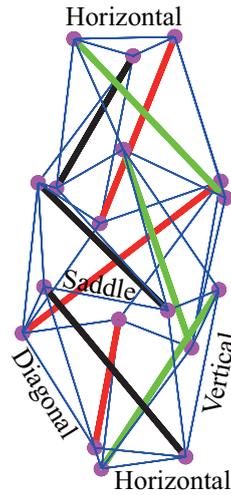
Figure 1. A 3-layer tensegrity tower model.

(a) Initial,   (b) Case 1   (c) Case 2   (d) Case 3
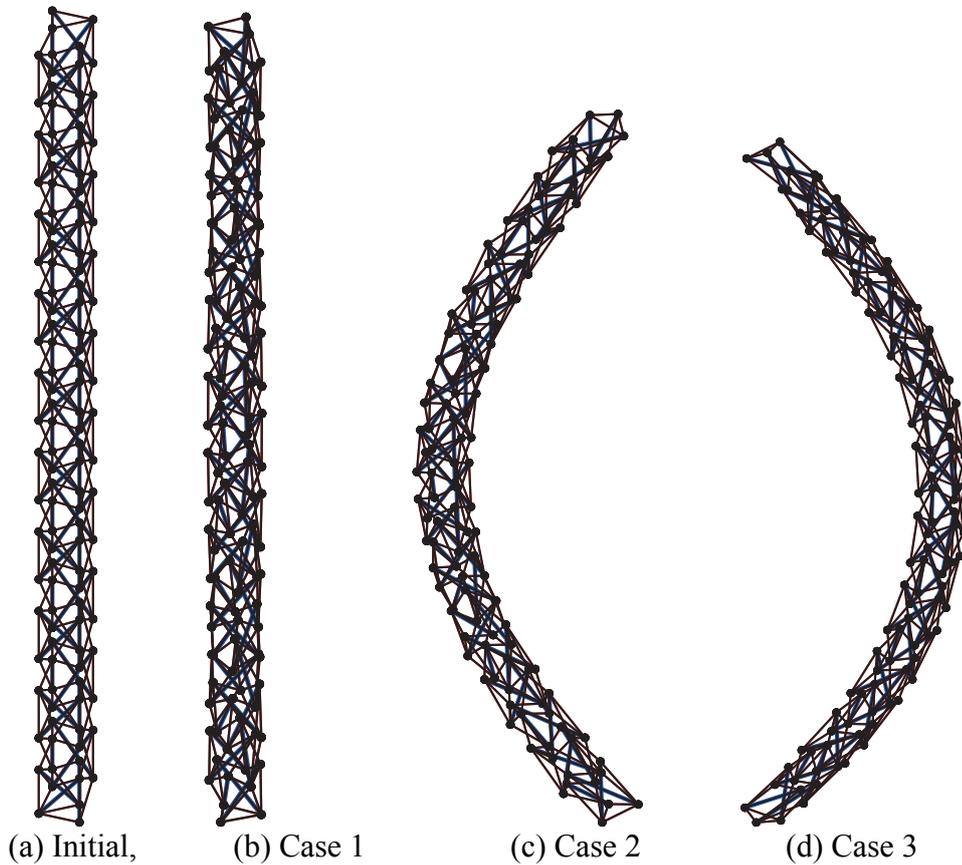
Figure 2. Initial and self-equilibrium shapes of a 20-layer tensegrity tower.

The unstressed lengths of cables and struts are assumed to be 80% and 100%, respectively, of the lengths of the members in the initial shape in Fig. 2(a). Let A_i and E_i denote the cross-sectional area and Young's modulus, respectively, of member i. The values of $A_i E_i$ for the fictitious material are 100000 for struts and 1000 for cables. Note that the unstressed lengths of cables should be sufficiently smaller than the initial lengths in Fig. 2(a) to obtain a stable equilibrium shape, and to find various shapes that are not close to the initial shape.

119

Case 1:

The equilibrium shape obtained by solving the unconstrained optimization problem is shown in Fig. 2(b). The maximum axial force among all cables is 362.9. In Case 1, the stiffness of the true material is the same as that of the fictitious material. The axial forces are divided by 100 so that the absolute values of axial forces are in the order of 1/1000 of $A_i E_i$. Eigenvalue analysis is carried out for K^* to find that the 6th and 7th smallest eigenvalues as listed in Table 1. Since the 7th eigenvalue is sufficiently larger than the 6th eigenvalue that is approximately equal to 0, the equilibrium state is stable with six zero eigenvalues corresponding to rigid-body motions.
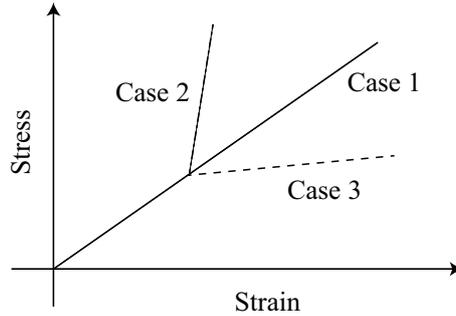
Figure 3. Bilinear stress-strain relations.

Case 2:

We next consider a fictitious material with bilinear stress-strain relation. The 60 vertical cables are classified into six groups connecting the nodes with the same xy-coordinates in the horizontal plane of the initial shape. Ten cables in one of six groups are selected to have the bilinear stress-strain relation as indicated as Case 2 in Fig. 3. The strain at the stiffness transition point is 0.1, and the value of $A_i E_i$ of the second part is $100 A_i E_i$. The equilibrium shape obtained by optimization is shown in Fig. 2(c). The minimum and maximum values of strains among the members with bilinear stress-strain relation are 0.1028 and 0.1030, which are close to 0.1. This way, a curved shape has been generated by assigning large stiffnesses for the cables that are vertically aligned at the initial shape.

We multiply 1/100 to axial forces of all members and carry out eigenvalue analysis of tangent stiffness matrix using the true material with constant stiffness $A_i E_i$ for all cables. The 6th and 7th eigenvalues are listed in Table 1, which shows that the structure is stable, although the true material has smaller stiffness than the fictitious material, and the sufficient condition (7) for stability is not satisfied. If we set the maximum member length $1.1 L_i^0$ and solve the constrained optimization problem, the same equilibrium shape as shown in Fig. 2(c) is obtained. The axial forces of the constrained members in layers 1, 3, and 5 are listed in Table 2, which confirms that the axial forces at equilibrium can be obtained as the sum of the differential coefficient $\partial S_i / \partial L_i$ and the Lagrange multiplier λ_i.

Table 1. Eigenvalues of tangent stiffness matrix using true material.

Case	6th	7th
1	6.135×10^{-8}	0.06125
2	-1.094×10^{-8}	0.02594
3	1.861×10^{-9}	0.02171

Table 2. Axial forces at equilibrium of constrained members in layers 1, 3, and 5.

Layer	Bilinear model	Constrained optimization		
		(A) Differential coefficient of strain energy	(B) Lagrange multiplier	(A) + (B)
1	397.6	100.0	300.1	400.0
3	380.2	100.0	282.6	382.6
5	380.8	100.0	283.2	383.2

Case 3:

Fictitious material property is given in the same vertical cables as Case 2. However, we decrease the value of A_iE_i of the second part of the vertical cables to $A_iE_i/100$ as indicated by Case 3 in Fig. 3. The equilibrium shape obtained by solving the unconstrained optimization problem is shown in Fig. 2(d). As seen from Figs. 2(c) and (d), the tower can be bent to opposite directions by increasing and decreasing the value of A_iE_i of the vertical cables in the specified group. The axial forces of the vertical cables with bilinear stress-strain relation are between 103 and 104, which are close to the specified value $0.1A_iE_i$. We multiply 1/100 to axial forces of all members and carry out eigenvalue analysis of tangent stiffness matrix. The 6th and 7th lowest eigenvalues are listed in Table 1, which confirms the stability of structure. Since the stiffness of the fictitious material is smaller than that of the true material, the equilibrium shape with the true material is stable, if the shape with fictitious material is stable.

Conclusions

The following conclusions have been obtained in this study:

1. Various equilibrium shapes can be obtained using the fictitios material with bilinear stress-strain relations. The equilibrium shape can be successfully found by solving an unconstrained optimization problem of minimizing the total strain energy.

2. A curved tensegrity tower can be generated by assigning fictitios materials for a group of vertically aligned vertical cables. It has been shown that the optimization problem with bilinear stress strain relation is equivalent to a constrained optimization problem with upper bound for the member lengths.

3. The equilibrium shape of the tensegrity structure is stable, if the stable equilibrium is found using a fictitious material with degrading bilinear stress-strain relation, and the true material has the constant stiffness that is equal to the initial stiffness of the fictitious material.

4. The rigid-body motions need not be constrained when solving the optimization problem using an SQP method, because the quadratic programming subproblem is automatically stabilized by assigning small positive values in the diagonals of the approximate Hessian of the Larangian.

References

Chen, Y, Feng, J. and Wu, Y. (2012) Novel form-finding of *tensegrity structures using ant colony systems, J. Mechanics and Robotics*, ASME **4**, No. 031001.

Gasparini, D., Klinka, K. K. and Arcaro V. F. (2011) A finite element for form-finding and static analysis of tensegrity structures, *J. Mechanics of Materials and Structures* **6**, 1239–1253.

Gill, P. E., Murray, W. and Saunders, M. A. (2002) SNOPT: An SQP algorithm for large-scale constrained optimization, *SIAM J. on Optimization* **12**: 979-1006.

Miki, M. and Kawaguchi, K. (2010) Extended force density method for form-finding of tension structures, *J. Int. Association for Shell and Spatial Structures* **51**, 291–303.

Zhang, J. Y. and Ohsaki. M. (2006) Adaptive force density method for form-finding problem of tensegrity structures, *Int. J. Solids and Structures* **43**, 5658–5673.

Zhang, J. Y. and Ohsaki, M. (2008) Numerical form-finding method for tensegrity tower, *Proc. IASS Symposium 2008*, Int. Assoc. for Shell and Spatial Struct., Acapulco, Mexico.

SPH simulation of sound propagation and interference

Y.O. Zhang[1,2], *T. Zhang[1,2], H. Ouyang[3] and T.Y. Li[1,2]

[1]School of Naval Architecture and Ocean Engineering, Huazhong University of Science and Technology, Wuhan 430074, China.
[2]Hubei Key Laboratory of Naval Architecture and Ocean Engineering Hydrodynamics, Huazhong University of Science and Technology, Wuhan 430074, China
[3]School of Engineering, University of Liverpool, The Quadrangle, Liverpool L69 3GH, UK.

*Corresponding author: zhangt7666@mail.hust.edu.cn

Abstract

Complicated domain topologies and moving boundaries in acoustic simulation are difficult to be described with mesh based methods. On the contrary, the meshfree SPH (Smoothed Particle Hydrodynamics) method does not have much trouble in dealing with these problems. Therefore, the present paper aims to simulate sound propagation and interference in time domain with the SPH method by solving linearized acoustic wave equations. Firstly, linearized acoustic wave equations are represented in the form of particle approximation. After that, a standard SPH numerical method for simulating sound waves in time domain is built by adding the leapfrog integration and the nearest neighbor particle searching method. Finally, both one dimensional sound propagation and interference models are simulated with the SPH method and results are validated and compared with theoretical data. Numerical results show that the SPH method can simulate acoustic waves accurately.

Keywords: SPH, sound propagation, sound interference, acoustic wave, time domain

Introduction

Mesh-based methods are widely used in modeling acoustic waves and these methods include some classic numerical methods such as the Finite Element Method (FEM) [Ihlenburg (1998)] and the Boundary Element Method (BEM) [Kythe (1995)]. However, these method is not perfect in solving problems with moving or deformable boundaries or interfaces due to its mesh-based properties.

Meshfree methods can handle these problems with a set of arbitrarily distributed nodes instead of mesh and many methods have been used in solving acoustic problems. As a meshfree, Lagrangian method, the SPH method not only has almost all advantages that meshfree methods have, but it is also suitable for solving problems with large ranges of density and object separation as shown in recent reviews by Springel [Springel (2010)], Liu and Liu [Liu and Liu (2010)] and Monaghan [Monaghan (2012)] due to its Lagrangian property. Introducing this method to acoustic simulation would also bring its advantages to some specific fields like combustion noise, bubble acoustic, sound propagation in multiphase flow et al. Therefore, this paper focuses on the application of SPH in the simulation of acoustic waves.

The SPH method was first pioneered independently by Lucy [Lucy (1977)] and Gingold and Monaghan [Gingold and Monaghan (1977)] in 1977 to solve astrophysical problems. It computes with a set of particles which possess individual material properties. Owing to its properties that have mentioned, the SPH method has been used in the fields of astrophysics, structure deformation, fluid dynamics etc. [Springel (2010); Liu and Liu (2010); Monaghan (2012)]. However, no literature is found about solving linearized acoustic equations with SPH, except for few researches [Wolfe and

Semwal (2007); Hahn and Negrut (2009)] discussed solving fluid dynamic equations to obtain sound waves.

Solving the fluid dynamic equations can model acoustic problems, but it is not the only way. Since there are large differences in the length scale between the values and variations of velocity and density, solving the linearized acoustic equations requires lower computational resource compared with solving the fluid dynamic equations and it has been widely used in modeling engineering problems [Bruneau (2010)]. However, no literature was found to use the SPH method to solve the linearized acoustic equations.

The paper is organized as follows. In section 2, the linearized acoustic equations are solved with the standard SPH theory and then the acoustic modeling method is built by adding the time integration and neighbor particles searching method. In section 3 and 4, one-dimensional sound propagation and interference model are simulated with the standard SPH method and the results are validated and compared with the theoretical solution.

SPH Formulations of Sound Waves

The linearized continuity and momentum equations governing sound waves can be written as

$$\frac{\partial(\delta\rho)}{\partial t} = -\rho\,\nabla\cdot\boldsymbol{u} \tag{1}$$

$$\frac{\partial\boldsymbol{u}}{\partial t} = -\frac{1}{\rho}\nabla p \tag{2}$$

The linearized state equation for ideal air is

$$p = c_0^2\delta\rho \tag{3}$$

where $\delta\rho$ is the change of density, ρ is the density, u is the velocity vector, t is the time, p is the sound pressure, c0 is the sound speed. The particle approximation equation of the continuity of acoustic waves is written as

$$\frac{\partial(\delta\rho_i)}{\partial t} = (\rho_0 + \delta\rho_i)\sum_{j=1}^{N}\frac{m_j}{(\rho_0 + \delta\rho_j)}\boldsymbol{u}_{ij}\nabla_i W_{ij} \tag{4}$$

The momentum equation in SPH method is obtained as

$$\frac{\partial\boldsymbol{u}_i}{\partial t} = \sum_{j=1}^{N}m_j\left[\frac{p_i}{(\rho_0 + \delta\rho_i)^2} + \frac{p_j}{(\rho_0 + \delta\rho_j)^2}\right]\nabla_i W_{ij} \tag{5}$$

Particle approximation of the equation of state is

$$p_i = c_0^2\delta\rho_i \tag{6}$$

The second order leap-frog integration [Kelager (2006)] is used in the paper. All-pair search approach [Liu and Liu (2003)], as a direct and simple algorithm, is used to realize the neighbor particles searching in acoustic waves simulation.

Sound Propagation

Sound Propagation Model

A one-dimensional sound propagation in a pipe with uniform cross section is used. The sound propagation model is shown in Figure 1.
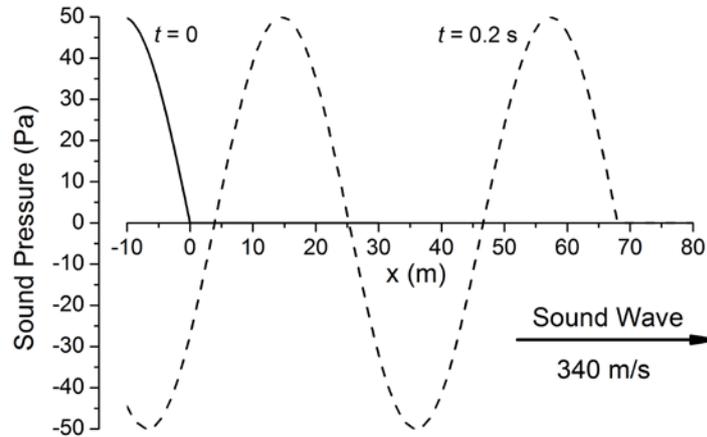
Figure 1 One-dimensional sound propagation model

The sound pressure of the acoustic wave transmitted in Figure 1 is

$$p(t, x < 0) = p_A \sin(wt - kx) \tag{}$$

where t is the time (propagation starts when t = 0), x is the geometric position, pA is the amplitude of the acoustic wave (in this section, pA = 50 Pa), w is the circular frequency of wave (in this section, w = 50 rad/s), k = w/c0, the sound speed c0 is 340 m/s and the density of the propagation medium is 1.0 kg/m3.

The sound propagates from x < 0 to x > 0 and the computational domain is from -10 m to 80 m. The simulation results at the time t = 0.2 s are used to compare with the theoretical resolution.

SPH Simulation

The simulation results of sound pressure at the time t = 0.2 s are shown in Figure 2 (a) while the theoretical solution is also plot in the figure. It can be seen from the line graph that there are two peaks appear in the propagation and one of them is shown in Figure 2 (b). At the same time, a detail view of the start of the sound is also given in Figure 2 (c).

(a) sound pressure at t = 0.2 s

124

(b) detail view 1   (c) detail view 2
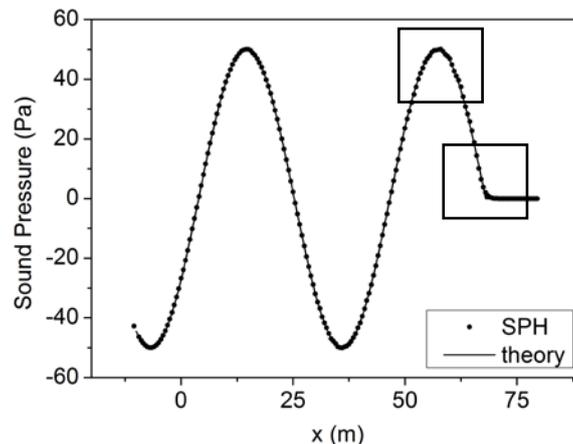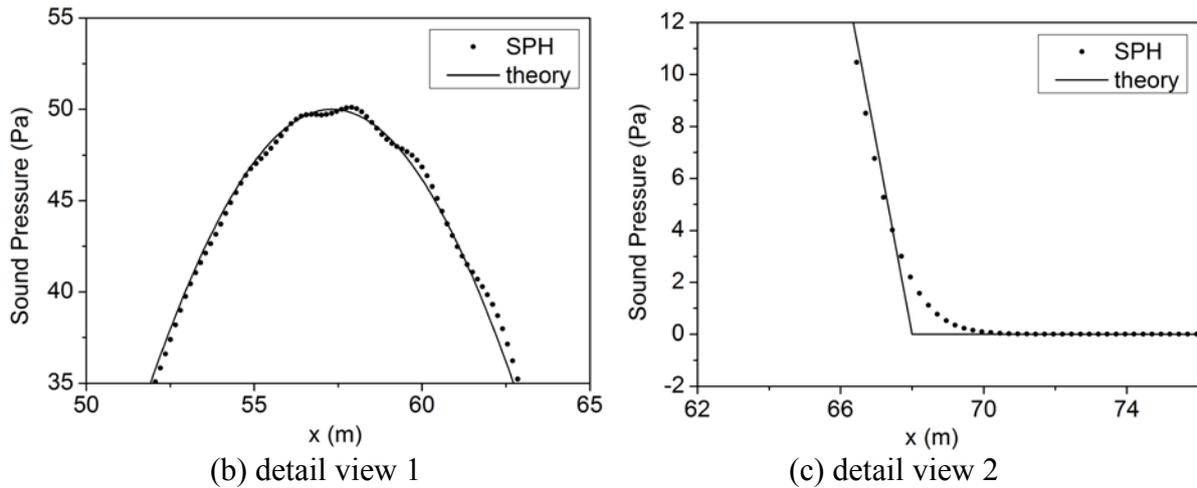
Figure 2 Sound pressure comparison between the SPH and the theoretical results at t = 0.2 s

As can be seen from the figure, the SPH simulation results have almost the same trend compared with the theoretical solution. Values of the sound pressure can also be obtained accurately by using the SPH method. However, Figure 2 (b) and (c) show the effects of unphysical oscillations and it mainly appears at the place with large changes of sound pressure.

Sound Interference

Sound Interference Model

An interference model of two different sound waves is used as shown in Figure 3.
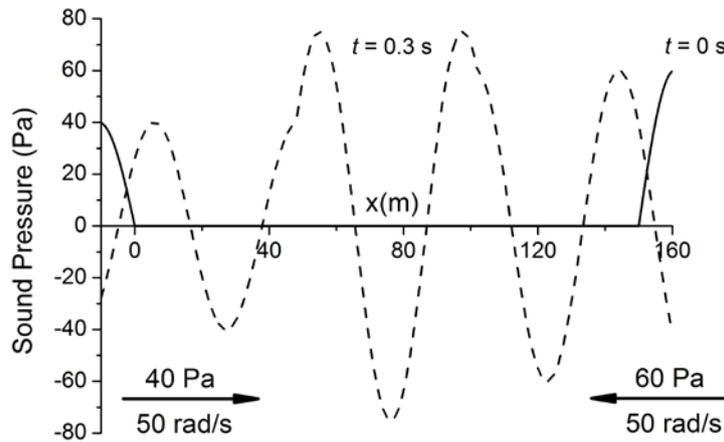
Figure 3 One-dimensional model of sound interference between two different sound waves

A sound wave with 40 Pa sound pressure and 50 rad/s circular frequency transmits from the left side while another sound wave with 60 Pa and 50 rad/s comes from the right side. After 0.3 s, the sound pressure along the x axis is shown in dash line in Figure 3.

SPH Simulation

The simulation results and theoretical solution of sound pressure at the time t = 0.3 s are shown in Figure 4 (a). Two detail views of a peak and a valley are given in Figure 4 (b) and (c).

(a) sound pressure at t = 0.3 s

(b) detail view 1
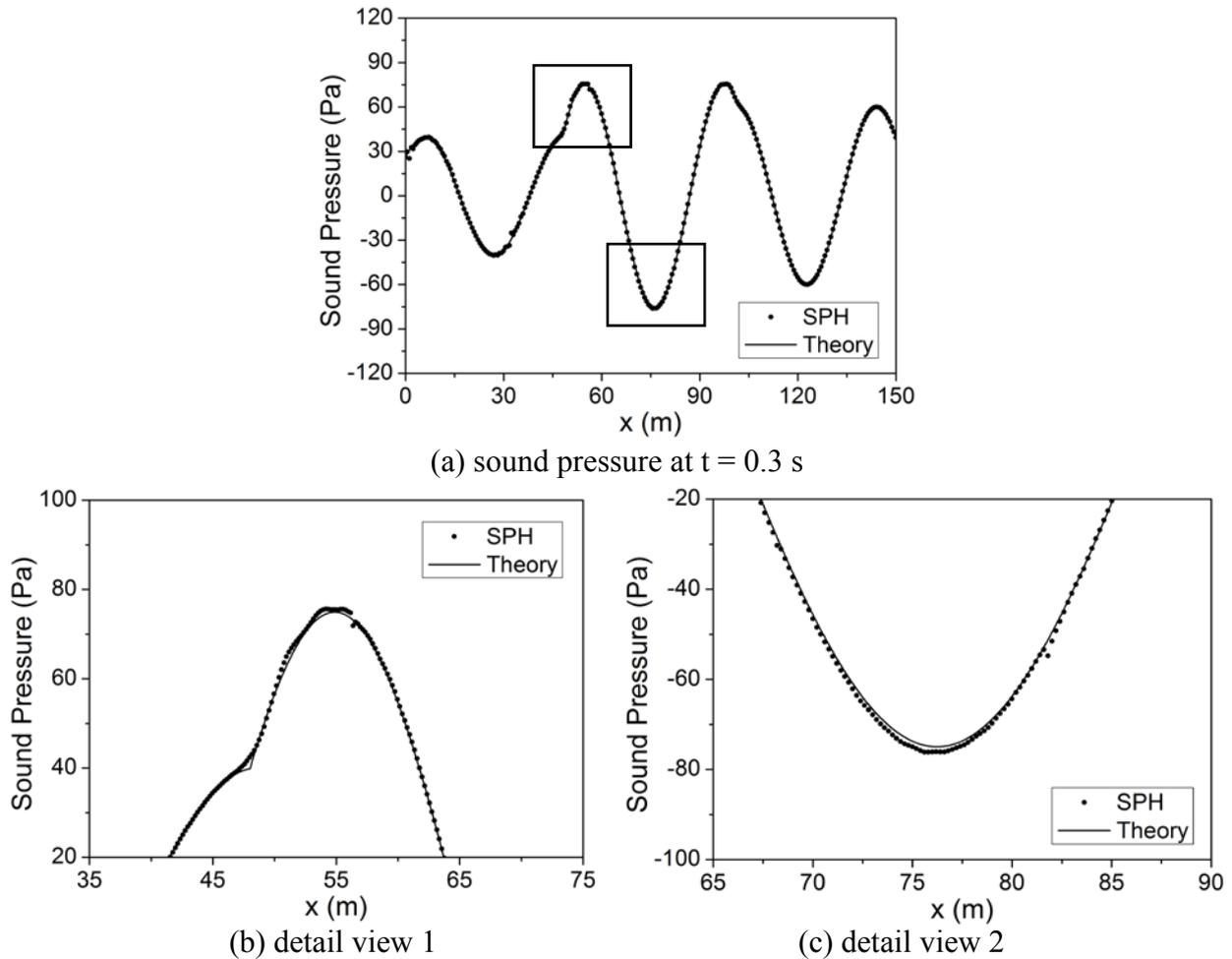
(c) detail view 2

Figure 4

Considering the computational time is 0.3 s, the interference happens at $48 < x < 102$ m. It can be seen from the line figure that the SPH simulation results agree well with the theoretical solution. As shown in the detail views, the standard SPH method can compute sound pressure accurate comparing with the theoretical results.

Conclusions

Linearized acoustic equations are solved with the standard SPH theory and the simulation method is built by adding the time integration and neighbor particles searching method. One dimensional sound propagation and interference models are simulated with the SPH method and computational results are compared with theoretical data. Sound pressure results show that the standard SPH method can achieve accurate solution, although unphysical oscillations cannot be ignored.

References

Bruneau M. (2010) *Fundamentals of Acoustics*, John Wiley & Sons.
Gingold R. A. and Monaghan J. J. (1977) Smoothed Particle Hydrodynamics - theory and application to non-spherical stars, *Monthly Notices of the Royal Astronomical Society*, **181**, 375-389.
Hahn P. and Negrut D. (2009) On the use of meshless methods in acoustic simulations, *ASME 2009 International Mechanical Engineering Congress and Exposition: American Society of Mechanical Engineers*, 185-199.
Ihlenburg F. (1998) *Finite Element Analysis of Acoustic Scattering*, Springer.
Kythe P. K. (1995) *An Introduction to Boundary Element Methods,* CRC press.
Kelager M. (2006) *Lagrangian Fluid Dynamics Using Smoothed Particle Hydrodynamics*, University of Copenhagen.

Lucy L. B. (1977) A numerical approach to the testing of the fission hypothesis, *The Astronomical Journal*, **82**, 1013-1024.

Liu G. R. and Liu M. B. (2003) *Smoothed Particle Hydrodynamics: A Meshfree Particle Method*, World Scientific.

Liu M. B. and Liu G. R. (2010) Smoothed Particle Hydrodynamics (SPH): an overview and recent developments, *Archives of Computational Methods in Engineering*, **17**, 25-76.

Monaghan J. J. (2012) Smoothed Particle Hydrodynamics and its diverse applications, *Annual Review of Fluid Mechanics*, **44**, 323-346.

Springel V. (2010) Smoothed Particle Hydrodynamics in astrophysics, *Annual Review of Astronomy and Astrophysics* **48**, 391-430.

Wolfe C. T. and Semwal S. K. (2007) *Acoustic Modeling of Reverberation Using Smoothed Particle Hydrodynamics*, University of Colorado at Colorado, Springs.

Self-equilibrium Analysis of Cable Structures based on Isogeometric Analysis

*Keigo Harada[1], Jingyao Zhang[2], Toshiyuki Ogawa[1]

[1]Department of Architecture and Building Engineering, Tokyo Institute of Technology, 2-12-1 O-okayama, Meguro-ku, 152-8550, Tokyo, Japan
[2]Department of Design and Architecture, Nagoya City University, 2-1-10 Kita-chikusa, Chikusa-ku, Nagoya, 464-0083, Japan

*Keigo Harada: harada.k.ad@m.titech.ac.jp

Abstract

Non-linear analysis of cable structures is computationally expensive due to large deformation against external loads. The Isogeometric Analysis method (IGA) initially developed by T.J.R. Hughes is considered to be more efficient than the existing numerical methods for large-deformation analysis of cable structures. Moreover, Isogeometric Analysis is well suited for the structures with curved configurations, because the same mathematical descriptions for the geometry in the design (CAD) and the modeling in the analysis (FEA) are used. In this paper, we consider the self-equilibrium analysis of catenary cables as well as parabolic cables by using Isogeometric Analysis. The results demonstrate effectiveness and accuracy of Isogeometric Analysis for large deformation analysis of unstable structures, compared to the existing analysis methods.

Keywords: Cable structures, Finite element analysis, Isogeometric analysis, B-spline curve, Self-equilibrium analysis, Singular value decomposition

Introduction

There is a big gap between (computer aided) design (CAD) and analysis in conventional finite element analysis (FEA). This comes from the fact that they are using different mathematical descriptions for the geometry. The gap becomes critical for curved structures, such as shells and cable structures, because their geometries are much complex. To solve the gap by using the same mathematical description for both design and analysis, Hughes et al. (2005; 2009) and thereafter many other researchers developed a new analysis tool, called Isogeometric Analysis method (IGA). Furthermore, smoothness in the curved structures can also be guaranteed in IGA.
The Isogeometric Analysis has been extensively applied for the studies on shell and plate structures, see for example those by Stefan et al. (2011) and Benson et al. (2010). However, there are only a limited number of studies on cables by using IGA. In this paper, we will apply IGA for self-equilibrium analysis of cables under gravity, and investigate its efficiency as well as accuracy by comparison with conventional FEA.

B-spline curve

IGA and conventional FEA share almost the same analysis procedure, except that they use different shape functions. The same mathematical descriptions in (CAD) design, for example B-spline or NURBS curves (surfaces), are used as shape functions in IGA. In the following, we adopt B-spline curves as shape functions, which are constructed by taking a linear combination of B-spline basis functions. The vector-valued coefficients of the basis functions are referred to as control points. A piecewise-polynomial B-spline curve is given by

$$C(\xi) = \sum_{i=1}^{n} N_{i,p}(\xi) \, \mathbf{B}_i \qquad (1)$$

where n is the number of control points, p is the polynomial order, ξ_i is the local coordinate of the ith knot, and $\mathbf{B}_i$ is the (global) coordinates of the ith control point. Moreover, the basis functions $N_{i,p}(\xi)$ are defined as follows:

$$N_{i,0}(\xi) = \begin{cases} 1 & \text{if } \xi_i \leq \xi < \xi_{i+1} \\ 0 & \text{otherwise} \end{cases}$$

$$N_{i,p}(\xi) = \frac{\xi - \xi_i}{\xi_{i+p} - \xi_i} N_{i,p-1}(\xi) + \frac{\xi_{i+p+1} - \xi}{\xi_{i+p+1} - \xi_{i+1}} N_{i+1,p-1}(\xi)$$

(2)

Eq. (2) is referred to as the Cox-de Boor recursion formula (Cox, 1971; de Boor, 1972). Piecewise linear interpolation of the control points gives the so-called control polygon.
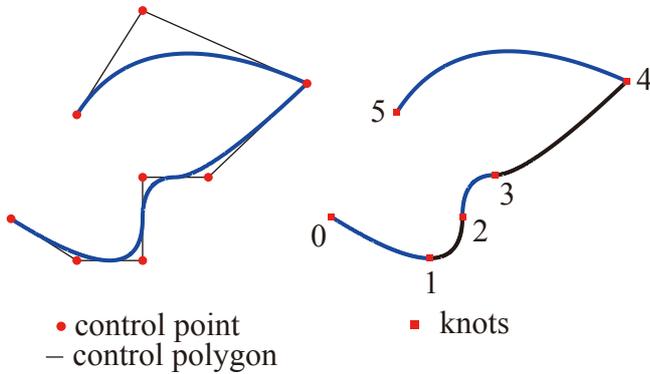
Figure 1. B-spline curve with control points, control polygon, and knots

$p = 2$, $\Xi = \{0, 0, 0, 1, 2, 3, 4, 4, 5, 5, 5\}$
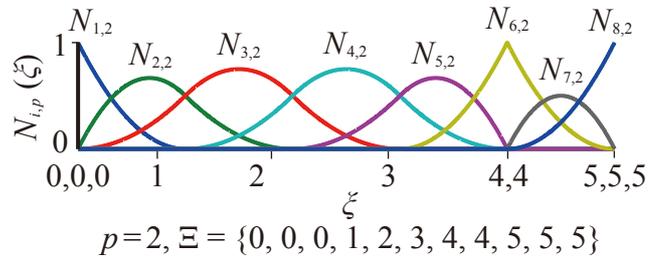
Figure 2. Quadratic B-spline basis function

An example B-spline curve is shown in Figure 1 with eight control points and $p = 2$; the resulting control polygon is shown in Figure1, and the B-spline basis functions are shown in Figure 2. Note that the curve is interpolatory at the first and last control points, due to the fact that the knot vector is open, and also at the sixth control point, due to the fact that the multiplicity of the knot $\xi = 4$ is equal to the polynomial order. Note also that the curve is tangent to the control polygon at the first, last and sixth control points. The curve is C^{p-1}-continuous everywhere except at the location of the repeated knot, $\xi = 4$, where it is $C^{p-2} (= C^0)$-continuous.

To describe a two-dimensional B-spline, it is convenient to summarize the basis functions and their first-order derivative in a matrix form as follows:

$$\mathbf{N} = \begin{bmatrix} N_{1,p} & 0 & N_{2,p} & 0 & \cdots & N_{i,p} & 0 & \cdots & N_{n,p} & 0 \\ 0 & N_{1,p} & 0 & N_{2,p} & \cdots & 0 & N_{i,p} & \cdots & 0 & N_{n,p} \end{bmatrix}$$

(3)

$$\dot{\mathbf{N}} = \begin{bmatrix} \dfrac{dN_{1,p}}{d\xi} & 0 & \dfrac{dN_{2,p}}{d\xi} & 0 & \cdots & \dfrac{dN_{i,p}}{d\xi} & 0 & \cdots & \dfrac{dN_{n,p}}{d\xi} & 0 \\[3mm] 0 & \dfrac{dN_{1,p}}{d\xi} & 0 & \dfrac{dN_{2,p}}{d\xi} & \cdots & 0 & \dfrac{dN_{i,p}}{d\xi} & \cdots & 0 & \dfrac{dN_{n,p}}{d\xi} \end{bmatrix}$$

(4)

where the components in $\dot{\mathbf{N}}$ are given as

$$\frac{d}{d\xi}N_{i,p}(\xi) = \frac{p}{\xi_{i+p} - \xi_i}N_{i,p-1}(\xi) - \frac{p}{\xi_{i+p+1} - \xi_{i+1}}N_{i+1,p-1}(\xi) \tag{5}$$

Formulations in large deformation

The tangent stiffness matrix $\mathbf{K}$ is the sum of the linear stiffness matrix $\mathbf{K}_E$ and the geometrical stiffness matrix $\mathbf{K}_G$:

$$\mathbf{K} = \mathbf{K}_E + \mathbf{K}_G \tag{6}$$

The B-spline curves are used as shape functions for analysis of cable structures, thus, the formulations for $\mathbf{K}_E$ and the geometrical stiffness matrix $\mathbf{K}_G$ are given as [Bathe (1995)]

$$\mathbf{K}_E = \frac{EAL_0}{2}\int_{\xi_1}^{\xi_n}\mathbf{B}^T\mathbf{B}\,d\xi \tag{7}$$

$$\mathbf{K}_G = \frac{AL_0}{2}\int_{\xi_1}^{\xi_n}\bar{\sigma}\mathbf{B}_{NL}^T\mathbf{B}_{NL}\,d\xi \tag{8}$$

$$\mathbf{B} = \frac{4}{L^2}\mathbf{X}^T\dot{\mathbf{N}}^T\dot{\mathbf{N}} \tag{9}$$

$$\mathbf{B}_{NL} = \frac{2}{L}\dot{\mathbf{N}} \tag{10}$$

$$\bar{\sigma} = E\bar{\varepsilon} = E\frac{4}{L^2}\mathbf{X}^T\dot{\mathbf{N}}^T\dot{\mathbf{N}}\mathbf{U} \tag{11}$$

$$\mathbf{U} = \mathbf{X} - \mathbf{X}^0 \tag{12}$$

$$\mathbf{X}^T = \left[x_1, y_1, \cdots, x_i, y_i, \cdots, x_n, y_n\right] \tag{13}$$

$$\left(\mathbf{X}^0\right)^T = \left[x_1^0, y_1^0, \cdots, x_i^0, y_i^0, \cdots, x_n^0, y_n^0\right] \tag{14}$$

Where E is Young's Modulus, A is the initial cross-sectional area, L_0 is the initial element length before deformation, L is the current length after deformation, $\bar{\sigma}$ is the axial stress in small deformation, $\bar{\varepsilon}$ is the axial strain in small deformation x_i, y_i are the current coordinates of the specified nodes of the element, and x_i^0, y_i^0 are the initial coordinates of the specified nodes of the element. For large deformation problems, the true axial strain has to be calculated from the extension of the cables, which is given as

$$\varepsilon = \frac{ds}{ds^0} - 1 = \sqrt{\left(\frac{dx}{ds^0}\right)^2 + \left(\frac{dy}{ds^0}\right)^2} - 1 = \frac{2}{L_0}\sqrt{\left(\frac{dx}{d\xi}\right)^2 + \left(\frac{dy}{d\xi}\right)^2} - 1 \tag{15}$$

Structural analysis by singular value decomposition

Tangent stiffness matrix $\mathbf{K}$ of an unstable structure is not invertible, because it is singular. To proceed the analysis for unstable structures ruling out the mechanisms as well as rigid-body motions, which cause singularity of $\mathbf{K}$, singular value composition of $\mathbf{K}$ turns out to be convenient for formulations as well as computations [Kawaguchi (2011)]. By using a unitary matrix $\mathbf{\Psi}$, a (symmetric) tangent stiffness matrix $\mathbf{K}$ is rewritten as follows:

$$\mathbf{K} = \mathbf{\Psi} \begin{pmatrix} \lambda_1 & & & & & \\ & \ddots & & & \mathbf{O} & \\ & & \lambda_i & & & \\ & & & \lambda_{i+1} & & \\ & \mathbf{O} & & & \ddots & \\ & & & & & \lambda_{dof} \end{pmatrix} \mathbf{\Psi}^{\mathbf{T}} \tag{16}$$

where **O** is a zero matrix, λ_i is the ith singular value of **K**, *dof* is the number of degrees of freedom of the system. The pseudo-inverse matrix $\mathbf{K}^{-}$ of the tangent stiffness matrix **K** is obtained as follows

$$\mathbf{K}^{-} = \mathbf{\Psi} \begin{pmatrix} 1/\lambda_1 & & & & & \\ & \ddots & & & \mathbf{O} & \\ & & 1/\lambda_i & & & \\ & & & 1/\lambda_{i+1} & & \\ & \mathbf{O} & & & \ddots & \\ & & & & & 1/\lambda_{dof} \end{pmatrix} \mathbf{\Psi}^{\mathbf{T}} \tag{17}$$

$$\text{if} \quad \lambda_i = 0 \quad \text{then} \quad 1/\lambda_i = 0$$

Subjected to the external load **F**, the displacements of control points of a (unstable) cable structure can be calculated by using the $\mathbf{K}^{-}$ defined in Eq. (17) as follows:

$$\mathbf{U} = \mathbf{K}^{-}\mathbf{F} \tag{18}$$

Accuracy evaluation and initial settings for analysis

In this paper, we analyze the self-equilibrium shapes of the cable structures subjected to gravity, and verify the accuracy of the analyses, which is evaluated by the mean square error (MSE) defined as

$$\text{RSE} = \sqrt{\frac{1}{m} \sum_{i=1}^{m} \left(\frac{y_i - \overline{y}_i}{f} \right)^2} \times 100 \quad [\%] \tag{19}$$

where *m* is the number of evaluation points, y_i, $\overline{y}_i$ are respectively the ith y-coordinate calculated by analysis and by theory, and *f* is sag of the cable.
In this paper, two cable structures with different initial shapes. Each of them are analyzed by different models:
· 9 two-node isoparametric elements with 10 nodes,
· 30 two-node isoparametric elements with 31 nodes,
· 9 four-node isoparametric elements with 10 (external) nodes,
· 30 four-node isoparametric elements with 31 (external) nodes,
· a single cubic B-spline curve with 10 control points, and
· a single cubic B-spline curve with 31 control points.
Two-node isoparametric elements are interpolated by straight lines, and four-node isoparametric elements are interpolated by cubic curves. To have the same (cubic) order for geometry description, the isogeometric elements are interpolated by the same polynomial order as four-node isoparametric

elements. Young's Modulus E is 205[GPa], the initial cross-sectional area A is 0.0001[m^2], spatial span is 30[m], the weight of the cable per unit length μ for catenary cables is 7.85[N/m], the vertical distributed load w_0 for parabolic cables is 7.85[N/m], and the number of evaluation points of mean square error is 3000 points. CPU is 2.8 GHz Intel Core i7, the memory of the CPU is 12GB, and analysis software is MATLAB R2007b provided by MathWorks Corporation.

Self-equilibrium analysis of catenary cable

The self-equilibrium shape of a single cable against its own weight becomes a catenary [Japan Society of Civil Engineers (2001)]. In this section, a catenary is used as the exact solution. The formulation of symmetric catenary cable is give as

$$y = \frac{T_0}{\mu} \cosh\left(\frac{\mu x}{T_0}\right) \tag{20}$$

where x, y is x-coordinate and y-coordinate respectively, T_0 is the horizontal tension.

(a) two-node elements (b) four-node elements (c) cubic B-spline elements

Figure 3. Initial catenary cable of object 1 with 10 nodes

(a) two-node elements (b) four-node elements (c) cubic B-spline elements

Figure 4. Final catenary cable of object 1 with 10 nodes

(a) two-node elements (b) four-node elements (c) cubic B-spline elements

Figure 5. Initial catenary cable of object 2 with 10 nodes

(a) two-node elements (b) four-node elements (c) cubic B-spline elements

Figure 6. Final catenary cable of object 2 with 10 nodes

(a) two-node elements (b) four-node elements (c) cubic B-spline elements

Figure 7. Initial catenary cable of object 1 with 31 nodes

(a) two-node elements (b) four-node elements (c) cubic B-spline elements

Figure 8. Final catenary cable of object 1 with 31 nodes

| (a) two-node elements | (b) four-node elements | (c) cubic B-spline elements |

Figure 9. Initial catenary cable of object 2 with 31 nodes

| (a) two-node elements | (b) four-node elements | (c) cubic B-spline elements |
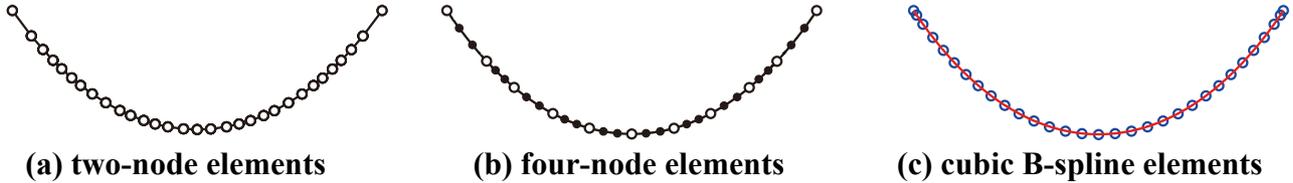
Figure 10. Final catenary cable of object 2 with 31 nodes

Table 1. Identified results of catenary cable

The number of nodes	Object	Element	The number of iterations	MSE[%]	Analysis time [s]
10 nodes	1	two-node	14 times	1.0620	0.2274
		four-node	27 times	0.0206	0.1843
		B-spline	26 times	0.0070	0.1247
	2	two-node	20 times	0.9425	0.3150
		four-node	93 times	0.5024	0.6957
		B-spline	37 times	0.2553	0.1767
31 nodes	1	two-node	21 times	0.0992	0.8055
		four-node	27 times	0.0023	1.1661
		B-spline	46 times	0.0019	2.0281
	2	two-node	79 times	0.0897	2.5160
		four-node	45 times	0.0165	1.9095
		B-spline	76 times	0.0052	3.2307

The initial shapes from which the large deformation analysis for different modeling are shown in Figures 3, 5, 7, and 9, and their corresponding final shapes due to gravity are respectively shown in Figures 4, 6, 8, and 10. Note that in (a) and (b) in these figures, ○ refers to element boundary node, ● refers to element internal node; and moreover, in (c) in these figures, ○ refers to control point.

Performances of the analyses using conventional FEA as well as IGA with different number of elements are summarized in Table 1. It was clear that IGA is more accurate compared to conventional FEA when the structure is modeled by using the same (external) nodes (or control points for IGA). On the other hand, convergence performance of IGA is not superior to that of conventional FEA.

Self-equilibrium analysis of parabolic cable

The self-equilibrium shape of cable with large vertical distributed load compared to its own weight becomes a parabolic cable. Parabolic cables are widely used in design of suspension bridges. In the analysis, the weight of the cable is regarded as zero and vertical distributed loads like floor slabs of the bridge are treated as loads applied to the nodes. The formulation of a symmetric parabolic cable is given as

$$y = \frac{w_0}{2T_0} x^2 \tag{21}$$

The initial shapes from which the large deformation analysis for different modeling are shown in Figures 11, 13, 15, and 17, and their corresponding final shapes due to gravity are respectively shown in Figures 12, 14, 16, and 18. Note that in (a) and (b) in these figures, ○ refers to element boundary node, ● refers to element internal node; and moreover, in (c) in these figures, ○ refers to control point.
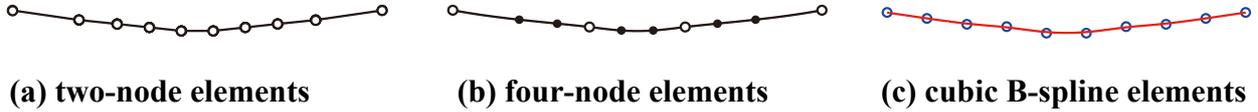
(a) two-node elements　　　**(b) four-node elements**　　　**(c) cubic B-spline elements**
Figure 11.　Initial parabolic cable of object 1 with 10 nodes
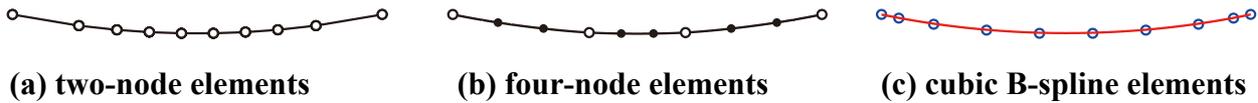
(a) two-node elements　　　**(b) four-node elements**　　　**(c) cubic B-spline elements**
Figure 12.　Final parabolic cable of object 1 with 10 nodes
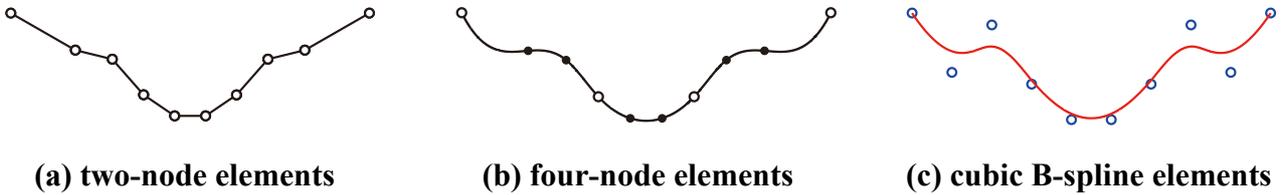
(a) two-node elements　　　**(b) four-node elements**　　　**(c) cubic B-spline elements**
Figure 13.　Initial parabolic cable of object 2 with 10 nodes
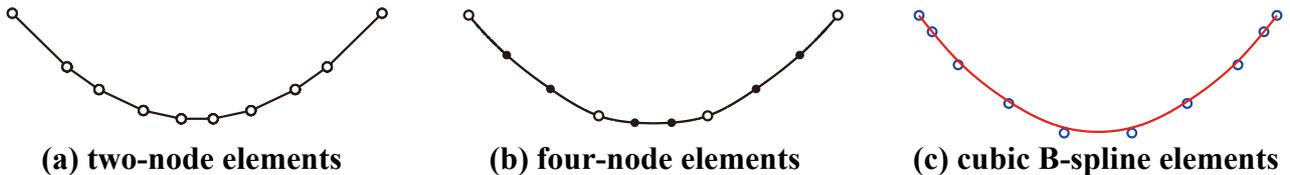
(a) two-node elements　　　**(b) four-node elements**　　　**(c) cubic B-spline elements**
Figure 14.　Final parabolic cable of object 2 with 10 nodes
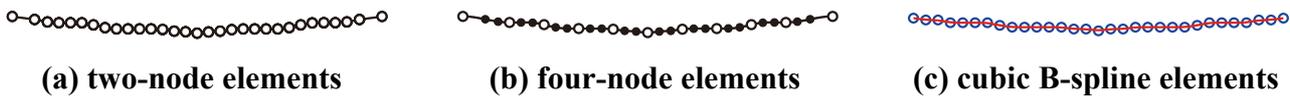
(a) two-node elements　　　**(b) four-node elements**　　　**(c) cubic B-spline elements**
Figure 15.　Initial parabolic cable of object 1 with 31 nodes
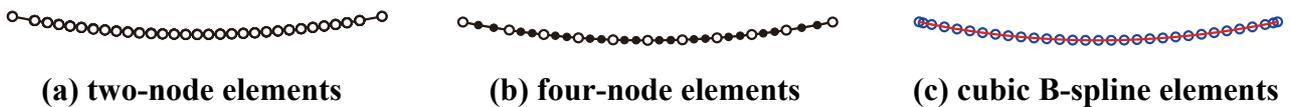
(a) two-node elements　　　**(b) four-node elements**　　　**(c) cubic B-spline elements**
Figure 16.　Final parabolic cable of object 1 with 31 nodes
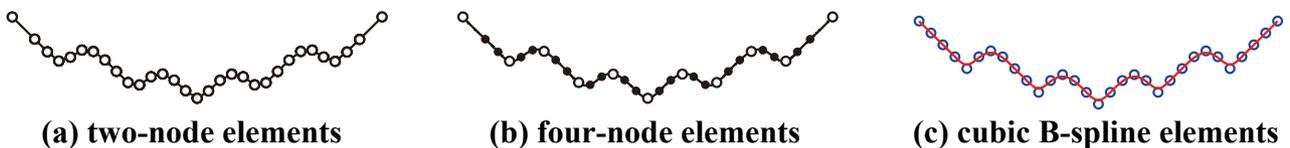
(a) two-node elements　　　**(b) four-node elements**　　　**(c) cubic B-spline elements**
Figure 17.　Initial parabolic cable of object 2 with 31 nodes
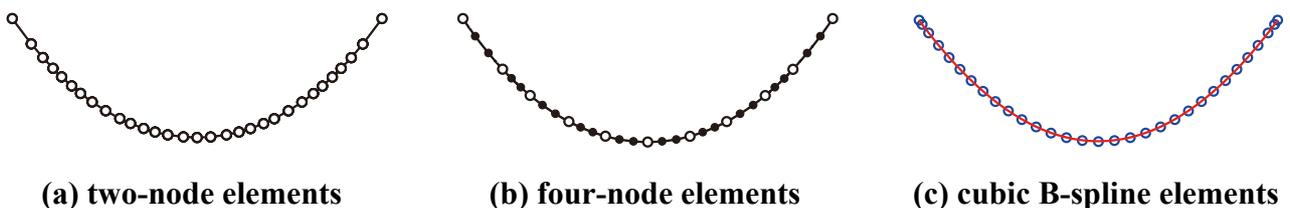
(a) two-node elements　　　**(b) four-node elements**　　　**(c) cubic B-spline elements**
Figure 18.　Final parabolic cable of object 2 with 31 nodes

Table 2. Identified results of parabolic cable

The number of nodes	Object	Element	The number of iterations	MSE[%]	Analysis time [s]
10 nodes	1	two-node four-node B-spline	14 times 26 times 25 times	1.0521 0.6397 0.3134	0.2424 0.1984 0.3075
	2	two-node four-node B-spline	23 times 85 times 42 times	0.8808 0.6589 0.1681	0.3816 0.6553 0.2186
31 nodes	1	two-node four-node B-spline	21 times 31 times 36 times	0.0979 0.0856 0.0838	0.9618 0.7811 1.4093
	2	two-node four-node B-spline	325 times 53 times 84 times	0.0949 0.0873 0.0860	14.3975 1.2747 3.2134

Performances of the analyses using conventional FEA as well as IGA with different number of elements are summarized in Table 2. It was clear that IGA performs better than conventional FEA in accuracy in all cases. However, the superiority of IGA in computation costs is not clear.

Conclusions

In this paper, we applied Isogeometric Analysis for self-equilibrium analysis of unstable cable structures and investigated its performances in accuracy as well as in efficiency. For all analysis cases in this paper both for catenary cables and parabolic cables, IGA is more accurate than conventional FEA. However, its performance on computational costs is not as clear as accuracy.

References

T.J.R. Hughes, J.A. Cottrell, Y. Bazilevs (2005) Isogeometric analysis: CAD, finite elements, NURBS, exact geometry and mesh refinement

J. Austin Cottrell, Thomas J. R. Hughes, Yuri Bazilevs (2009) Isogeometric Analysis: Toward Integration of CAD and FEA

Stefan Hartmann, David J. Benson, David Lorenz (2011) About Isogeometric Analysis and the new NURBS-based finite Elements in LS-DYNA, *8th European LS-DYNA Users Conference*

D.J. Benson, Y. Bazilevs, M.C. Hsu, T.J.R. Hughes (2010) Isogeometric shell analysis: Yhe Reissner-Mindlin shell, *Compt. Methods Appl. Mech. Engrg.* **199**, 276-289

Klaus-Jurgen Bathe (1995) Finite Element Procedures

Japan Society of Civil Engineers (2001) Fundamentals and Applications of cable space structures

Ken'ichi Kawaguchi (2011) generalized inverse matrix and application to structural engineering (Computational Engineering Series)

PARAMETER IDENTIFICATION OF FLUID VISCOUS DAMPERS

*Rita GRECO [1] and Giuseppe Carlo MARANO[2]

[1] DICATECH, Department of Civil Engineering, Environmental, Territory, Building and Chemical, Technical University of Bari, via Orabona 4 - 70125, Bari -Italy.

[2] DICAR, Department of civil Engineering and Architecture, Technical University of Bari, via Orabona 4 - 70125, Bari -Italy

*Corresponding author: rita.greco@poliba.it

This paper focuses on parameter identification of Fluid Viscous Dampers, comparing different existing literature models, with the aim to recognize ability of these models to match experimental loops under different test specimens. Identification scheme is developed evaluating the experimental and the analytical values of the forces experienced by the device under investigation. The experimental force is recorded during the dynamic test, while the analytical one is obtained by applying a displacement time history to the candidate mechanical law.

Identification procedure furnishes device mechanical parameters by minimizing a suitable objective function, which represents a measure of difference between analytical and experimental forces. To solve optimization problem, the Particle Swarm Optimization is adopted, and the results obtained under various test conditions are shown. Some considerations about the agreement of different models with experimental data are furnished, and the sensitivity of identified parameters of analyzed models against frequency excitation is evaluated and discussed..

Keywords: Fluid Viscous Damper, parameters identification, Kelvin-Voigt model, Particle Swarm Optimization

Introduction

In recent years, several devices have been proposed to reduce the effects of dynamic loads in civil structures and infrastructures. In this paper, the attention is focused on Fluid Viscous Dampers (FVD), generally viewed as passive dissipation elements [1], widely adopted in many civil engineering applications to reduce the vibration level and to increase structural protection level against wind and earthquake forces (see for instance [2],[3]). Among the most interesting features of viscous dampers, one should mention low maintenance costs, usability for several earthquakes without damage and viscous forces out-of-phase with the elastic ones.

Viscous dampers utilized in civil structures to control seismic, wind induced and thermal expansion motions, are usually arranged in one of the following configurations: a diagonal or chevron bracing element within a steel or concrete frame, as a part of the cable stays of long-span bridges, as a part of tuned mass dampers, as a part of a base isolation system to increase the energy dissipation and as a device to allow free thermal movements [4]. Viscous dampers can be efficiently used in the construction of new buildings or in retrofitting existing structures. The importance of viscous dampers in vibration control has increased thanks to their energy dissipation capability and wide range of applications.

A viscous fluid damper typically consists [1] of a piston within a damper housing, filled with a compound of silicone or similar type of oil. The fluid passes through several small orifices from one side of the piston to the other; therefore, the energy is dissipated through the concept of fluid orificing. The fluid damper produces a force that is not always proportional to velocity [5], depending on the type of orifice used. The orifice utilizes a series of passages to alter flow characteristics with fluid speed. The "fluid control orifice" provides forces proportional to , where α is a coefficient varying in the range [0.5 ÷ 1]. When α=1, the behavior of FVD is linear and in earthquake engineering applications this is the most desired circumstance. Actually, FVDs contain valves instead of the piston within orifices. These valves are opened once the transmitted force exceeds a certain design limit. However, the force produced by FVD is not proportional to velocity, and also in this case the valves provide forces proportional to .

Since the applications of viscous dampers are growing very fast, the exact recognition of their mechanical behavior is of primary importance to provide a reliable support to design an efficient seismic protection strategy. Current identification techniques for viscous dampers are mostly based

on parametric models. Although parametric identification techniques have been successfully used to identify viscous dampers, non-parametric identification techniques are more suitable in structural health monitoring, because the system characteristics may continuously vary over time, both quantitatively as well as qualitatively.

Several identification approaches, both parametric and nonparametric, are compared in [4], by using real data carried out from full-scale nonlinear viscous dampers, commonly used in large flexible bridges. About the parametric techniques, the capability of the Adaptive Random Search is explored in [4]: the authors solved an optimization problem in which the numerical values of the unknown model parameters were estimated by minimizing an objective function based on the normalized mean square error between the measured and identified damper responses, evaluated as displacement/velocity, and obtained integrating dynamic equilibrium equations of FVD constitutive law, under experimental applied force history.

In this field, also soft-computing techniques, fuzzy inference systems and neural networks have been applied to model a Magneto Rheological Fluid Damper [6][7]. Evolutionary computation methods, e.g., Genetic Algorithms (GA) [8][9], have been widely applied in parameter identification applications and many others. Among different nonlinear models, especially the Bouc-Wen has been identified thanks to its versatility. In [10], the GA was employed to identify a mechatronic system of unknown structure. In this framework, a real-coded GA has been recently adopted in [11] to identify a piezoelectric actuator, whose hysteretic behavior has been modeled by the Bouc-Wen nonlinear law. A magneto-rheological fluid damper behavior has been recognized by [12], with reference to a non-symmetric version of the original Bouc-Wen model and by using a real coded GA. The final algorithm is very similar to the GA, but its efficiency has been improved in virtue of a selection procedure embedded into crossover and mutation genetic operators. The GA has been widely adopted to fit the Bouc–Wen model to hysteresis loops experimentally obtained for composite materials [13], non-linear degrading structures [14], magneto-rheological fluid dampers [15][16][17] or bolted-welded connections [18]. In [19], a new method based on GA is developed to identify the Bouc–Wen model parameters from experimental hysteretic loops, obtained from periodic loading tests.

Among evolutive algorithms, the Particle Swarm Optimization (PSO) [20] has been recognized as a promising candidate in parameter identification. The PSO is based on the multi-agent or population based philosophy (the particles) which mimics the social interaction in bird flocks or schools of fish, by incorporating the search experience of individual agents. Moreover, the PSO is effective in exploring the solution space in a relatively small number of iterations. PSO has been used in the design of PID controllers [21] and electro-magnetic [22]. The PSO convergence characteristic was analyzed in [23], where algorithm control settings were also proposed. In [24], a PSO algorithm is employed using experimental force–velocity data, obtained from various operating conditions, to identify the model parameters of a magneto rheological fluid damper.

In [25] a parameter identification for basic and generalized Kelvin–Voigt and Maxwell models for FVD is carried out. The identification procedure developed by means of particle swarm optimization gives the best mechanical parameters by minimizing a suitable objective function that represents a measure of difference between analytical and experimental applied forces. Results are obtained under various test conditions, comparing the agreement of various models with experimental data.

This paper focuses on parameter identification of FVD: the identification process is developed comparing the experimental and the analytical values of the forces experienced by the device under investigation. The experimental value of the force is recorded during the dynamic test, while the analytical one is obtained by applying the time history of displacements to the candidate mechanical law. In this way, a measure of the "distance" between experimental and analytical results is introduced, as the integral of the difference along the whole experiment. The optimal parameter set is thus derived by minimizing this distance using an evolutionary algorithm. For the parametric identification of FVD, the authors adopt an evolutive algorithm, the Particle Swarm Optimization. Different analytical models, characterized by increasing complexity, are considered and then are identified. The sensitivity against test conditions is also assessed.

The next of the paper is organized as follows: in section 2 there is a selection of models adopted in this study for FVD modeling; in section 3, the identification scheme is posed and in section 4, some remarks of PSO algorithm are given. Moreover, in section 5 some specifications of experimental tests are furnished; section 6 reports the results of identified parameters, which are discussed in section 7. Some conclusions are finally given in section 8.

MECHANICAL MODELS FOR FLUID-VISCOUS DAMPERS

System identification involves creating a model for a system that, with the same input as the original system, the model will produce an output that matches the original system output with a certain degree of accuracy. The input or excitation of the system and model, and their corresponding outputs, are used to create and tune the model until a satisfactory degree of accuracy is reached.

The application of non-classical methods for the parametric identification of viscous dampers requires: (i) the definition of an appropriate single-degree-of-freedom mechanical model and (ii) the formalization of the objective (or cost) function to be minimized. This section deals with the first aspect.

Generally, the system to be identified could be modeled by physical laws that reflect the dynamics of the system. A model created by laws, which reflect the physical properties of the system is called a white-box model. However, creating a white box model for real-world (complex) systems is a challenging task.

In structural applications, the selection of a proper model for FVD plays a central role to predict the real structural response after the identification. Generally, the description of FVD requires a suitable mechanical model, made of a set of springs and dashpots appropriately connected each other. In this study, different classical and generalized mechanical models are selected to identify a viscous device using experimental data. The main difference between classical and generalized models is that the generalized one incorporates a nonlinearity in spring and viscous elements; in addition, the resistant forces of generalized models have fractional exponential coefficients.

Linear viscous model
The simplest way to model a velocity dependent mechanical law is by means of the standard linear viscous model. The equation of the motion of a FVD modeled in this way and subject to a time-varying force *p* is:

$$m\ddot{y} + C\dot{y} = p \tag{1}$$

This basic model has the main advantage to be extremely simple, but sometimes it is too poor for a reasonable representation of real mechanical behavior. For this reason, it has been updated by the non-linear viscous model that depends on a fractional exponent of the velocity instead of a simple linear relationship. Generalized non-linear viscous model is described below:

Generalized viscous model
 It is a two parameters model proposed by Constantinou [26], [27] whose law is:

$$m\ddot{y} + C\,\mathrm{sgn}(\dot{y})\left|\dot{y}\right|^{\alpha} = p \tag{2}$$

where α is the damping term exponent, whose value lies between 0 and 1. Various mechanical behaviors are associated to different values of α. For instance, if $\alpha = 1$ the linear viscous damping law corresponds; if $\alpha = 0$ the dry friction appears (consequently, the force increases quickly for small velocity values, and becomes almost constant for large velocity values). This damping law has been widely adopted by various authors thanks to its ability in structural behavior modeling. For example, Lin and Chopra [28] make use of this constitutive law in the investigation of the earthquake induced response. In addition, this law is adopted in many structural computer codes.

However, experimental studies demonstrated that the resistance force of viscous dampers depends not only on damper velocity, but also on damper deformation. This mechanical property may be mathematically modeled connecting a spring element and a viscous element, respectively. If these two elements are connected in parallel, the family of Kelvin-Voigt models is obtained. For example, if a linear spring is connected in parallel with the simple linear dashpot, the basic Kelvin-Voigt

model is derived. When non-linear springs are connected with generalized non-linear viscous models, other behaviors are obtained. In [29] Terenzi investigated linear and parabolic models for the elastic force ψ_e:

$$\psi_e = K_1 y \tag{3}$$

$$\psi_e = K_2 y^2 + K_1 y + K_0 \tag{4}$$

where K_1 is the elastic stiffness, K_2 and K_0 are two constants. In [29], the authors stated that the parabolic function reproduces better the shape of the test cycles, but the linear function may be preferable, because it is simpler and yields a comparable energy balance.

Generalized viscous – linear elastic model
By combining Eq.(2) and Eq.(3), the equation of motion of a generalized Kelvin-Voigt model, subjected to a time-varying force p is derived:

$$m\ddot{y} + C \operatorname{sgn}(\dot{y})|\dot{y}|^\alpha + K_1 y = p \tag{5}$$

Generalized viscous – quadratic elastic model
In this model, the parabolic form in Eq.(4) is considered without the constant K_0:

$$m\ddot{y} + C \operatorname{sgn}(\dot{y})|\dot{y}|^\alpha + K_1 y + K_2 y^2 = p \tag{6}$$

IDENTIFICATION: OPTIMIZATION PROBLEM

The second step of parameter identification requires the formalization of a suitable objective function to be minimized.
The model parameters $\mathbf{x}$ of the viscous damper are identified by solving the following single-objective optimization problem:

$$\min_{\mathbf{x}} \{f(\mathbf{x})\}$$
$$\text{s.t. } \mathbf{x}^l \le \mathbf{x} \le \mathbf{x}^u$$

where $\mathbf{x} = \{x_1,\ldots,x_j,\ldots,x_n\}$ is a set of real parameters (in this case $\mathbf{x}$ collects the mechanical model parameters), $\mathbf{x}^l = \{x_1{}^l,\ldots,x_j{}^l,\ldots,x_n{}^l\}$ and $\mathbf{x}^u = \{x_1{}^u,\ldots,x_j{}^u,\ldots,x_n{}^u\}$ are lower and upper bounds of $\mathbf{x}$, respectively. The solution that minimizes the objective function (OF) $f(\mathbf{x})$ is $\mathbf{x}^*$.

The following integral is assumed as measure to define the OF in the identification problem:

$$f(\mathbf{x}) = \frac{1}{\sigma_{p_m}(t_{end} - t_{start})} \int_{t_{start}}^{t_{end}} (p_m - p_e(\mathbf{x})) dt \tag{7}$$

where t_{start} and t_{end} are the start and end time records, $p_m(t)$ is the force measured, while $p_e(t)$ is the force estimated. This is obtained by numerical differentiation of experimental displacement time history with a 3rd order algorithm to limit numerical noise. One should point out that the evaluation of this OF is extremely computational cheap if compared with alternative approaches, in which the duality of starting from an experimental force leads to the theoretical displacement, obtained by integration as a solution of the differential equation. The optimization problem is solved by Particle Swarm Optimization (PSO).

Experimental studies

Test apparatus

The 750 kN viscous damper was tested at SISMALB srl laboratory in Taranto, Italy. The test setup (Figure 1) consists of a high resistance steel frame to withstand loads of tension and compression of 2200 KN. The device is anchored to the structure by means of a pin, and is stilled to the servant cylinder by means of a threaded connection. The movements are generated by a servant cylinder of 1400 KN, controlled in force and/or displacement. Between the servant cylinder and the device a load cell of 2500 KN is located, which acquires the forces applied to the device during the entire duration of the experiment. In a displacement imposed test, the device movements are controlled by a transducer mounted on the device. The control and data acquisition system is able to generate a real time analysis of device displacements, by instantaneously variation of applied forces by the servant cylinder by means of a computer automatic control hydraulic pressure system. The displacement time history can be imposed with different laws, from sinusoidal, triangular, or through a generator step of generic ones. This system is able to control applied forces in real time according to the imposed displacement or force imposed test. Acquiring system has 30 channels and can command 2 actuators at the same time.

Table 1 shows the design characteristics of the tested FVD.

Figure 1. View of the viscous test machine and fluid viscous damper
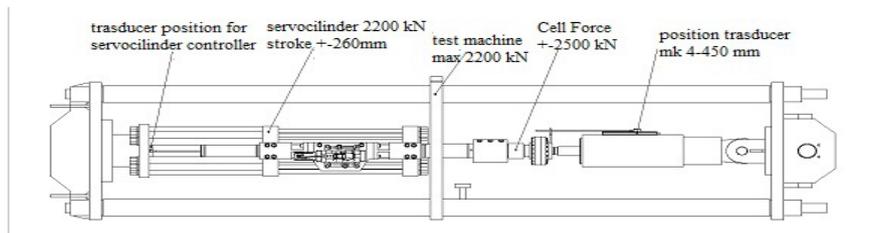
Figure 2. A photo of the test apparatus with the fluid viscous damper

Table 1 Fluid Viscous Damper Design Condition

F [kN]	Stroke [mm]	C [kN/(mm/s)]	V [mm/s]	α
750	± 100	406.24	460	0.1

Test cases

Four experiments were performed to obtain dynamic response of the viscous damper. The experiments were designed to determine the dynamic characteristics of the damper at varying velocities and to evaluate the effective energy dissipation of the device. The damper was subjected to multiple sets of monotonic sinusoidal excitations, at peak velocities of 92 mm/s, 230 mm/s, 460 mm/s (% refers to design velocity 460 mm/s) The first three tests have a 3-cycle excitation period, while the fourth test (energy dissipation test) has a 10-cycle period. The test specifications are summarized in Table 2.

Table 2. Fluid viscous damper test condition

No.	Test Type	Load (kN)	Test stroke (±mm)	Velocity (mm/s)	Cycle
1		750	20	92 (20%)	3
2	Constitutive law test	750	20	230 (50%)	3
3		750	20	460 (100%)	3
4	Damping efficiency test	750	17	460 (100%)	10

Parametric identification

For the evaluation of optimal values of the unknown parameters in Equations (1), (2), (5), (6) the parametric identification performed by PSO, was applied with a population size N=50 and maximum number of iterations L=100. The parametric identification has been performed by solving the single-objective optimization problem, whose objective function is given by Equation (7). The algorithms have been performed fifty times, and the best solution has been carried out as the final identification result.

Identification results

This subsection shows the identified parameter values that best fit the test results for the four analyzed models. Table 3, Table 4, Table 5 and Table 6 show the best (Min), worst (Max), mean and standard deviation (Std) values of the OF obtained under different numerical tests, for the four analyzed models. Data are represented also in Figure 3.

Table 3. Objective Function results obtained from the PSOA using the linear viscous mechanical model for four different experimental tests

Mechanical Model: *Linear viscous*				
Test	**Mean**	**Max**	**Min**	**Std**
Test 1	0.324322	0.324322	0.324322	0
Test 2	0.363997	0.363997	0.363997	2.8E-16
Test 3	0.272685	0.272685	0.272685	1.68E-16
Test 4	0.297829	0.297829	0.297829	1.68E-16

Table 4 Objective Function results obtained from the PSOA using the Generalized viscous mechanical model for four different experimental tests

Mechanical Model: *Generalized viscous*				
Test	Mean	Max	Min	Std
Test 1	0.254494	0.254494	0.254494	4.26E-14
Test 2	0.332256	0.332257	0.332256	1.39E-07
Test 3	0.264244	0.26426	0.264243	2.99E-06
Test 4	0.28234	0.28234	0.28234	2.45E-09

Table 5 Objective Function results obtained from the PSOA using the Generalized viscous – linear elastic mechanical model for four different experimental tests

Mechanical Model: *Generalized viscous- linear elastic*				
Test	Mean	Max	Min	Std
Test 1	0.162356	0.163188	0.162077	0.000298
Test 2	0.203976	0.204116	0.203949	3.45E-05
Test 3	0.153384	0.153388	0.153384	7.23E-07
Test 4	0.127699	0.127699	0.127699	1.41E-12

Table 6. Objective Function results obtained from the PSOA using the Generalized viscous – quadratic elastic mechanical model for four different experimental tests

Mechanical Model: *Generalized viscous- quadratic elastic*				
Test	Mean	Max	Min	Std
Test 1	0.173636	0.254494	0.158448	0.022962
Test 2	0.208454	0.21712	0.203949	0.006284
Test 3	0.160706	0.26426	0.153025	0.026845
Test 4	0.12752	0.127699	0.126207	0.00049

Tables 7-10 show the values of identified parameters obtained for each mechanical model, where mean, max, min and std indicate the values which correspond to mean, max, min and std of OF in previous tables. Results of identification are represented also in Figures 4-7.

Table 7. Values of mechanical parameters obtained in four different test types, using the linear viscous mechanical model of FVD

Mechanical Model: *Linear viscous*				
Parameters	Test Type N.1	Test Type N.2	Test Type N.3	Test Type N.4
	v=92mm/s	v=230mm/s	v=460mm/s	v=460mm/s
M(mean) - [kg]	0	0	0	0
M(max) - [kg]	0	0	0	0
M(min) - [kg]	0	0	0	0
C(mean) - [kN/(mm/s)]	6.308518	9.955068	2.950677	3.599261
C(max) - [kN/(mm/s)]	6.308518234	9.955068455	2.95067697	3.599260974
C(min) - [kN/(mm/s)]	6.308518234	9.955068455	2.95067697	3.599260974
C(std)-[kN/(mm/s)]	3.32E-14	0	1.93E-15	3.15E-14

Table 8. Values of mechanical parameters obtained in four different test types, using the fractional viscous mechanical model of FVD

Mechanical Model: *Fractional viscous*				
Parameters	Test Type N.1 v=92mm/s	Test Type N.2 v=230mm/s	Test Type N.3 v=460mm/s	Test Type N.4 v=460mm/s
M(mean) - [kg]	1.75E-14	1.45E-11	0	0
M(max) - [kg]	8.74059E-13	7.26404E-10	0	0
M(min) - [kg]	0	0	0	0
M(std) - [kg]	1.24E-13	1.03E-10	0	0
C(mean) - [kN/(mm/s) $^\alpha$]	321.4664	101.8108	20.93332	60.02495
C(max) - [kN/(mm/s) $^\alpha$]	321.4663828	102.5398101	22.44238445	60.02544199
C(min) - [kN/(mm/s) $^\alpha$]	321.4663828	101.058709	20.75427774	60.01439848
C(std) - [kN/(mm/s)$^\alpha$]	1.05E-10	0.254748	0.284589	0.001661
α(mean)	0.121515	0.456479	0.647184	0.472998
α(max)	0.121514934	0.458176548	0.648755563	0.473033897
α(min)	0.121514934	0.454813372	0.634798957	0.472996579
α(std)	6.82E-14	0.00058	0.002352	5.61E-06

Table 9. Values of mechanical parameters obtained in four different test types, using the fractional viscous –linear elastic mechanical model of FVD

Mechanical Model: Fractional viscous- linear elastic				
Parameters	Test Type N.1 v=92mm/s	Test Type N.2 v=230mm/s	Test Type N.3 v=460mm/s	Test Type N.4 v=460mm/s
M(mean) - [kg]	2.034798	1.820115	0.000221	4.88E-12
M(max) - [kg]	2.198171813	1.904731018	0.004116766	2.43898E-10
M(min) - [kg]	1.810732517	1.602184231	0	0
M(std) -[kg]	0.089231	0.077191	0.000761	3.45E-11
C(mean) - [kN/(mm/s) $^\alpha$]	52.61233	24.70355	2.924908	3.575181
C(max) - [kN/(mm/s) $^\alpha$]	58.98786914	24.94752828	2.925077333	3.575181353
C(min) - [kN/(mm/s) $^\alpha$]	48.52647178	24.41265785	2.924898421	3.575181353
C(std) - [kN/(mm/s) $^\alpha$]	3.512028	0.125295	3.31E-05	8.73E-11
α(mean)	0.510677	0.768888	1	1
α(max)	0.52834009	0.772247572	1	1
α(min)	0.484805623	0.766278406	1	1
α(std)	0.014462	0.001346	0	0
K_0(mean) - [kN/mm]	70.59402	41.37259	9.3253	13.89884
K_0(max) - [kN/mm]	75.50983233	42.59740601	10.00317587	13.89884259
K_0 (min) - [kN/mm]	63.74648345	38.15868057	9.286881568	13.89884253
K_0(std) - [kN/mm]	2.749654	1.129993	0.132487	9.43E-09

Table 10. Values of mechanical parameters obtained in four different test types, using the fractional viscous – quadratic elastic mechanical model of FVD

Mechanical Model: Fractional viscous- quadratic elastic				
Parameters	Test Type N.1 v=92mm/s	Test Type N.2 v=230mm/s	Test Type N.3 v=460mm/s	Test Type N.4 v=460mm/s
M(mean) - [kg]	1.528561	1.22022	0.007768	1.00E-15
M(max) - [kg]	2.180005282	1.921826379	0.189771629	4.99811E-14
M(min) - [kg]	0	0	0	0
M(std) -[kg]	0.861915	0.886947	0.032415	7.07E-15
C(mean) - [kN/(mm/s) ^ α]	67.00402	25.31663	4.324775	3.574265
C(max) - [kN/(mm/s) ^ α]	321.4663828	26.52458809	22.43076879	3.57518146
C(min) - [kN/(mm/s) ^ α]	48.53968825	24.39242143	2.924898421	3.567545202
C(std) - [kN/(mm/s) ^ α]	52.75564	0.696476	4.821501	0.002507
α(mean)	0.480449	0.764121	0.972133	1
α(max)	0.527975613	0.771759295	1	1
α(min)	0.121514934	0.753137579	0.634903225	0.999999994
α(std)	0.077104	0.005187	0.095558	7.80E-10
K_1(mean) - [kN/mm]	55.07667	32.25728	10.0696	13.94087
K_1(max) - [kN/mm]	74.9994301	42.88675282	42.85875522	14.24910384
K_1(min) - [kN/mm]	0	13.76515596	0	13.89884243
K_1(std) - [kN/mm]	26.29288	13.44016	6.269791	0.114977
K_2(mean) - [kN/mm^2]	0.007748	0	0.002121	0.010372
K_2(max) - [kN/mm^2]	0.082770673	0	0.036533794	0.086436416
K_2(min) - [kN/mm^2]	0	0	0	0
K_2 (std) - [kN/mm^2]	0.023509	0	0.008492	0.028374

COMPARISON OF HYSTERESIS LOOPS PREDICTED BY VARIOUS MODELS

In figures 3-6 the experimental hysteresis loops of the damper under investigation are compared with those simulated by the selected models previous described, for load application velocities V_1, V_2 , V_3 and V_4. More precisely, in figures 3 and 4, the relationships between displacement and forces are shown, whereas figures 5 and 6 illustrate the relationships between force and velocity. The dotted lines represent the experimental loops, while the solid lines are the theoretical loops obtained by using the identified parameters for each assessed model.

From these plots one can notice that the experimental and theoretical loops have exactly the same relative displacement (and velocity), whereas the damper force of the theoretical loop is computed according to each model. The experiment loops in Figures 3 and 4 show that, under harmonic excitation, the hysteresis loop of the damper changes when load application velocity increases. The comparison between theoretical and simulated loops points out that the simulated results obtained by the *generalized viscous – linear elastic* model ((b) in figure 3) match well with the experimental loops under all the excitation frequencies. The agreement of this model with experimental loop is better with respect to the linear *viscous elastic one* ((a) in figure 3). On the other hand, the other

analyzed models lead to elliptic hysteresis loops. For this reason, these cannot match well with the experimental loops for all the frequencies, because the loop changes its shape from low to high frequencies. For example, the *linear viscous model* underestimates the force for all frequencies and especially at low frequency.

With reference to *generalized viscous – linear elastic* ((c) in figure 4) and *generalized viscous – quadratic elastic* ((d) in figure 4) models, one can observe a good match with experimental loops for all velocities of the load application. The third and the fourth models predict well the force; in effect, one should consider another aspect, i.e. the area of the loop, which represents the amount of dissipated energy in the cycle. The plots point out that the *generalized viscous – linear elastic* model overestimates the amount of dissipated energy for all velocities of load application. On the contrary, the *generalized viscous – quadratic elastic* predicts fine the dissipated energy, especially for high load application velocity. The same observation can be pointed out with reference to *generalized viscous – quadratic elastic* model.
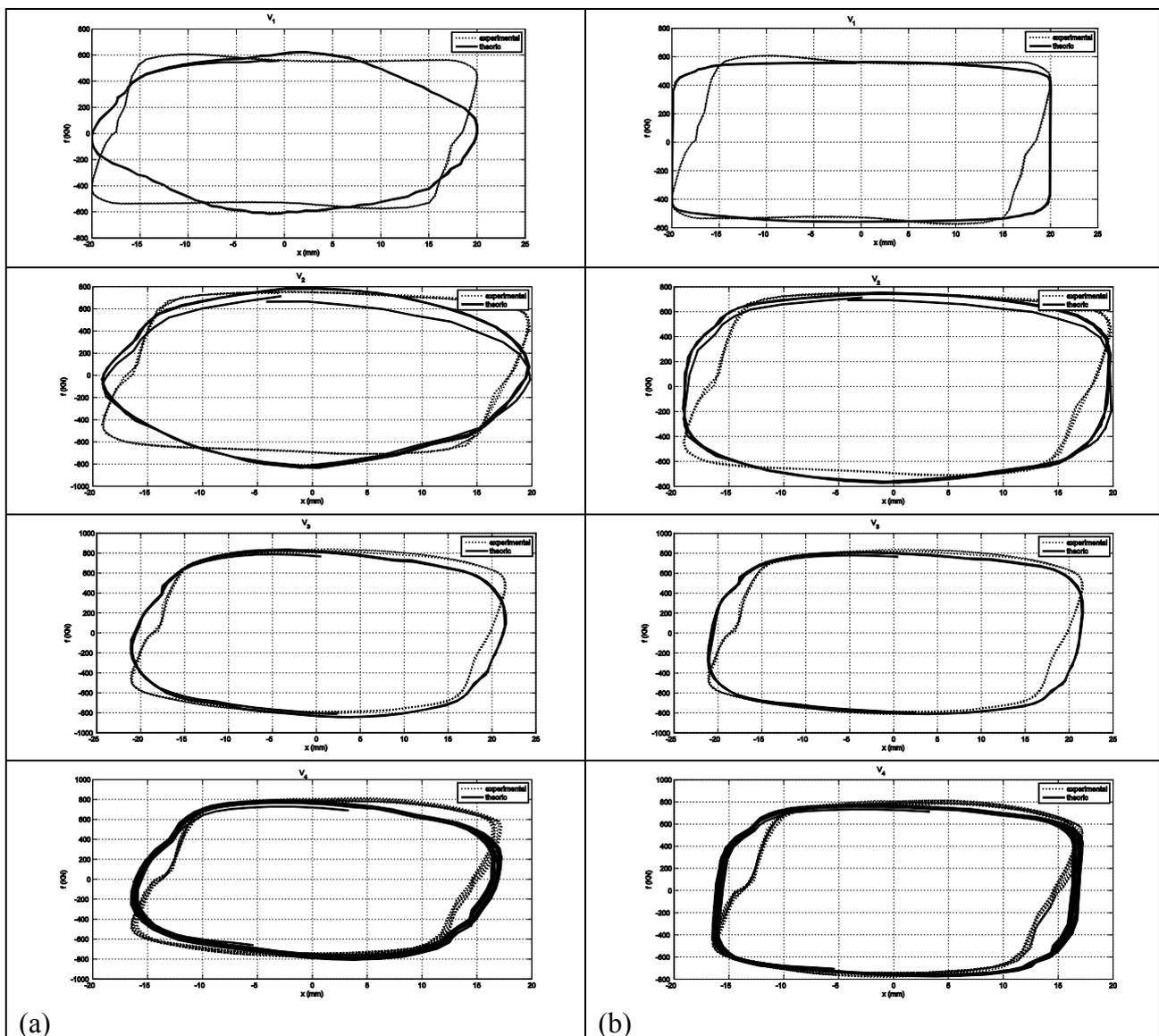
Figure 3. Comparison between theoretical and experimental force- displacement relationship: a) Linear viscous model, b) Generalized viscous model.

In figures 5 and 6 the relationships between the force and the velocity are shown. The first and the second model don't predict absolutely the experimental force -velocity experimental loop, wearers the third and the fourth model match satisfactorily the experimental loop, especially for high excitation frequency.

Because the matching of the identified model with the experimental ones depends on the excitation frequency, it is interesting to evaluate the sensitivity of identified parameters against the frequency excitation. For this purpose, for each model, the mean value $\bar{p}$ of each identified parameter p, evaluated from the four tests is extrapolated; the range of variation $\Delta \bar{p} = \bar{p}_{max} - \bar{p}_{min}$ and the ratio $\Delta \bar{p} / \bar{p}$ are furnished (table 11-14) to quantity the variability of mentioned parameters with respect to the test conditions. From numerical data in tables 11-14, one can deduce that, except for the *linear viscous model*, the parameter C exhibits the highest variability against the velocity of the external excitation application. Anyway, all analyzed models present almost a comparable variability of involved parameters.
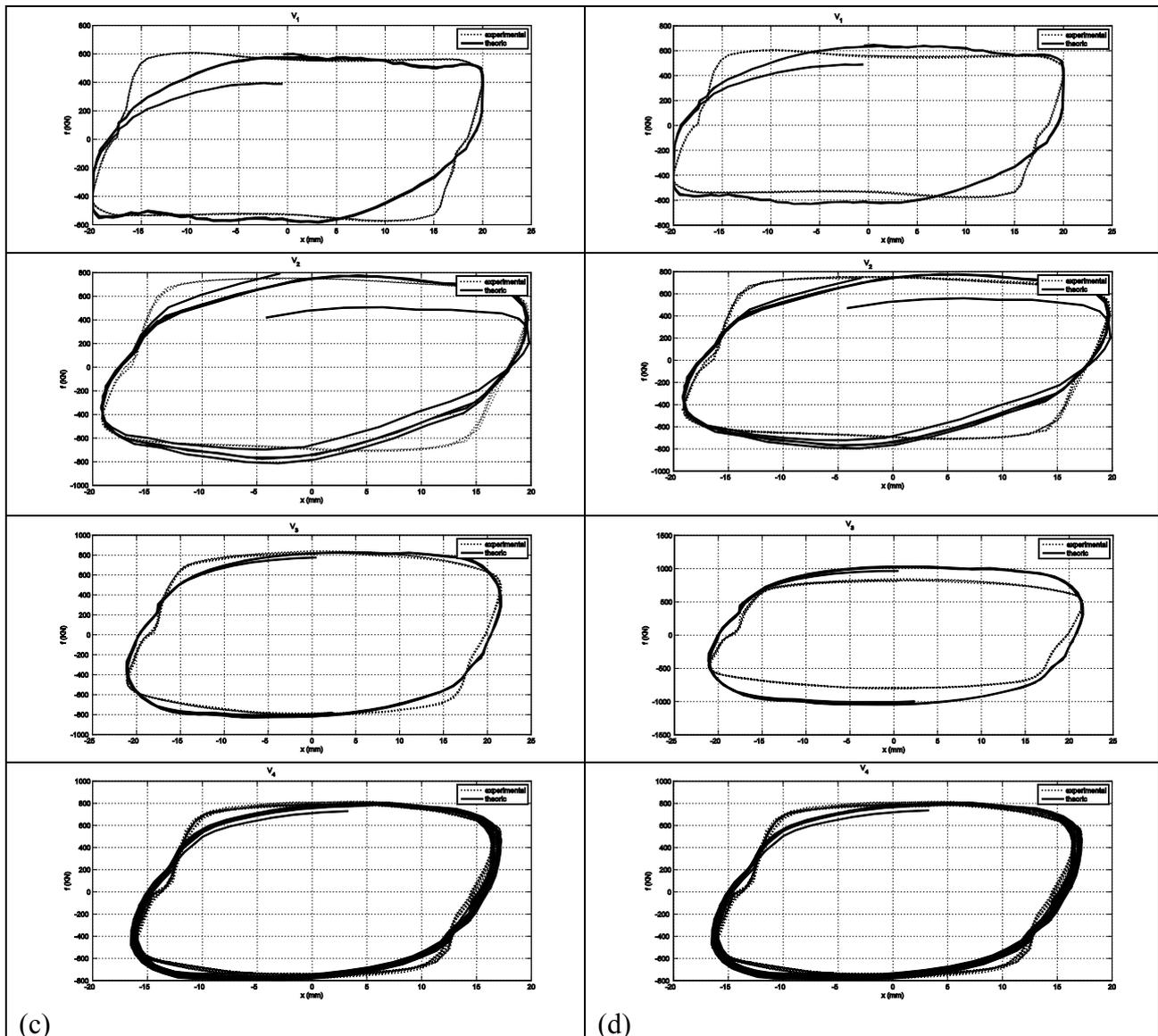
Figure 4. Comparison between theoretical and experimental force- displacement relationship: c) Generalized viscous- linear elastic, d) Generalized viscous- quadratic elastic.
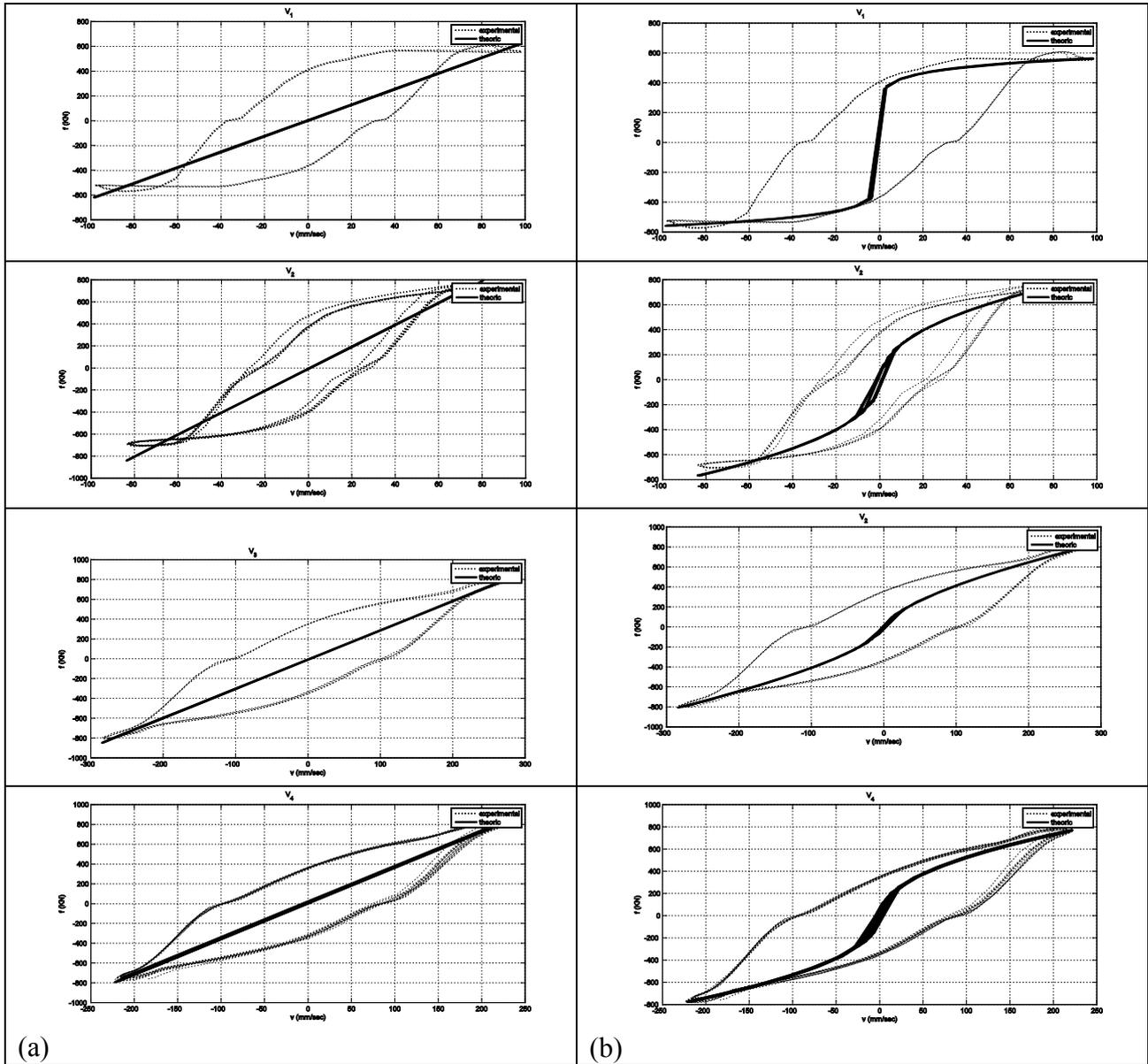
Figure 5. Comparison between theoretical and experimental force- velocity relationship: a) Linear viscous model, b) Generalized viscous model.

Table 11. Parameters sensitivity of Linear viscous mechanical model

$\bar{M}$ [kg]	0
$\Delta\bar{M}$	0
$\Delta\bar{M}\,/\,\bar{M}$	0
$\bar{C}$ [kN/(mm/s)]	5,703381
$\Delta\bar{C}$	7,004391
$\Delta\bar{C}\,/\,\bar{C}$	1,228112

(c)

(d)
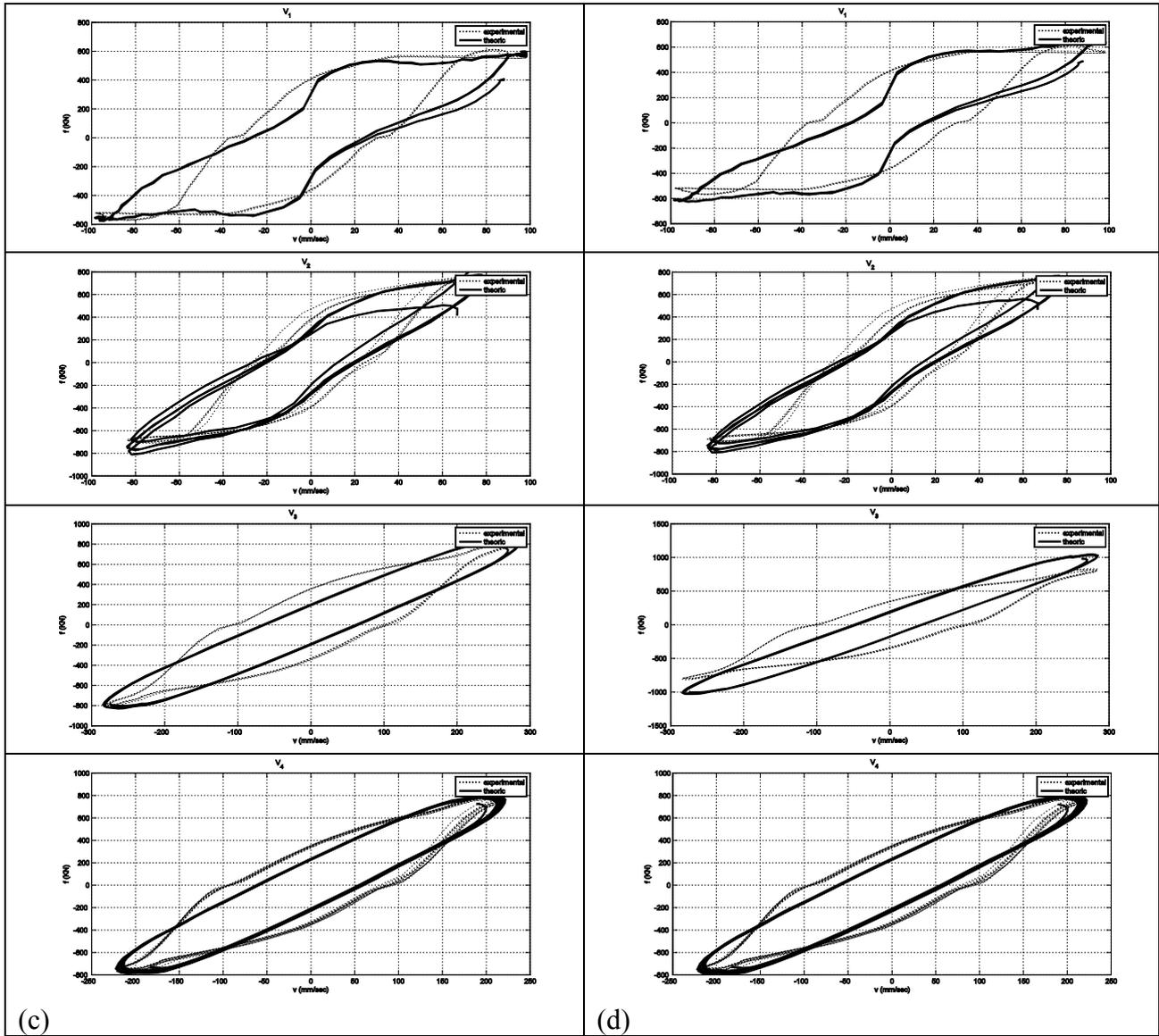
Figure 6. Comparison between theoretical and experimental force- velocity relationship: c) Generalized viscous- linear elastic, d) Generalized viscous- quadratic elastic.

Table 12. Parameters sensitivity of generalized viscous mechanical model

$\overline{C}$ - [kN/(mm/s) ^ α]	126,0589
$\Delta\overline{C}$	300,5331
$\Delta\overline{C} / \overline{C}$	2,384069
$\overline{\alpha}$	0,424544
$\Delta\overline{\alpha}$	0,525669
$\Delta\overline{\alpha} / \overline{\alpha}$	1,238197

Table 13: Parameters sensitivity of Generalized viscous- linear elastic mechanical model

$\bar{M}$ [kg]	0,963784
$\Delta\bar{M}$	2,03E+00
$\Delta\bar{M} / \bar{M}$	2,11E+00
$\bar{C}$ [kN/(mm/s)) ^ α]	20,95399
$\Delta\bar{C}$	49,68742
$\Delta\bar{C} / \bar{C}$	2,371263
$\bar{\alpha}$	0,819891
$\Delta\bar{\alpha}$	0,489323
$\Delta\bar{\alpha} / \bar{\alpha}$	0,596815
$\bar{K}_0$ - [kN/mm]	33,79769
$\Delta\bar{K}_0$	61,26872
$\Delta\bar{K}_0 / \bar{K}_0$	1,812808

Table 14: Parameters sensitivity of Generalized viscous- quadratic elastic mechanical model

$\bar{M}$ [kg]	0,689137
$\Delta\bar{M}$	1,53E+00
$\Delta\bar{M} / \bar{M}$	2,22E+00
$\bar{C}$ [kN/(mm/s) ^ α]	25,05492
$\Delta\bar{C}$	63,42976
$\Delta\bar{C} / \bar{C}$	2,531628
$\bar{\alpha}$	0,804176
$\Delta\bar{\alpha}$	0,519551
$\Delta\bar{\alpha} / \bar{\alpha}$	0,646066
$\bar{K}_1$ - [kN/mm]	27,83611
$\Delta\bar{K}_1$	45,00707
$\Delta\bar{K}_1 / \bar{K}_1$	1,616859
$\bar{K}_2$ - [kN/mm^2]	0,00506
$\Delta\bar{K}_2$	0,010372
$\Delta\bar{K}_2 / \bar{K}_2$	2,049701

Conclusions

This study concentrates on classical and generalized mechanical models for FVD. The focal difference between classical and generalized models is that the generalized ones incorporate nonlinearity in spring and viscous elements; in addition, the resistant forces in generalized models have fractional exponential coefficients. To evaluate the effectiveness of diverse models to catch the hysteretic behavior of real FVDs, diverse analytical models have been identified on the basis of experimental tests. The identification procedure is performed comparing the experimental and the analytical values of the forces experienced by the device under investigation. The experimental forces have been recorded during the dynamic test and the analytical ones have been evaluated by imposing the time history of displacement to the candidate mechanical law. The parametric identification of a real FVD has been developed by Particle Swarm Optimization. The identification process furnishes the best mechanical parameters by minimizing the difference between analytical and experimental applied forces. Four experiments have been performed to obtain the dynamic response of the viscous damper under investigation, varying the velocity of the load application.

The results show that the analytical results obtained by the generalized viscous − linear elastic model match well the experimental loops, under all the excitation frequencies, better with respect the linear viscous elastic one. Moreover, with reference to generalized viscous − linear elastic and generalized viscous − quadratic elastic it has been observed a good match with experimental loops for all velocities of the load application. The generalized viscous − linear elastic model and the generalized viscous − quadratic elastic model one predict well the force, but the generalized viscous − linear elastic overestimates the amount of dissipated energy for all velocities of the load application. On the contrary, the generalized viscous − quadratic elastic predicts well the energy dissipated, especially for high velocity of load application. The same observation can be made with reference to generalized viscous − quadratic elastic model. Moreover, the sensitivity of identified parameters against the frequency excitation has been investigated. Results showed that, except for the linear viscous model, the parameter C exhibits the highest variability against the velocity of the external excitation application. Anyway, all analyzed models present almost a comparable variability of involved parameters.

Acknowledgement

References

[1] Soong TT, Dargush GF (1997) Passive energy dissipation systems in structural engineering. John Wiley & Sons, New York.
[2] Marano GC,Trentadue F, Greco R (2007) Stochastic optimum design criterion for linear damper devices for seismic protection of buildings. Structural and Multidisciplinary Optimization *33*:441-455
[3] Marano GC, Trentadue F, Greco R (2007) Stochastic optimum design criterion of added viscous dampers for buildings seismic protection. Structural Engineering and Mechanics 25(1): 21-37
[4] Yun HB, Tasbighoo F, Masri SF, Caffrey JP, Wolfe RW, Makris N, Black C (2008) Comparison of Modeling Approaches for Full-scale Nonlinear Viscous Dampers. Journal of Vibration and Control 14:51-76
[5] Constantinou MC and Symans MD (1993) Experimental study of seismic response of buildings with supplemental fluid dampers. Structural Design of Tall Buildings 2:93–132
[6] Schurter KC, Roschke PN (2000) Fuzzy modeling of a magnetorheological damper using ANFIS.Proceedings of the Ninth IEEE International Conference on Fuzzy Systems, San Antonio, TX, 122–127.
[7] Wang DH, Liao WH (2001) Neural network modeling and controllers for magnetorheological fluid dampers. Proceedings of the 2001 IEEE International Conference on Fuzzy Systems, Melbourne, Australia, 1323–1326.
[8] Man KF, Tang KS, Kwong S (1996) Genetic algorithms: concepts and applications. IEEE Trans. Ind. Eletron 43:519–534
[9] Goldberg DE (1989) Genetic Algorithms in Search, Optimization and Machine Learning. Addison-Wesley, Reading, MA
[10] Iwasaki M, Miwa M, Matsui N (2005), GA-based evolutionary identification algorithm for unknown structured mechatronic systems. IEEE Trans. Ind. Electron 52: 300–305.

[11] Ha JL, Kung YS, Fung RF, Hsien SC (2006) A comparison of fitness functions for the identification of a piezoelectric hysteretic actuator based on the real-coded genetic algorithm. Sensors and Actuators A 132:643-650

[12] Kwok NM, Ha QP, Nguyen MT, Li J, Samali B (2007) Bouc-Wen model parameter identification for a MR fluid damper using computationally efficient GA. ISA Transactions 46:167-179

[13] Hornig KH, Flowers GT (2005) Parameter characterization of the Bouc–Wen mechanical hysteresis model for sandwich composite materials using real coded genetic algorithms. The International Journal of Acoustics and Vibration 10(2):73–81

[14] Ajavakom N, Ng CH, Ma F (2008) Performance of nonlinear degrading structures: identification, validation, and prediction. Computer and Structure 86:652–662

[15] Giuclea M, Sireteanu T, Stancioiu D, Stammers CW (2004a) Modelling of magnetorheological damper dynamic behavior by genetic algorithms based inverse method. Proceedings of the Romanian Academy Series A 5(1):55–63

[16]   Giuclea M, Sireteanu T, Stancioiu D, Stammers CW, (2004b) Model parameter identification for vehicle vibration control with magnetorheological dampers using computational intelligence methods. Proceedings of the institution of mechanical engineers. Part I: Journal of Systems and Control Engineering 218(7):569–581

[17] Kwok NM, Ha QP, Nguyen MT, Li J, Samali B (2007) Bouc–Wen model parameter identification for a MR fluid damper using computationally efficient GA. ISA Transactions, 46(2):167–179

[18] Charalampakis AE, Koumousis VK (2008) Identification of Bouc–Wen hysteretic systems by a hybrid evolutionary algorithm. Journal of Sound and Vibration 314:571–585

[19] Sireteanu T, Giuclea M, Mitu AM (2010) Identification of an extended Bouc–Wen model with application to seismic protection through hysteretic devices. Computational Mechanics 45(5):431–441

[20] Kennedy J, and Eberhart R C (1995) Particle Swarm Optimization. Proceedings of IEEE International Conference on Neural Networks, Perth, Australia 1942–1948

[21] Gaing ZL (2004) A particle swarm optimization approach for optimum design of PID controller. AVR system, IEEE Trans. Energy Conversion 19:384–391

[22] Ciuprina G, Loan D, Munteanu I (2002) Use of intelligent-particle swarm optimization in electromagnetics. IEEE Trans. Magn. 38:1037–1040

[23] Trelea IC (2003). The particle swarm optimization algorithm: convergence analysis and parameter selection. Inf. Process. Lett. 85:317–325.

[24] Kwok MN, Ha QP, Nguyen TH, Li J, Samali B (2006) A novel hysteretic model for magnetorheological fluid dampers and parameter identification using particle swarm optimization. Sensors and Actuators A 132: 441–451

[25] Greco R, Marano GC (2013) Identification of parameters of Maxwell and Kelvin-Voigt generalized models for fluid viscous dampers. Journal of Vibration and Control doi: 10.1177/1077546313487937.

[26] Constantinou MC and Symans MD (1993) Experimental study of seismic response of buildings with supplemental fluid dampers. Structural Design of Tall Buildings 2:93–132

[27] Constantinou MC, Symans MD, Tsopelas P, Taylor DP (1993), Fluid viscous dampers in applications of seismic energy dissipation and seismic isolation. in Proceedings of ATC 17-1 Seminar on Seismic Isolation, Passive Energy Dissipation and Active Control V. 2:581–592

[28] Lin WH, Chopra AK (2003) Asymmetric One-Storey Elastic Systems with Non-linear Viscous and Viscoelastic Dampers.Earthquake Response, Earthquake Engineering and Structural Dynamics 32:555-577

[29] Terenzi G (1999) Dynamics of SDOF Systems with Nonlinear Viscous Damping. Journal of Engineering Mechanics 125:956-963

[30] Sh, Y, Eberhart R (1998) A modified particle swarm optimizer. IEEE World Congress on Computational Intelligence, Anchorage, AK, USA, May 4-9

[31] Ratnaweera A, Halgamuge SK, Watson HC (2004) Self-Organizing Hierarchical Particle Swarm Optimizer With Time-Varying Acceleration Coefficients. IEEE Transactions on Evolutionary Computation 8:240-255

Generalized Irwin plastic zone correction of a sub-interface Zener-Stroh crack in a coating-substrate system

J. Zhuang, and *Z. M. Xiao

School of Mechanical and Aerospace Engineering

Nanyang Technological University, Singapore

*Corresponding author: mzxiao@ntu.edu.sg

Abstract

Elastic-plastic stress analysis of a Zener-Stroh crack paralleling to the interface of a coating-substrate system has been carried out in this work. The sum of the Burgers vectors of the climb and the glide dislocations along the crack line account for the stress field around its blunt tip where dislocation enters, and the sharp tip where crack propagates. Firstly, Gauss-Chebyshev quadrature technique is applied to solve the governing equation of dislocation density functions constrained by load-free crack faces. When taking plasticity into account at both crack tips where stresses are high, the generalized Irwin plastic zone correction is recommended. Plastic zone size (PZS) for both tips and crack tip opening displacement (CTOD) for the sharp tip are then obtained. The effects of coating thickness, crack depth, material mismatch and displacement loads ratio onto PZSs and CTOD have been analyzed in detail.

Keywords: Zener-Stroh crack, bi-material coating-substrate composite, singular integral equations, Gauss-Chebyshev quadrature technique, PZS, CTOD.

Introduction

Apart from the well-known Griffith crack, there is another mechanism of cracking as a result of edge dislocations in solids, firstly realized by Zener and Stroh [Stroh (1954); Zener (1948)] in literature. They proposed that the edge dislocations of a pile-up that are stopped at an obstacle, such as a grain boundary (GB), could coalesce into a crack nucleus (Fig. 1). Some situations in which massive Zener-Stroh cracks are coalesced have been recognized: Noticing solids with smaller grain size will possess more GBs, as well as less amount of possible pile up of dislocations at each boundary. More GBs lead to frequent occurrences of dislocation pile-up and more potential sources of crack nucleuses; while less pile up of dislocations accumulated at each location make it harder for dislocations to be repelled and overcome the energetic barrier for diffusion across a GB. That's why GBs are major sinks of dislocations as well. Therefore, knowing more about Zener-Stroh cracks' behaviors in micro- or nano-scale structures is of much significance.
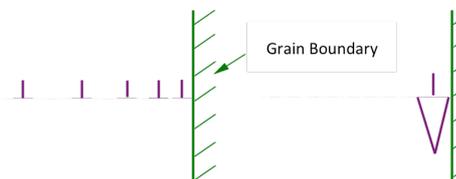
Figure 1. Zener-Stroh crack initiation

Many attractive features of nanocrystalline (nc) and microcrystalline (mc) metals, such as high strength and hardness, and improved resistance to wear and corrosion damage compared to conventional metals have been fully discovered by researchers [Kumar et al. (2003); Zhang et al. (2005)]. However, due to the presence of high-density ensembles of GBs as stoppers for lattice dislocations, nc becomes quite easy

to break, especially those super-refined into ultrasmall grain sizes [Pozdnyakov (2003)]. High stress concentration at GBs will facilitate crack nucleation and growing process, degrading the fracture toughness of the material. Moreover, bulk nanostructured materials usually have disappointingly low ductility. They exhibit a very low uniform elongation due to low work-hardening rate. Localized deformation (necking) under tensile stress often occurs very abruptly because of nc's low dislocation accumulation capability [Zhao et al. (2006)]. In order to enhance both toughness and ductility of nc and mc, without sacrificing their high yield strengths, numerous methodologies and techniques have been suggested [Kuntz et al. (2004); Wang et al. (2002)]. No matter how different they look like, people do believe that fabrication of micro- or nano-composites is the best solution up to now in fulfilling practical needs where both strength and toughness are highlighted.

Although literatures commence to study on the effects of nanocomposites onto fracture toughness as a whole body, the localized behavior, such as how certain types of microcrack are initiated and propagating through the composite is lack of information, especially with plastic zone correction at crack tips. Therefore we manage to start with the investigation of a Zener-Stroh crack lying in a semi-infinite substrate covered by a coating with finite thickness, and check around its crack tips to see how certain properties are improved from single-phase structures. To our best knowledge, most of the time, ductility of nanocomposites, though enhanced, still remains limited compared to their corresponding values of traditional coarse-grained materials. That's why for most cases, the size of the process zone and the plastic region around the crack tip is sufficiently small, so the small-scale yielding assumption is applicable to account for crack tip plasticity [Koch (2007)]. With the additional concern of more complicated configuration and stress field, we proposed a generalized Irwin model in dealing with mode I and mode II stress intensity factors simultaneously. The advantage of this model is that the model itself is intuitive and the procedure can be easily adopted by engineers. Results include the plastic zone size, the crack tip opening displacement, and effective stress intensity factors of mode I and mode II in different scenarios.

The generalized Irwin model of a sub-interface Zener-Stroh crack

The plastic zone size

Current physical problem is depicted in Fig. 2a. Stress fields ahead of the crack tips along x axis can be expressed as [Anderson (2005)]

$$\sigma_{xx}^{(t_m)} = \sigma_{yy}^{(t_m)} = \frac{K_I^{(t_m)}}{\sqrt{2\pi r}}, \sigma_{xy}^{(t_m)} = \frac{K_{II}^{(t_m)}}{\sqrt{2\pi r}}, \sigma_{zz}^{(t_m)} = v_2\left(\sigma_{xx}^{(t_m)} + \sigma_{yy}^{(t_m)}\right) = \begin{cases} 0 & \text{Plane stress} \\ \dfrac{2v_2 K_I^{(t_m)}}{\sqrt{2\pi r}} & \text{Plane strain} \end{cases}, m = 1,2. \quad (1)$$

Here (t_1), (t_2) stands for the blunt and sharp crack tips, respectively. The subscript 2 refers to the substrate material. v_2 represents its Poisson's ratio. Due to the Von Mises yield criterion, yielding will occur if the equivalent stress σ_e reaches the yielding stress of material 2, σ_{ys}

$$\sigma_e^{(t_m)} = \sqrt{\frac{\left(\sigma_{xx}^{(t_m)} - \sigma_{yy}^{(t_m)}\right)^2 + \left(\sigma_{xx}^{(t_m)} - \sigma_{zz}^{(t_m)}\right)^2 + \left(\sigma_{zz}^{(t_m)} - \sigma_{yy}^{(t_m)}\right)^2 + 6(\sigma_{xy}^{(t_m)})^2}{2}} = \sigma_{ys} = \frac{K_e^{(t_m)}}{\sqrt{2\pi r}}. \quad (2)$$

The equivalent stress intensity factors K_e are then obtained by

$$K_e^{(t_m)} = \begin{cases} \sqrt{(K_I^{(t_m)})^2 + 3(K_{II}^{(t_m)})^2} & \text{Plane stress} \\ \sqrt{\left(1 - 2v_2\right)^2 (K_I^{(t_m)})^2 + 3(K_{II}^{(t_m)})^2} & \text{Plane strain} \end{cases} \quad (3)$$

From Eq. (2), the first order estimation of PZS can be expressed with respect to σ_{ys},

$$r_y^{(t_m)} = \frac{(K_e^{(t_m)})^2}{2\pi\sigma_{ys}^2}. \quad (4)$$

Due to stress relaxation around crack tips, it is clear that the actual plastic strain will be extended to a larger zone. See from Fig. 2b, the 2nd order estimation of PZS, known as plastic zone correction, has the following form:

$$r_p^{(t_m)} = \frac{(K_e^{(t_m)})^2}{\pi\sigma_{ys}^2}. \quad (5)$$

Crack tip opening displacement

The crack tip opening displacement of a Zener-Stroh crack under the generalized Irwin model is shown in Fig. 2c. Although we can see faces are completely open throughout the crack, propagation will be initiated only at the sharp tip due to the existence of tensile stress, not at the blunt tip that has been compressed and stabilized. As a result, only CTOD at the sharp tip will be discussed hereafter. CTOD at the sharp tip δ is given in literature as [Anderson (2005)]

$$\delta = \frac{\kappa_2 + 1}{\mu_2} K_I \sqrt{\frac{r_y^{t_2}}{2\pi}}, \quad (6)$$

μ_2 is the shear modulus. $\kappa_2 = \frac{3 - v_2}{1 + v_2}$ for plane stress, and $\kappa_2 = 3 - 4v_2$ for plane strain. Substitute Eq. (4) into (6), with the universal relation $E_2/2\mu_2 = 1 + v_2$, we have

$$\delta = \frac{4}{\pi E_2'} \cdot \frac{K_I^{t_2} K_e^{t_2}}{\sigma_{ys}}, \quad (7)$$

in which $E_2' = E_2$ for plane stress, $E_2' = \frac{E_2}{1 - v_2^2}$ for plane strain. E_2 is the elastic modulus of the substrate.

(a)

(b)

(c)
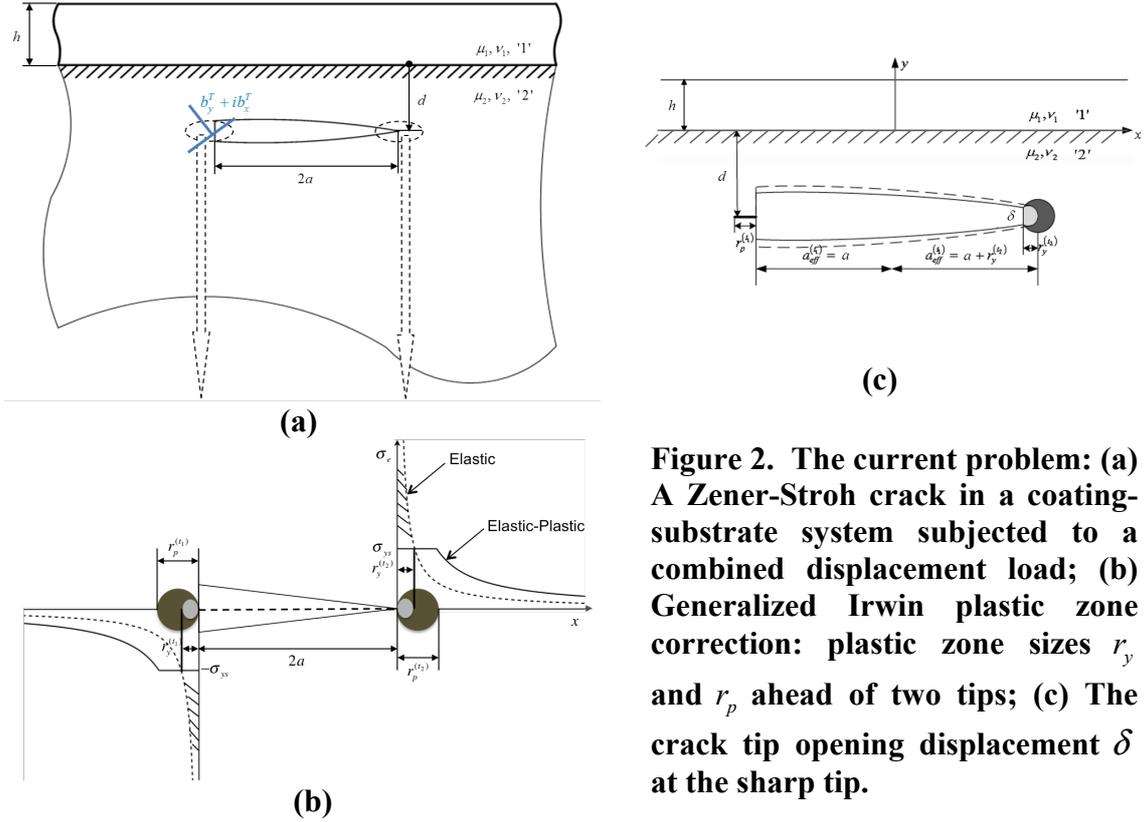
Figure 2. The current problem: (a) A Zener-Stroh crack in a coating-substrate system subjected to a combined displacement load; (b) Generalized Irwin plastic zone correction: plastic zone sizes r_y and r_p ahead of two tips; (c) The crack tip opening displacement δ at the sharp tip.

The effective stress intensity factor

Let's begin with the Zener-Stroh crack of length $2a$ without plastic zone correction. Concentrated climb and glide edge dislocations at the blunt tip would lead to an array of dislocations emitted along the crack line. Due to traction free condition on the crack faces, governing equation of combined distributed dislocation density $B(\xi)$ turns out to be

$$\sigma_{yy}(x) + i\sigma_{xy}(x) = 2\int_{-a}^{a}\frac{\overline{B}(\xi)}{x-\xi}d\xi + \int_{-a}^{a}B(\xi)F_1(x-\xi)d\xi + \int_{-a}^{a}\overline{B}(\xi)F_2(x-\xi)d\xi = 0, \qquad |x| < a \qquad (8)$$

where $\overline{(\ )}$ denotes the complex conjugate. The kernels $F_1(x-\xi)$ and $F_2(x-\xi)$ are given in the literature [Lu and Lardner (1992)]. The boundary conditions are

$$\int_{-a}^{a}B_x(\xi)d\xi = \frac{\mu_2}{\pi(1+\kappa_2)}b_x^T, \quad \int_{-a}^{a}B_y(\xi)d\xi = \frac{\mu_2}{\pi(1+\kappa_2)}b_y^T, \qquad (9)$$

in which $B_x(\xi)$ and $B_y(\xi)$ are the glide and climb dislocation densities, respectively. b_x^T and b_y^T are the corresponding total sum of Burgers vector in the x and y directions.

Since the dislocation density tends to go infinity in a square root singular manner, B_x and B_y can be rewritten into $B_x(s) = \dfrac{1}{\sqrt{1-s^2}} \phi_x(s)$ and $B_y(s) = \dfrac{1}{\sqrt{1-s^2}} \phi_y(s)$, where $\phi_x(s)$ and $\phi_y(s)$ are unknown regular functions. Substituting B_x and B_y into Eqs. (8) and (9), four singular integral equations with Cauchy kernels are obtained. Gauss-Chebyshev quadrature technique is then implemented to solve them numerically, thus B_x and B_y can be obtained [Zhuang et al. (2013)]. Mode I and mode II stress intensity factors at each crack tip can be derived in the following form [Weertman (1996)]

$$K_I^{(t_1)} = -2\pi\sqrt{\pi a}\,\phi_y(-1), \quad K_I^{(t_2)} = 2\pi\sqrt{\pi a}\,\phi_y(+1),$$
$$K_{II}^{(t_1)} = -2\pi\sqrt{\pi a}\,\phi_x(-1), \quad K_{II}^{(t_2)} = 2\pi\sqrt{\pi a}\,\phi_x(+1). \tag{10}$$

Here $\phi_x(\pm 1)$ and $\phi_y(\pm 1)$ are values of regular functions at blunt (-1) and sharp (+1) crack tips after the half-crack length a has been normalized to 1.

When we improve our analysis to investigate the elastic-plastic fracture behaviors of the Zener-Stroh crack, plastic zone correction needs to be imposed at both crack tips. The elongated, effective half-crack length is given

$$a_{eff}^{(t_m)} = a + r_y^{(t_m)} = a + \frac{(K_e^{(t_m)})^2}{2\pi\sigma_{ys}^2}. \tag{11}$$

Remember that a Zener-Stroh crack can only propagate from the sharp tip, let's focus on investigation of effective stress intensity factors at that tip. Hence, we get

$$K_I^{eff} = 2\pi\sqrt{\pi a_{eff}^{(t_2)}}\,\phi_y{}'(+1), \quad K_{II}^{eff} = 2\pi\sqrt{\pi a_{eff}^{(t_2)}}\,\phi_x{}'(+1), \tag{12}$$

where $\phi_x{}'(+1)$ and $\phi_y{}'(+1)$ are values of regular functions at the sharp tip after the effective half-crack length $a_{eff}^{(t_2)}$ has been normalized to 1.

Numerical examples and discussion

Some numerical examples and discussions for the plastic zone size, the crack tip opening displacement and effective stress intensity factors of a Zener-Stroh crack of length $2a$ are given. The crack is embedded in a coating-substrate without external loading. The total sum of the Burgers vector throughout the crack $b_y^T + ib_x^T$ ensures faces are fully open. For the ease of assessment, PZS and CTOD are normalized by:

$$K_I^0 = \frac{2\mu_2 b_y^T}{(1+\kappa_2)\sqrt{\pi a}}, \quad K_{II}^0 = \frac{2\mu_2 b_x^T}{(1+\kappa_2)\sqrt{\pi a}}, \quad K_e^0 = \begin{cases} \sqrt{(K_I^0)^2 + 3(K_{II}^0)^2} & \text{Plane stress} \\ \sqrt{(1-2\nu_2)^2(K_I^0)^2 + 3(K_{II}^0)^2} & \text{Plane strain} \end{cases},$$

$$r_0 = \frac{(K_e^0)^2}{\pi\sigma_{ys}^2}, \qquad \delta_0 = \frac{4K_I^0 K_e^0}{\pi E_2{}'\sigma_{ys}}, \tag{13}$$

where K_I^0, K_{II}^0, K_e^0, r_0 and δ_0 are the mode I, mode II, equivalent stress intensity factors, PZS and CTOD respectively for the same Zener crack that is embedded in a homogeneous infinite plate of material '2'. The dependence of the normalized plastic zone size $r_p^{(t_m)}/r_0$, normalized crack tip opening displacement δ/δ_0, and normalized effective stress intensity factors $K_I^{eff}/K_I^{(t_2)}$ and $K_{II}^{eff}/K_{II}^{(t_2)}$ on the normalized coating thickness h/a, normalized crack depth d/a, the Dundurs' parameter α, as well as displacement loads ratio b_x^T/b_y^T are shown in Tables 1-2.

Normalized PZS and normalized CTOD

In Table 1, normalized PZS at the sharp (r_{sp}/r_0) and blunt tip (r_{bt}/r_0), and normalized CTOD (δ/δ_0) at the sharp tip of the Zener crack are depicted in different scenarios. For the case of $b_x^T = 0$, we may find the same α leads to a pair of identical plastic zones around two tips. In the most special situation $\alpha = 0$, if coating thickness h is very large compared to half-crack length a ($h = 10a$), it is verified from second column that values of r_p and δ converge to their corresponding values r_0 and δ_0 (they are called "reference values" in the context), no matter how far the crack is located beneath the interface.

Comparing figures in second and third columns, we observe that with a decreasing coating thickness, PZS at both tips, and CTOD at the sharp tip will be decreased. And the trend becomes more significant in plane stress than plane strain condition. This observation tells us a fact that increasing the volume fraction of added material (the coating) will enhance the ductility of the original structure (the substrate) in manner of magnifying the plastic deformation region around the crack tips.

Effects of the crack depth d can be viewed from third column ($h = 2a$), where different material mismatches have different reactions from a decreased crack depth. Softer coatings ($\alpha < 0$) shrink PZS and CTOD values lower than the reference while stiffer coatings ($\alpha > 0$) result in higher-than-reference plasticity quantities. This indicates a fact that when the crack gets nearer to the interface, it becomes easier to propagate with a softer coating covered on top, but stabilized under the protection of a stiffer coating.

Last two columns show continuous influence of displacement loads ratio b_x^T/b_y^T onto PZS and CTOD. Supposing that crack depth d can be either $0.5a$ or $5a$. If $x-dir$ displacement load gradually increases from $0.1 \times b_y^T$ to b_y^T, normalized r_p and δ will be increased without exceptions. As long as the crack is far away from the interface (for example $d/a = 5$), crack tip parameters become converge even within different material mismatches. But when the crack locates nearer, a thorough examination at different material mismatches tells that: a larger α always results in higher sensitivities of normalized r_p and δ along with the changing b_x^T/b_y^T. Besides, this effect onto plane strain cases is more significant than it does on plane stress cases.

Table 1. Normalized PZS and normalized CTOD, with $\beta = 0$
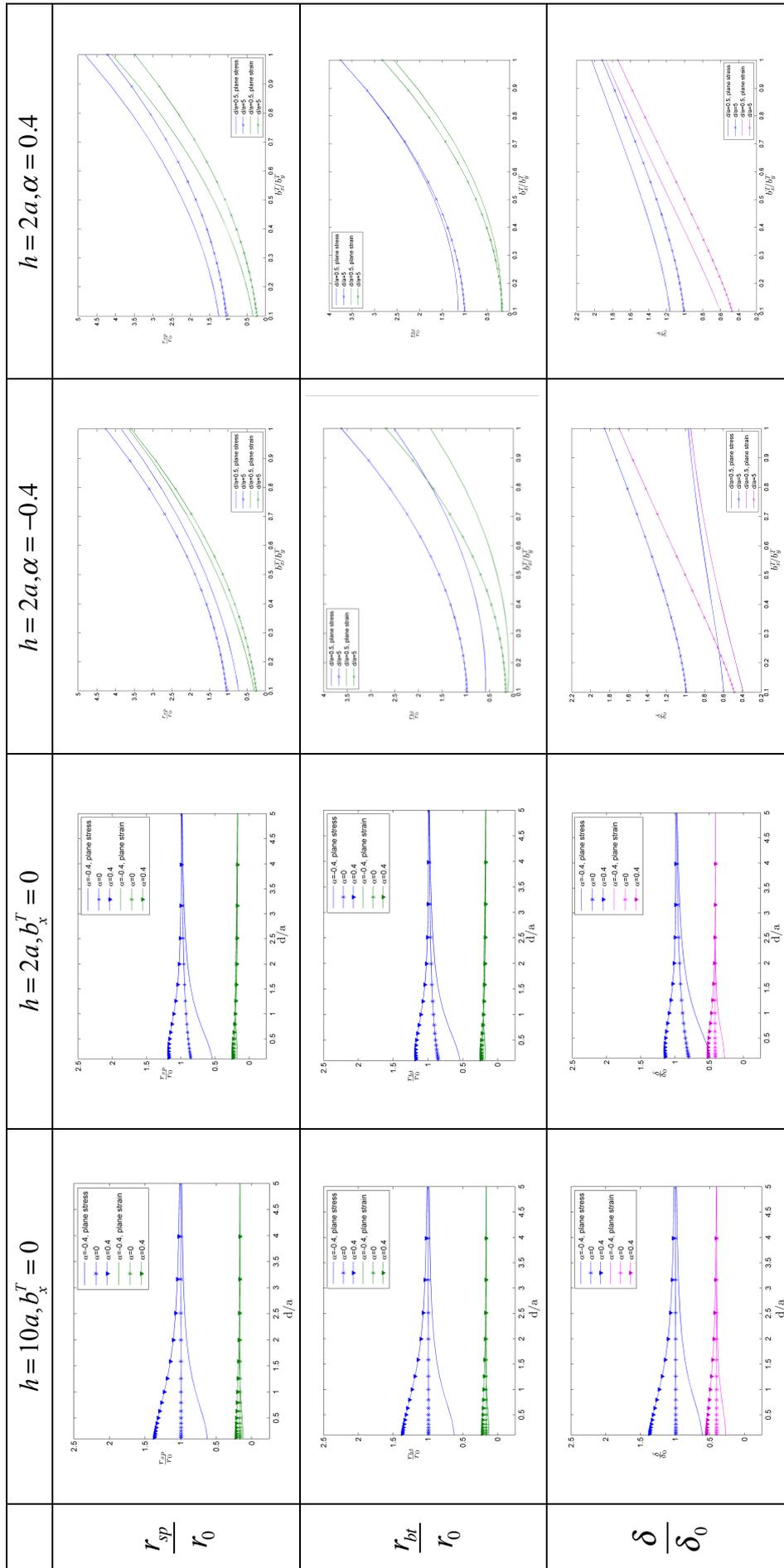
The effective stress intensity factors K_I^{eff} and K_{II}^{eff}

Due to the inclusion of the 1st order plastic zone size, we can see from Table 2, generally speaking, K_I^{eff}/K_I and K_{II}^{eff}/K_{II} are lower than 1. However the scale of decrement depends on material mismatches, crack depths and many more. One can see that K_I^{eff}/K_I and K_{II}^{eff}/K_{II} deduce more when the coating is stiffer ($\alpha > 0$), meaning that the correction of K_I and K_{II} is more necessary if the substrate is coated with stiffer material.

Moreover, when $\alpha < 0$, crack with greater depth shows smaller values of K_I^{eff}/K_I and K_{II}^{eff}/K_{II}. Similar trends can be seen when $\alpha > 0$ and crack gets closer to the interface. This illustrates that cases of shallower crack under stiffer coating, or deeper crack under softer coating, are in greater need of correction in terms of stress intensity factors. It is also not difficult to find K_I^{eff} and K_{II}^{eff} of the sharp tip of a Zener crack under plane stress differ more from K_I and K_{II} respectively than the results shown for plane strain condition.

Table 2. Effective stress intensity factors K_I^{eff} and K_{II}^{eff}, with $\beta = 0$, $h = 2a$, $b_x^T / b_y^T = 0.5$

d/a			0.1	0.2	0.3	0.4	0.5	0.6
$\alpha = -0.4$	$\dfrac{K_I^{eff}}{K_I}$	Plane stress	0.99827	0.99822	0.99818	0.99815	0.99812	0.99808
		Plane strain	0.99857	0.99853	0.99851	0.99850	0.99848	0.99847
	$\dfrac{K_{II}^{eff}}{K_{II}}$	Plane stress	0.99783	0.99777	0.99772	0.99767	0.99763	0.99758
		Plane strain	0.99820	0.99816	0.99813	0.99811	0.99809	0.99807
$\alpha = 0.4$	$\dfrac{K_I^{eff}}{K_I}$	Plane stress	0.99682	0.99694	0.99704	0.99713	0.99720	0.99727
		Plane strain	0.99791	0.99802	0.99810	0.99818	0.99824	0.99830
	$\dfrac{K_{II}^{eff}}{K_{II}}$	Plane stress	0.99625	0.99631	0.99637	0.99642	0.99648	0.99653
		Plane strain	0.99753	0.99761	0.99767	0.99773	0.99779	0.99784

Conclusions

In the present work, plastic zone size, crack tip opening displacement and effective stress intensity factors for a sub-interface Zener-Stroh crack in a coating-substrate system under combined displacement load $b^T = b_y^T + i b_x^T$ are investigated by a generalized Irwin model. In the numerical examples, we specifically describe the dependence of normalized plastic zone size for sharp tip r_{sp}/r_0, for blunt tip r_{bt}/r_0, normalized crack tip opening displacement for sharp tip δ/δ_0, as well as normalized effective stress intensity factors for sharp tip K_I^{eff}/K_I, K_{II}^{eff}/K_{II} on normalized crack depth d/a, normalized coating thickness h/a, Dundurs' parameter α, and displacement loads ratio b_x^T/b_y^T. According to the results obtained and discussed, following conclusions can be made:

1. Either the Zener-Stroh crack exists in an infinite bi-material composite without mismatches ($\alpha = \beta = 0$), or it locates far from the interface ($d/a > 5$)

in a coating-substrate system with mismatches (arbitrary α and β), the current physical problem reduces to the corresponding problem of the same crack in a homogeneous material.

2. Normalized PZS and normalized CTOD will be increased with the increasing coating thickness. When coating thickness is fixed, a Zener crack moves nearer to the interface will experience higher PZS and CTOD values if the substrate is coated with stiffer material, but lower PZS and CTOD if it has a softer coating instead.

3. These are the circumstances shall we need to produce the effective stress intensity factors: 1) if the coating is softer than the substrate and the crack is relatively deep beneath the interface; 2) if the coating is stiffer than the substrate and the crack locates near the interface; 3) choose substrate with stiffer-coating system to evaluate when the other conditions are the same; 4) choose plane stress structure to evaluate when the other conditions are the same.

4. When the coating thickness and crack depth are fixed, with the increasing displacement loads ratio b_x^T / b_y^T, normalized PZS and normalized CTOD grows more rapidly and around higher values if the substrate is coated with stiffer materials, especially in plane strain condition.

Reference

Anderson, T. L. (2005). *Fracture Mechanics: Fundamentals and Applications*: Taylor & Francis.

Koch, C. C. (2007). Structural nanocrystalline materials: an overview. *Journal of materials science, 42*(5), 1403-1414.

Kumar, K., Van Swygenhoven, H., & Suresh, S. (2003). Mechanical behavior of nanocrystalline metals and alloys. *Acta Materialia, 51*(19), 5743-5774.

Kuntz, J. D., Zhan, G.-D., & Mukherjee, A. K. (2004). Nanocrystalline-matrix ceramic composites for improved fracture toughness. *Mrs Bulletin, 29*(01), 22-27.

Lu, H., & Lardner, T. J. (1992). Mechanics of Subinterface Cracks in Layered Material. *International Journal of Solids and Structures, 29*(6), 669-688.

Pozdnyakov, V. (2003). Nanostructural crack arrest. *Technical Physics Letters, 29*(2), 151-153.

Stroh, A. N. (1954). The formation of cracks as a result of plastic flow. *Proc. Roy. Soc. London, 223A*, 404-414.

Wang, Y., Chen, M., Zhou, F., & Ma, E. (2002). High tensile ductility in a nanostructured metal. *Nature, 419*(6910), 912-915.

Weertman, J. (1996). *Dislocation Based Fracture Mechanics*: World Scientific.

Zener, C. (1948). *The micro-mechanism of fracture*. Paper presented at the Fracturing of Metals, Cleveland.

Zhang, S., Sun, D., Fu, Y., & Du, H. (2005). Toughening of hard nanostructural thin films: a critical review. *Surface and Coatings Technology, 198*(1), 2-8.

Zhao, Y.-H., Liao, X.-Z., Cheng, S., Ma, E., & Zhu, Y. T. (2006). Simultaneously increasing the ductility and strength of nanostructured alloys. *Advanced Materials, 18*(17), 2280-2283.

Zhuang, J., Yi, D. K., & Xiao, Z. M. (2013). Elastic–plastic analysis of a sub-interface crack in a coating-substrate composite. *International Journal of Solids and Structures, 50*(2), 414-422. doi: http://dx.doi.org/10.1016/j.ijsolstr.2012.10.018

Determining the Residual Stresses of Circular Weld Bead with Eigenstrain BIE as an Inverse Approach

* H. Ma[1], Y.J. Tang[2], and C. Yan[3]

[1]College of Sciences, Shanghai University, Shanghai 200444, China.
[2]Shanghai Institute of Applied Mathematics and Mechanics, Shanghai 200072.
[3]School of Chemistry, Physics and Mechanical Engineering, Queensland University of Technology, QLD 4001, Australia.

*Corresponding author: hangma@staff.shu.edu.cn

Abstract

In the present work, based on the computational model of the eigenstrain boundary integral equations (BIE) as an inverse problem, the algorithm is investigated to determine the circular weld bead induced residual stresses, where the eigenstrain is considered to be the origin of residual stresses in structures. In order to reduce the number of unknowns and to consider the stability of inverse problem, the eigenstrains are approximated in terms of low-order polynomials in the local area, which is divided by cells, around welded zones according to the features of welding. The corresponding domain integrals with polynomial eigenstrains in each cell are transformed into the boundary integrals to preserve the favorable features of the boundary-only discretization in the numerical solutions. The sensitivity matrix in the inverse approach for evaluating the eigenstrain fields represented by the coefficients of polynomials is constructed with the aid of measured stresses in the domain after welding over a few selected measuring points. In the numerical examples, the residual stresses of circular weld beads in both the finite and infinite plates are evaluated with the proposed procedure, verifying the feasibility and effectiveness of the present algorithm.

Keywords: residual stress, circular weld bead, eigenstrain, boundary integral equation, inverse approach

Introduction

The circular weld beads are used quite frequently in the welded structures as well as in the repair weld and in the test pieces for evaluating the effects of stress corrosion cracks, so that the residual stress fields such formed become one of the primary concerns of engineers and researchers in this field. As the residual stresses have a significant influence on the performance of related components in service [Masubuchi (1980)], when such a component is in service, the associated residual stresses may superimpose on the applied stress to influence the deformation behavior of components, which induce distortion during further machining and cause unexpected failure or reduce the service time of components. Since the nature of residual stress is in self-equilibrium, however, its determination is not an easy task, especially with the mechanical techniques [Prime (1999)].

There are a great number of techniques to detect the residual stresses in a solid which can be classified as three major groups: physical, mechanical and numerical techniques. In the physical techniques, for example, the variation of inter-crystal distances can be detected by X-ray diffraction [Korsunsky et al (2006)] or by sound speed changes in acoustoelasticity or by magnetic techniques, most of them depending on certain material properties. In contrast, in the mechanical techniques, since the direct detection is impossible, parts of the material have to be removed from the solid to disturb the stress balance while the response of the specimen is measured in terms of either strains or shape changes at some other locations on the surface of the body. The blind-hole drilling may be the most commonly used residual stress measuring method in practice and consists essentially of drilling a small blind-hole on the surface of solid and measuring the strain field induced by material removal, usually by means of electrical resistance strain gauges. In addition to the use of strain

gauges, the strain fields can also be measured by photo means such as the electronic speckle pattern interferometer [Suterio et al (2006)] or the digital image correlation in recent advances. In addition to the cost of mechanical methods, all of them are more or less destructive to the measured component by material removals due to the nature of self-balancing. It is evident that the mechanical techniques always provide a limited level of detail, due to the finite number of discrete data points that restricts the possibility of reconstructing full-field stress distributions.

Numerical methods present a supplementary but effective means for determining the residual stresses. However, the detailed modeling of the process of residual stress generation requires the knowledge of numerical models for analyzing sophisticated coupled microstructural and thermo-mechanical behaviors, which rely deeply on the understanding of constitutive laws and material parameters. As is widely accepted, residual stresses in components at service are caused by incompatible internal permanent strains, named originally as the inherent strains [Ueda et al (1986); Ma et al (1998a)] and lately as the eigenstrains [Jun et al (2010)], induced by any inhomogeneous inelastic deformation, temperature gradients or phase transformations during manufacturing and processing of the components. By making use of the information observed from the experiment at a number of selected points, the unknown eigenstrain distributions can be retrieved using the finite element method (FEM) [Lee et al] or the boundary element method (BEM) [Cao et al (2002)], following the mathematical framework of the inverse problem of eigenstrain theory to obtain the whole field of residual stresses.

In spite of the inelastic origin of eigenstrains, the inherent state of residual stress fields falls really into elastic regime so that the BEM would be the most efficient numerical means to deal with the residual stress problems [Qian et al (2004; 2005)]. Based on the concept of eigenstrain, a straightforward computational model as an inverse approach was proposed with the eigenstrain formulations of boundary integral equations to determine the welding residual stresses [Ma et al (2012)]. In the present work, the eigenstrains are approximately expressed in terms of low-order polynomials in the local area around the heat affected zones of circular weld beads, which is divided by cells, according to the features of welding. The corresponding domain integrals with polynomial eigenstrains are transformed into the boundary integrals [Ma et al (1998b)] so that the attractive features of the boundary-only discretization are reserved in the process of numerical solutions using the boundary point method (BPM) [Ma et al (2010)]. The sensitivity matrix in the inverse approach for evaluating the eigenstrain fields represented by the coefficients of polynomials is constructed with the aid of measured stresses in the domain after welding over a few selected measuring points. In the numerical examples, the residual stresses of circular weld beads in both the finite and infinite plates are evaluated with the proposed procedure, verifying the feasibility and effectiveness of the present algorithm.

Computational Model

Eigenstrain Boundary Integral Equations

The displacements and the stresses of a weld plate, Ω, in the static state without body force can be described by the eigenstrain boundary integral equations as follows [Ma et al (2012)]:

$$Cu_i(p) + \int_\Gamma u_j(q)\tau_{ij}^*(p,q)d\Gamma(q) = \int_\Gamma \tau_j(q)u_{ij}^*(p,q)d\Gamma(q) + \int_{\Omega_I} \varepsilon_{jk}^0(q)\sigma_{ijk}^*(p,q)d\Omega(q) \quad (1)$$

$$C\sigma_{ij}(p) = \int_\Gamma \tau_k(q)u_{ijk}^*(p,q)d\Gamma(q) - \int_\Gamma u_k(q)\tau_{ijk}^*(p,q)d\Gamma(q)$$

$$+ \int_{\Omega_I - \Omega_\varepsilon} \varepsilon_{kl}^0(q)\sigma_{ijkl}^*(p,q)d\Omega(q) + \varepsilon_{kl}^0(p)O_{ijkl}^* \quad (2)$$

where Ω_I ($\Omega_I \in \Omega$) represents the local area having unknown eigenstrains, ε_{ij}^0, around the weld bead in Ω, since it is generally true that the position of the local area is known a priori in welding. p and q are the source and field points, u_{ij}^*, τ_{ij}^*, and σ_{ij}^* represent the fundamental solutions for displacement, traction and stress, respectively. u_{ijk}^*, τ_{ijk}^* and σ_{ijk}^* are the related derivatives. C is the conventional boundary shape coefficient, $C=1/2$ if p is on the smooth boundary Γ. Ω_ε is a small region of radius ε around point p when $p \in \Omega_I$, and O_{ijkl}^* is the corresponding free term resulted from the domain integral in (2) which can be derived using the conventional limiting techniques with a small ε region since the kernel of this domain integral is strong singular. A square plate $2w \times 2w$ with a circular weld bead is shown in Fig. 1a.

(a)   (b)   (c)
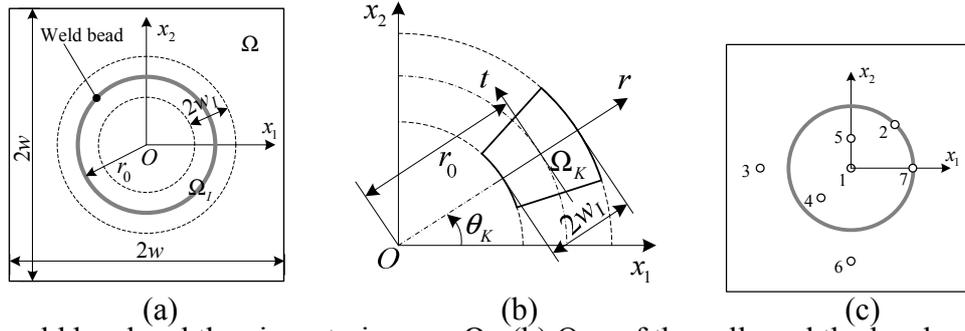
(a) The weld bead and the eigenstrain zone Ω_I; (b) One of the cells and the local coordinate for the eigenstrain zone; (c) The position of measuring points

Figure 1. The plate with a circular weld bead

Eigenstrain Representation

It is obvious from (1) and (2) that once the distributions of the eigenstrains ε^0_{ij} in the domain integrals are known, the unknown boundary displacements can be solved using (1) and the total fields of stresses can be computed using (2). Considering the features of thermal cycles of welding, the distributions of eigenstrains can be approximately expressed in terms of low-order polynomials in the local area Ω_I:

$$\varepsilon^0_{ij} = \sum_{m=0,n=0}^{m+n=M} \alpha^{mn}_{ij} x^m_1(q) x^n_2(q) \qquad (3)$$

where M is the number of terms of polynomials and $\alpha_{ij}{}^{mn}$ the coefficients to be identified. m and n are integers. In fusion welding, the eigenstrain distributions can be expected to be smooth since the temperature field in welding can be expressed by smooth functions especially during the cooling stage. Owing to the similar reason, the eiegenstrain can be assumed to be zero at the boundary of the eiegenstrain domain. The polynomial representations (3) for eigenstrain are inherently smooth, giving a smooth constraint on the eigenstrain field. The domain integrals with polynomial eigenstrains in (1) and (2) can be transformed into the boundary integrals by introducing the two-point variables

$$x_i = x_i(q) - x_i(p) \qquad (4)$$

With this definition, the domain integrals with certain term of polynomials in (1) and (2), respectively, can be expressed in the form of the two-point polynomials as follows:

$$\alpha^{mn}_{kl} \left\{ \int_{\Omega_I} x^m_1(q) x^n_2(q) \sigma^*_{ijk} d\Omega \right\}$$

$$= \alpha^{mn}_{kl} \left\{ \sum_{s=0}^m \sum_{t=0}^n \frac{m!n!}{(m-s)!s!(n-t)!t!} \left[x_1(p) \right]^{m-s} \left[x_2(p) \right]^{n-t} \int_{\Omega_I} x^s_1 x^t_2 \sigma^*_{ijk} d\Omega \right\} \qquad (5)$$

$$\alpha^{mn}_{kl} \left\{ \int_{\Omega_I - \Omega_\varepsilon} x^m_1(q) x^n_2(q) \sigma^*_{ijkl} d\Omega + x^m_1(p) x^n_2(p) O^*_{ijkl} \right\}$$

$$= \alpha^{mn}_{kl} \left\{ \sum_{s=0}^m \sum_{t=0}^n \frac{m!n!}{(m-s)!s!(n-t)!t!} \left[x_1(p) \right]^{m-s} \left[x_2(p) \right]^{n-t} \times \int_{\Omega_I - \Omega_\varepsilon} x^s_1 x^t_2 \sigma^*_{ijkl} d\Omega + x^m_1(p) x^n_2(p) O^*_{ijkl} \right\} \quad (6)$$

where m, n, s and t are all integers. x_i, $x_i(p)$ and $x_i(q)$ are defined in (4). Then the domain integrals at the right hand sides in (5) and (6) with eigenstrains in the form of two-point polynomials can be transformed into the boundary integrals [Ma et al (1998b)], respectively. In this way, the favorable features of the boundary-only discretization are reserved. However, considering the difficulty of representing eigenstrains with low-order polynomials in a ring area formed by the circular weld bead as shown in Fig. 1a and for the purpose of reducing the number of unknown coefficients, the eigenstrain zone is divided into cells in the present work, one of them being shown in Fig. 1b. The low-order polynomials in each cell Ω_K are represented using the local polar coordinates so that all of the polynomials in the cells are the same. That is, the polynomials in each cell have the same number of terms with the same coefficients owing to the circular weld bead. The domain integrals in (1) and (2) become

$$\int_{\Omega_I} \varepsilon_{jk}^0 \sigma_{ijk}^* d\Omega = \sum_{K=1}^{N_I} \int_{\Omega_K} \varepsilon_{jk}^0 \sigma_{ijk}^* d\Omega \tag{7}$$

$$\int_{\Omega_I - \Omega_\varepsilon} \varepsilon_{kl}^0 \sigma_{ijkl}^* d\Omega = \sum_{K=1}^{N_I} \int_{\Omega_K} \varepsilon_{kl}^0 \sigma_{ijkl}^* d\Omega \tag{8}$$

respectively, where N_I is the number of cells divided and the eigenstrains and the kernels in each Ω_K are also computed in the local polar coordinates.

Inverse Approach

In the inverse approach, the information from experiments is required to identify the unknown coefficients α_{ij}^{mn} in (3), the stresses measured after welding in the domain at a number of selected measuring points, $x^{(k)}$

$$\sigma_{ij}\left(x^{(k)}\right) = \sigma_{ij}^{0(k)}, \qquad x^{(k)} \in \Omega, \, k=1,2,\ldots, M_S \tag{9}$$

where $\sigma_{ij}^{0(k)}$ is the measured stresses and M_S the number of measuring points of stresses. Since the residual stresses of weld plates have three components, σ_{11}^0, σ_{12}^0 and σ_{22}^0, at one point for the two-dimensional problem, the number of known information from experiments is $3M_S$. By employing the BPM [Ma et al (2010)] and noticed the traction-free boundary conditions in the residual stress problem, the displacement equation (1) combined with (7) can be written after discretization in matrix from as

$$\mathbf{Hu} = \mathbf{B\alpha}, \quad \mathbf{B} = \sum_{K=1}^{N_I} \mathbf{T}^K \mathbf{B}^K \tag{10}$$

where $\mathbf{u}$ is the vector of displacements at all the N nodal points on the boundary Γ, $\mathbf{\alpha}$ the vector of unknown coefficients, and $\mathbf{H}$ and $\mathbf{B}$ the corresponding coefficient matrices. Similarly, the discrete stress equation (2) combined with (8) can be used to compute the stresses at selected points as follows

$$\mathbf{\sigma} = \mathbf{Fu} + \mathbf{D\alpha} = \left(\mathbf{FH}^{-1}\mathbf{B} + \mathbf{D}\right)\mathbf{\alpha} = \mathbf{S\alpha}, \quad \mathbf{D} = \sum_{K=1}^{N_I} \mathbf{T}_\sigma^K \mathbf{D}^K \tag{11}$$

where $\mathbf{S}$ is the so-called sensitivity matrix, $\mathbf{F}$ and $\mathbf{D}$ the corresponding coefficient matrices. The transformation matrices in (10) and (11) are defined respectively as

$$\mathbf{T}^K = \begin{bmatrix} \cos\theta_K & -\sin\theta_K \\ \sin\theta_K & \cos\theta_K \end{bmatrix}, \quad \mathbf{T}_\sigma^K = \begin{bmatrix} \cos^2\theta_K & -2\cos\theta_K \sin\theta_K & \sin^2\theta_K \\ \cos\theta_K \sin\theta_K & \cos^2\theta_K - \sin^2\theta_K & -\cos\theta_K \sin\theta_K \\ \sin^2\theta_K & 2\cos\theta_K \sin\theta_K & \cos^2\theta_K \end{bmatrix} \tag{12}$$

$\mathbf{B}^K$ and $\mathbf{D}^K$ are formed from the kernels of domain integrals in (7) and (8), respectively, which are computed by line integrals after the transformations using (5) and (6). The unknown coefficients of eigenstrains $\mathbf{\alpha}$ can be obtained using the least square method by minimizing the object function, $\mathbf{\Phi}$, defined as follows

$$\mathbf{\Phi} = \frac{1}{2}\left\| \mathbf{S\alpha} - \mathbf{\sigma}^0 \right\|^2 \tag{13}$$

where $\mathbf{\sigma}^0$ represents the vector of measured stresses. The unknown eigenstrain coefficients can be computed by the minimizing condition of (13) as $\mathbf{S}^{\mathrm{T}}(\mathbf{S\alpha} - \mathbf{\sigma}^0) = 0$ so that to obtain

$$\mathbf{\alpha} = \left(\mathbf{S}^{\mathrm{T}}\mathbf{S}\right)^{-1} \mathbf{S}^{\mathrm{T}}\mathbf{\sigma}^0 \tag{14}$$

Numerical Examples

Conditions of Computation

Both the finite and infinite plates with circular weld bead are considered in the numerical examples, corresponding to the cases of test pieces and repair welds, respectively. The finite plate is shown in Fig. 1a with the width of localized area Ω_I being set as $w_I=0.3w$ expressed in dashed lines where the eigenstrains are distributed. This width is somewhat wider than that of the heat affected zone (HAZ) according to the parameters of the material and welding, since the HAZ refers as to the narrow band with changes in microstructures of the material near fusion line while the eigenstrain domain corresponds to the zone undergoing plastic tensions in cooling stage following compressive

deformations in heating stage in the welding thermal cycle. The boundary Γ is discretized by N=100 nodes and the eigenstrain zone Ω_I is divided equally by N_I=40 cells, one of them as shown in Fig. 1b, in the application of the BPM.

Only the normal components of eigenstrains in each cell are considered in the analysis and are approximately expressed in terms of low-order polynomials in the polar coordinate as follows

$$\varepsilon_{rr}^0\left(r\right)/\varepsilon_S = -4 + 0.78r^2 \tag{15}$$

$$\varepsilon_{tt}^0\left(r\right)/\varepsilon_S = -1 + 0.44r^2 \tag{16}$$

where ε_S stands for the material constant, or the yield strain, defined as the strain when the Von Mises stress reaches the yield strength, σ_S, of the material. The eigenstrains given in (15) and (16) satisfy approximately both the zero condition at the border and the maximum value at the center of Ω_I following the features of welding, which are used to compute the control values of stresses such as the measured stresses. The positions for the stress measuring points are shown in Fig. 1c, where the idealized measuring stresses are computed using the BPM with the values of eigenstrains in (15) and (16). With these idealized data, the residual stress can be reconstructed after solving the eigenstrains using the inverse approach stated above. However, as there are always errors in the experimental measurements, 10% random noises are introduced into the idealized data as follows

$$\boldsymbol{\sigma}_{noise}^0 = \left(1 \pm 0.1 ran\right)\boldsymbol{\sigma}^0 \tag{17}$$

where *ran* represents the random function varying between 0 and 1. With these noisy data, the residual stress can also be reconstructed after solving the eigenstrains using the inverse approach. For the infinite plates, the solution procedure and all parameters are as the same with those of the finite plates except that there is no outer boundary Γ so that the boundary integrals in (1) and (2) vanish.

Computed Results

In all of the following figures, the stress distributions are shown along the x_1 axis. The computed stresses with the inverse approaches are computed using noisy data of three measuring points. The computed results of the infinite and the finite plates are presented in Figs. 2 and 3, respectively, showing the feasibility and effectiveness of the present algorithm.

(a) r_0/w=0.3 (b) r_0/w=0.5 (c) r_0/w=0.8
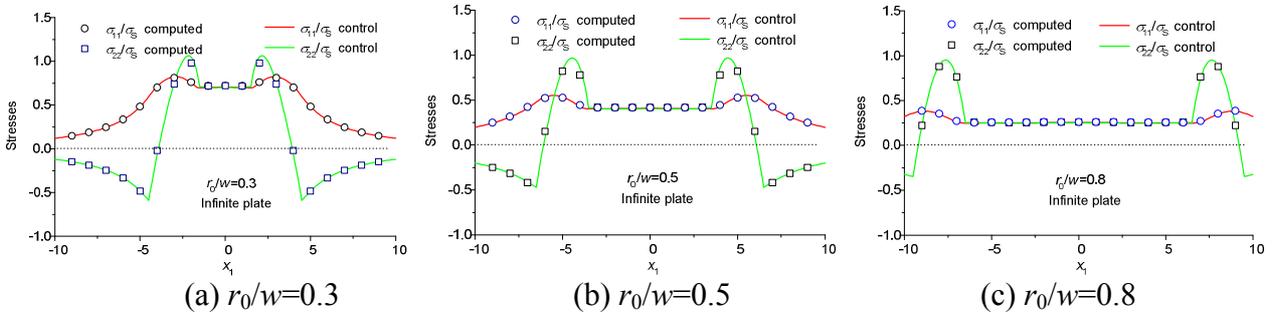
Figure 2. **The residual stress distributions in infinite plates**

(a) r_0/w=0.3 (b) r_0/w=0.5 (c) r_0/w=0.8
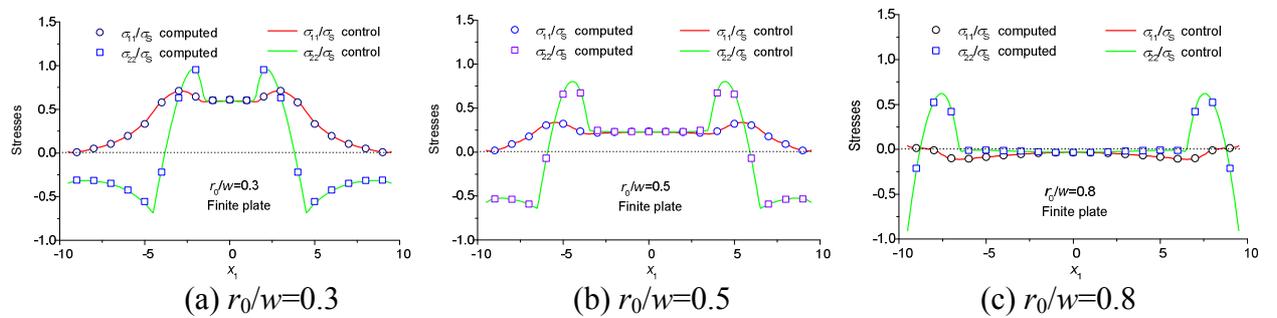
Figure 3. **The residual stress distributions in finite plates**

It can be seen from Figs. 2 and 3 that there are equally biaxial stress fields (σ_{11}=σ_{22}) inside the circular weld beads. The values of these biaxial stresses decrease with the increase of the radius r_0/w, formed by the constrained shrinkages of the welding plastic zone during cooling stages in both

the transverse and the longitudinal directions, or radius and circular directions respectively in the case of circular weld beads. In general, the transverse shrinkage plays the principal role since the gratitude of it is greater than that of the longitudinal shrinkage so that there are generally the equally tensile biaxial stresses inside the circular weld beads. However, the opposite situation can occur as shown in Fig. 3c that the equally compress biaxial stresses exist when r_0/w is relatively large in the finite plate, since the stress field is formed primarily by the longitudinal shrinkage just like an iron hoop fasten the plate owing to almost the null outer constraint in this case. The longitudinal stresses (σ_{22}) reach the peak values at the weld beads owing to the longitudinal shrinkage.

| (a) RMS errors | (b) Max. errors of σ_{11} | (c) Max. errors of σ_{22} |
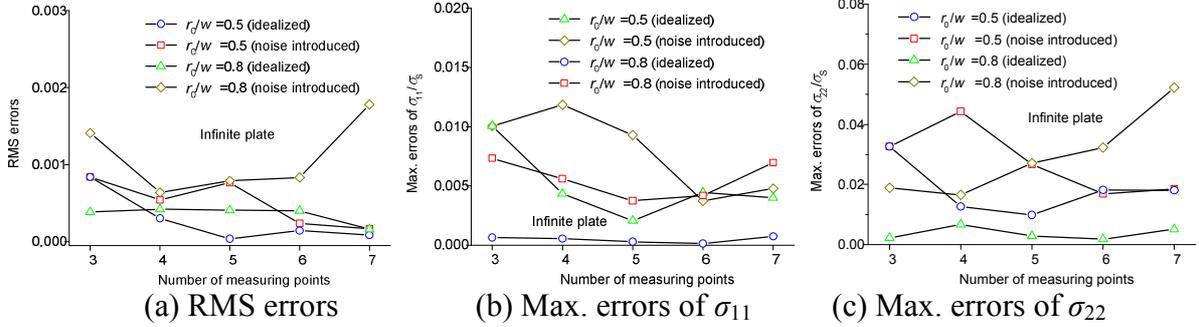
Figure 4. **The errors in infinite plates**

| (a) RMS errors | (b) Max. errors of σ_{11} | (c) Max. errors of σ_{22} |
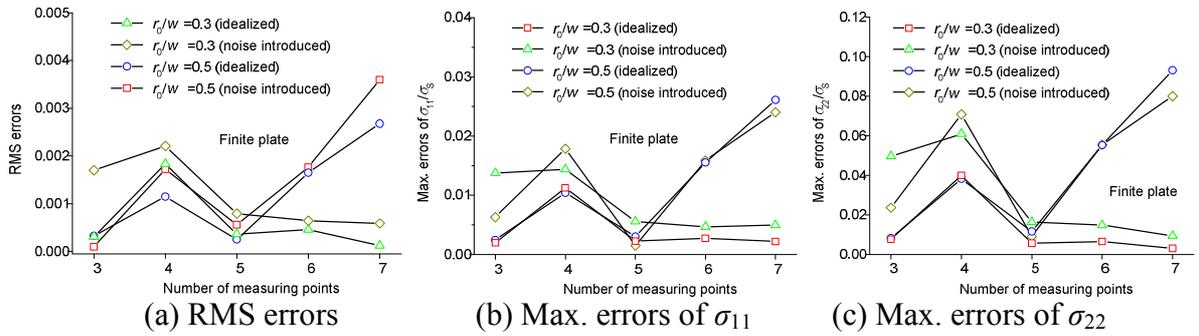
Figure 5. **The errors in finite plates**

The errors using the idealized and the noise introduced data for the infinite and the finite plates are given in Figs. 4 and 5, respectively, showing that the inverse approach is not too sensitive to the noises in the stress measurement. The residual stress fields can be reconstructed with the proposed approach by using only a small number of selected measuring points, for example, three pointes 1, 2, and 3 as shown in Fig. 1c, with which the computed stresses are drawn in Fig. 2. The computed results verify the feasibility and effectiveness of the present algorithm.

Conclusions

Using the computational model based on the eigenstrain boundary integral equations (BIE), an algorithm of inverse problem is investigated to determine the circular weld bead induced residual stresses, where the eigenstrain is considered to be the origin of residual stresses in structures. In order to reduce the number of unknowns and to consider the stability of inverse problem, the eigenstrains are approximated by low-order polynomials in the local area, divided into cells, around welded zones according to the features of welding. The sensitivity matrix in the inverse approach for evaluating the eigenstrain fields represented by the coefficients of polynomials is constructed with the aid of measured stresses in the domain after welding over a few selected measuring points. The residual stresses of circular weld beads in both the finite and infinite plates are evaluated in the numerical examples, showing that the proposed inverse approach is not too sensitive to the noises in the stress measurement, verifying the feasibility and effectiveness of the proposed approach.

References

Cao, Y.P., Hu, N., Lu, J., Fukunaga, H., and Yao, Z.H. (2002) An inverse approach for constructing the residual stress field induced by welding, *Journal of Strain Analysis for Engineering Design* **37**, 345-359.

Jun, T-S., and Korsunsky, A.M. (2010) Evaluation of residual stresses and strains using the eigenstrain reconstruction method, *International Journal of Solids and Structures* **47**, 1678-1686.

Korsunsky, A.M., Regino, G.M., Latham, D., Liu, J., Nowell, D., and Walsh, M. (2006) Eigenstrain analysis of synchrotron X-ray diffraction measurement of residual strains in machined nickel alloy plates, *Journal of Strain Analysis* **41**, 381-395.

Lee, C.H, and Chang, K.H. (2008) Three-dimensional finite element simulation of residual stresses in circumferential welds of steel pipe including pipe diameter effects, *Materials Science and Engineering* **A487**, 210-218.

Ma, H., and Deng, H.L. (1998a) Nondestructive determination of welding residual stresses by boundary element method, *Advances in Engineering Software* **29**, 89-95.

Ma, H., Kamiya, N., and Xu, S.Q. (1998b) Complete polynomial expansion of domain variables at boundary for two-dimensional elasto-plastic problems, *Engineering Analysis with Boundary Elements* **21**, 271-275.

Ma, H., Zhou, J., and Qin, Q.H. (2010) Boundary point method for linear elasticity using constant and quadratic moving elements, *Advances in Engineering Software* **41**, 480-488.

Ma, H., Wang, Y., and Qin, Q.H. (2012) Determination of welding residual stresses by inverse approach with eigenstrain formulations of BIE, *Applicable Analysis* **91**, 807-821.

Masubuchi, K. (1980) *Analysis of Welded Structures-Residual Stresses, Distortion and Their Consequences*, Pergamon Press, New York.

Prime, M.B. (1999) Residual stress measurement by successive extension of a slot: the crack compliance method, *Applied Mechanics Review* **52**, 75-96.

Qian, X.Q., Yao, Z.H., Cao, Y..P, and Lu, J. (2004) An inverse approach for constructing residual stress using BEM, *Engineering Analysis with Boundary Elements* **28**, 205-211.

Qian, X.Q., Yao, Z.H., Cao, Y.P., and Lu, J. (2005) An inverse approach to construct residual stresses existing in axisymmetric structures using BEM, *Engineering Analysis with Boundary Elements* **29**, 986-999.

Suterio, R., Albertazzi, A., and Amaral, F.K. (2006) Residual stress measurement using indentation and a radial electronic speckle pattern interferometer-recent progress, *Journal of Strain Analysis* **41**, 515-524.

Ueda, Y., Jirn, Y., and Yuan, M. (1986) Determination of residual stresses based on the sources of their formation, *Transactions of the Japan Welding Society* **6**, 59-64.

Three-dimensional welding residual stresses evaluation based on the eigen-strain methodology via X-ray measurements at the surface

*M. Ogawa[1]

[1]Faculty of Engineering, Yokohama National University, 79-1Tokiwadai, Hodogaya-ku, Yokohama-shi, Kanagawa-ken, 240-8501, Japan

*Corresponding author: mogawa@ynu.ac.jp

Abstract

In order to assure structural integrity for operating welded structures, it is necessary to evaluate crack growth rate and crack propagation direction for each observed crack nondestructively. Here, three dimensional welding residual stresses must be evaluated to predict crack propagation. Today, X-ray diffraction is used and the ultrasonic method has been proposed as nondestructive method to measure residual stresses. However, it is impossible to determine residual stress distributions in the thickness direction. Although residual stresses through a depth of several tens of millimeters can be evaluated nondestructively by neutron diffraction, it cannot be used as an on-site measurement technique. It is because neutron diffraction is available only in special irradiation facilities. Author pays attention to the bead flush method based on the eigen-strain methodology. In this method, three-dimensional welding residual stresses are calculated by an elastic FEM (Finite Element Method) analysis from eigen-strains which are evaluated by an inverse analysis from released strains by strain gauges in removal of reinforcement of weld. Here, the removal of the excess metal can be regarded as nondestructive treatment essentially because toe of weld which may become crack starters can be eliminated. The effectiveness of the method has been proved for welded plates and pipes even with relatively lower bead height. In actual measurements, stress evaluation accuracy becomes poorer because measured values of strain gauges are affected by processing strains on the machined surface. In the previous studies, the author has developed the bead flush method that is free from the influence of the affecting strains by using residual strains on surface by X-ray diffraction. However, stress evaluation accuracy is not good enough because of relatively poor measurement accuracy of X-ray diffraction. In this study, a method to improve the estimation accuracy of residual stresses in this method is formulated, and it is shown numerically that inner welding residual stresses can be estimated accurately from the measured residual strains by X-ray diffraction.

Keywords: Eigenstrain, Weld, Residual stress, X-ray diffraction, Bead flush method, Three-dimensional evaluation,

Introduction

In order to assess structural integrity for operating welded structures, it is important to evaluate three-dimensional welding residual stresses non-destructively to predict crack propagating for observed cracks in in-service inspection. Today, there are some techniques to estimate three-dimensional residual stresses such as neutron diffraction methods [Suzuki and Akita (2009)], welding simulation via thermal elastic-plastic FEM analysis [Yaghi et al. (2013)] and techniques based on the eigen-strain methodology [Mura (1978)]. However, neutron diffraction is unavailable to use as an on-site measurement application because it can be used only in special irradiation facilities. Furthermore, measured stresses from diffraction methods including X-ray diffraction and high energy X-ray diffraction techniques cannot be input into the FEM model that has been used in assessment of structural integrity at the time of the design. It is because all the 6 stress components

that satisfy the self-equilibrium condition cannot be measured. Also, it is difficult to predict crack propagation via FEM [Kikuchi et al. (2009)] from estimated residual stresses by diffraction methods. Although welding residual stresses can be estimated non-destructively by using welding simulation, estimation accuracy may be poorer due to the difficulty of determining the parameters depended on temperature. To make matters worse, piece-to-piece variations have to be neglected in qualitative evaluation via thermal elastic-plastic FEM analysis. On the other hand, three-dimensional residual stress distribution can be estimated quantitatively by FEM analysis by using the eigen-strain methodology. For example, the cutting method [Ueda et al. (1975); Ueda et al. (1979)] based on the eigen-strain methodology has been proposed. In this method, residual stresses are determined by an elastic FEM analysis from eigen-strains which are calculated by an inverse analysis [Kubo (1992)] from released strains through sectioning. Here, eigen-strains are defined as a sum of inelastic strains [Mura (1978)] and can be regarded as the cause of residual stresses and elastic strains. Note that they are not always equal to inherent strains which are a total of physical inelastic strains such as thermal, plastic and transformation strains [Masuda and Nakamura (2010a; 2010b)]. Although structures have to be wasted by the cutting method, welding residual stresses can be evaluated non-destructively by the bead flush method [Nakamura et al. (1995)]. In this method, eigen-strains are estimated from released strains in removal of reinforcement of weld. Since toe of weld may become crack starters, the removal of the excess weld metal can be regarded as a preferable treatment. The effectiveness of this method has been proved numerically for welded plates [Kumagai et al. (2000)]. In addition, statistical range of residual stress distributions has been accumulated successfully for welded pipes even with lower bead height [Ogawa and Nakamura (2011a; 2011b)]. In actual measurement, however, processing strains are created after machining the reinforcement of the weld. In this case, stress evaluation accuracy becomes poorer because measured values of strain gauges are affected by the processing strains. In order to solve the difficulties, the bead flush method has been developed to be free from the influence of the affecting strains [Ogawa (2013)]. In this method, not only welding eigen-strains but also processing strains are estimated non-destructively from residual strains on surface by X-ray diffraction instead of released strains by strain gauges (Fig. 1). However, estimation accuracy in this method is not higher due to relatively poor measurement accuracy of X-ray diffraction.

In this study, numerical formula to be able to use the measured residual strains on the weld metal after the removal as additional source of information is shown. And, numerical simulation is carried out to prove the effectiveness in this method.
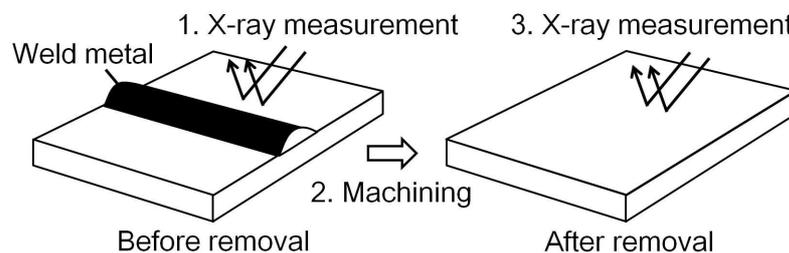
Figure 1. Procedures in the advanced bead flush method

Analytical Procedures

Formulation of the Bead Flush Method

In general, the elastic strains $\{\varepsilon_e\}$ of the concerned elements and the eigen-strains $\{\varepsilon_e^*\}$ can be related as:

$$\{\varepsilon_e\} = [R_e]\{\varepsilon_e^*\} \qquad (1)$$

where $[\boldsymbol{R}_e]$ is an elastic response matrix. And, the i-th column of it can be obtained by imposing an unit eigen-strain vector to an i-th component of $\{\boldsymbol{\varepsilon}_e^*\}$ as shown below:

$$\{{}_{\text{unit}}\boldsymbol{\varepsilon}_e^*\}_i = \{0, \cdots, \varepsilon_{ei}^* = 1, \cdots, 0\}^{\text{T}} \tag{2}$$

Therefore, elastic strains before and after removals of excess metal can be described as follows:

$$\{\boldsymbol{\varepsilon}_{\text{eb}}\} = [\boldsymbol{R}_{\text{eb}}]\{\boldsymbol{\varepsilon}_{\text{eb}}^*\} \tag{3}$$

$$\{\boldsymbol{\varepsilon}_{\text{ea}}\} = [\boldsymbol{R}_{\text{ea}}]\{\boldsymbol{\varepsilon}_{\text{ea}}^*\} \tag{4}$$

where the subscripts b and a denote the before and after removals, respectively. Since it is based on the assumption that eigen-strains are constant through machining, the released strain vectors $\{\Delta\boldsymbol{\varepsilon}_e\}$ are given by the following equations:

$$\{\Delta\boldsymbol{\varepsilon}_e\} = \{\boldsymbol{\varepsilon}_{\text{ea}}\} - \{\boldsymbol{\varepsilon}_{\text{eb}}\}$$

$$= ([\boldsymbol{R}_a] - [\boldsymbol{R}_b])\{\boldsymbol{\varepsilon}^*\}$$

$$= [\boldsymbol{R}]\{\boldsymbol{\varepsilon}^*\} \tag{5}$$

where $[\boldsymbol{R}] = [\boldsymbol{R}_a] - [\boldsymbol{R}_b]$ and $\{\boldsymbol{\varepsilon}^*\} = \{\boldsymbol{\varepsilon}_b^*\} = \{\boldsymbol{\varepsilon}_a^*\}$. In actual measurements, measured released stains by strain gauges include measurement errors $\{\Delta\boldsymbol{\varepsilon}_{\text{err}}\}$. In this case, measured released strain vector $\{\Delta\boldsymbol{\varepsilon}_{\text{em}}\}$ is written as follows:

$$\{\Delta\boldsymbol{\varepsilon}_{\text{em}}\} = [\boldsymbol{R}]\{\boldsymbol{\varepsilon}^*\} + \{\Delta\boldsymbol{\varepsilon}_{\text{err}}\} \tag{6}$$

The most probable values of estimated eigen-strain vector $\{\boldsymbol{\varepsilon}_{\text{est}}^*\}$ is described by the least square method as follows:

$$\{\boldsymbol{\varepsilon}_{\text{est}}^*\} = [\boldsymbol{R}]^+\{\Delta\boldsymbol{\varepsilon}_{\text{em}}\} \tag{7}$$

where $[\boldsymbol{R}]^+$ is the Moore and Penrose generalized inverse matrix [Kubo (1992)] of $[\boldsymbol{R}]$, and it is written as:

$$[\boldsymbol{R}]^+ = [\boldsymbol{R}]^{\text{T}}[\boldsymbol{R}]([\boldsymbol{R}]^{\text{T}}[\boldsymbol{R}][\boldsymbol{R}]^{\text{T}}[\boldsymbol{R}])^-[\boldsymbol{R}]^{\text{T}} \tag{8}$$

Improvement of the Bead Flush Method

In the conventional bead flush method, excess metal has to be eliminated without affecting strains. Once processing strains $\{\boldsymbol{\varepsilon}_p^*\}$ are created on a sample, measurement accuracy of release stains is worsened as shown below:

$$\{\Delta\boldsymbol{\varepsilon}_{\text{em}}\} = [\boldsymbol{R}]\{\boldsymbol{\varepsilon}^*\} + \{\Delta\boldsymbol{\varepsilon}_{\text{err}}\} + [\boldsymbol{R}_a]\{\boldsymbol{\varepsilon}_p^*\} \tag{9}$$

In order to improve this problem, the author has proposed the following equations instead of Eq. (7) [Ogawa (2013)].

$$\{\boldsymbol{\varepsilon}_{\text{est}}^* \ \boldsymbol{\varepsilon}_{\text{p\_est}}^*\}^{\text{T}} = [\boldsymbol{R}_{\text{ab}}]^+\{\boldsymbol{\varepsilon}_{\text{ebm}} \ \boldsymbol{\varepsilon}_{\text{eam}}\}^{\text{T}} \tag{10}$$

$$[\boldsymbol{R}_{\text{ab}}] = \begin{bmatrix} \boldsymbol{R}_b & \boldsymbol{0} \\ \boldsymbol{R}_a & \boldsymbol{R}_a \end{bmatrix} \tag{11}$$

where $\{\boldsymbol{\varepsilon}_{\text{ebm}}\}$ and $\{\boldsymbol{\varepsilon}_{\text{eam}}\}$ are measured residual strains before and after removals, respectively. And, these two residual strains can be measured non-destructively by X-ray diffraction. Therefore, it is

possible to obtain estimated values of both welding eigen-strains $\{\boldsymbol{\varepsilon}_{\text{est}}^{*}\}$ and processing strains $\{\boldsymbol{\varepsilon}_{\text{p\_est}}^{*}\}$ non-destructively by using Eqs. (10) and (11).

Additionally, in this study, measured strains on the weld metal after machined $\{\boldsymbol{\varepsilon}_{\text{wam}}\}$ are added to Eqs. (10) and (11) to increase measurement information as shown below:

$$\{\boldsymbol{\varepsilon}_{\text{est}}^{*}\boldsymbol{\varepsilon}_{\text{p\_est}}^{*}\}^{\text{T}} = [\boldsymbol{R}_{\text{abw}}]^{+}\{\boldsymbol{\varepsilon}_{\text{ebm}}\boldsymbol{\varepsilon}_{\text{eam}}\boldsymbol{\varepsilon}_{\text{wam}}\}^{\text{T}} \tag{12}$$

$$[\boldsymbol{R}_{\text{abw}}] = \begin{bmatrix} \boldsymbol{R}_{\text{b}} & \boldsymbol{0} \\ \boldsymbol{R}_{\text{a}} & \boldsymbol{R}_{\text{a}} \\ \boldsymbol{R}_{\text{a}} & \boldsymbol{R}_{\text{a}} \end{bmatrix} \tag{13}$$

Note that residual strains on surface can be obtained non-destructively by the EBSD (electron backscatter diffraction patterns) method [Wilkinson (1996)] instead of X-ray diffraction.

Numerical Simulation

In this study, numerical simulation in the bead flush method based on the eigen-strain method is conducted to show the effectiveness of this method.

FEM Model

As shown in Fig. 2, a half of a butt-welded plate without geometrical restrictions at both ends was used as FEM model. The plate length, thickness and width are 120mm, 10mm and 60mm, respectively. The bead width is 8mm and its height is 0.3mm. Solid element that has 8 nodes and 3 degrees of freedom were applied. The total nodes and elements of the model are 3349 and 2544, respectively. Young's modulus and Poisson's ratio were set at 200GPa and 0.265, respectively. A commercial software, ANSYS (CYBERNET SYSTEMS CO., LTD., Japan), was used here.
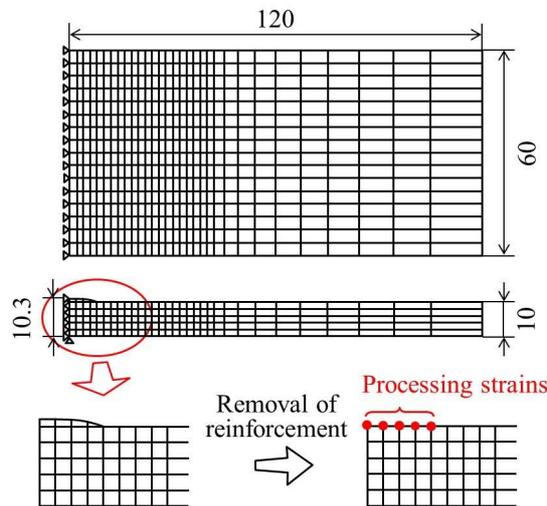
Figure 2. FEM model [Ogawa (2013)]

Exact Distribution

Exact eigen-strain distributions assumed in this simulation were quoted from the research report by Kumagai et al. [Kumagai et al. (1999)] in which eigen-strains were determined on the basis of the experimental results in the cutting method (Fig. 3). Here, the exact eigen-strains are distributed uniformly in the welding and thickness directions. Three dimensional exact residual stress distributions can be calculated from exact eigen-strains by elastic FEM analysis. For example, exact

stresses at the middle in the welding line on the bottom surface (*x*=30mm and *z*=0mm) are seen in Fig. 4. Here, *x*, *y* and *z* directions are the welding, perpendicular the welding and thickness directions, respectively.
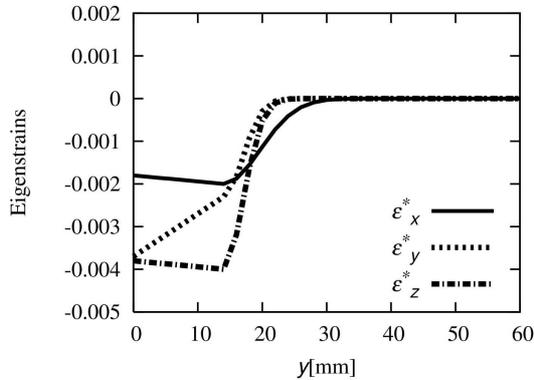
Figure 3. Exact eigen-strain distributions
[Ogawa (2013)]
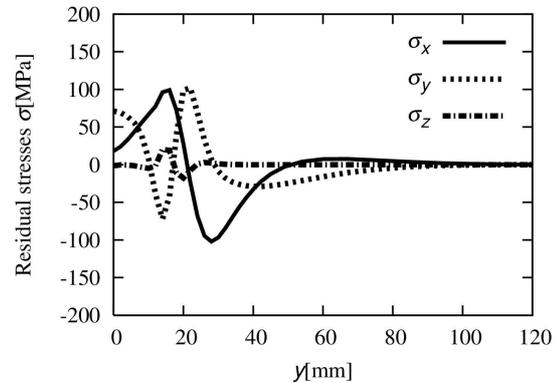
Figure 4. Exact residual stress distributions
(*x*=30mm on the bottom surface) [Ogawa (2013)]

Procedure to Evaluate Estimation Accuracy

First, exact residual strain distributions for the whole structure are calculated by inputting exact eigen-strains in the FEM model as initial strains. Second, measured residual strains in the *x* and *y* directions at measurement points are obtained by adding measurement errors to exact residual strains. Here, measurement points on the top surface (*z*=10mm) on the base metal and weld metal are shown in Fig. 5 and Fig. 6, respectively. Third, estimated eigen-strains were computed by an inverse analysis. And, residual stresses on the bottom surface (*z*=0mm) are calculated to compare exact residual stresses. In this analysis, -500μ eigen-strains in the *x* direction were added evenly on the machined surface as processing strains. It is based on the assumption that micro cutter was used to remove the reinforcement of the weld [Chen et al. (1996)].
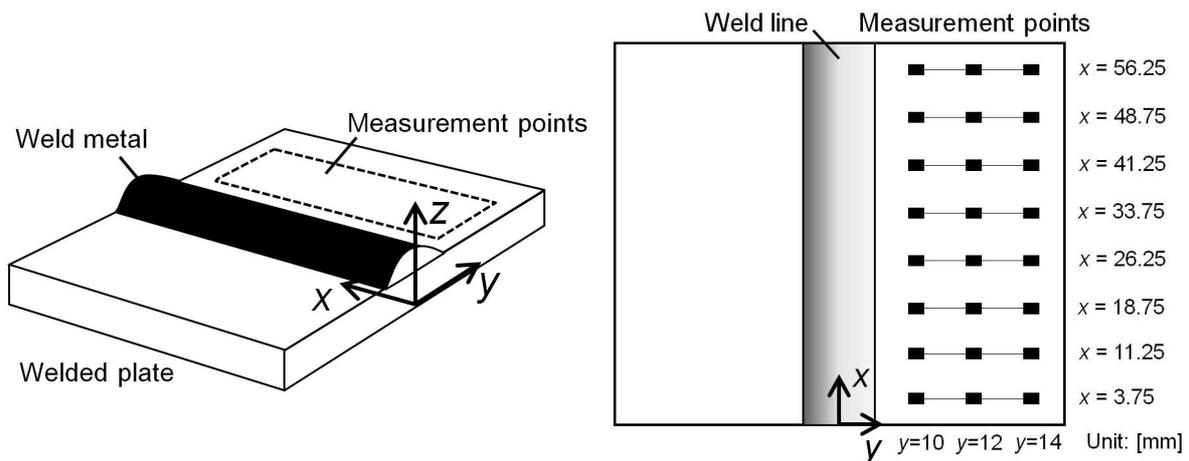
Figure 5. Measurement points on the base metal

Stabilization of Solution in Inverse Analysis

In order to reduce unknown parameters in this inverse analysis, welding eigen-strain distributions in each direction were expressed by the four logistic functions [Kumagai et al. (1999)] as:

Measurement points on the weld metal
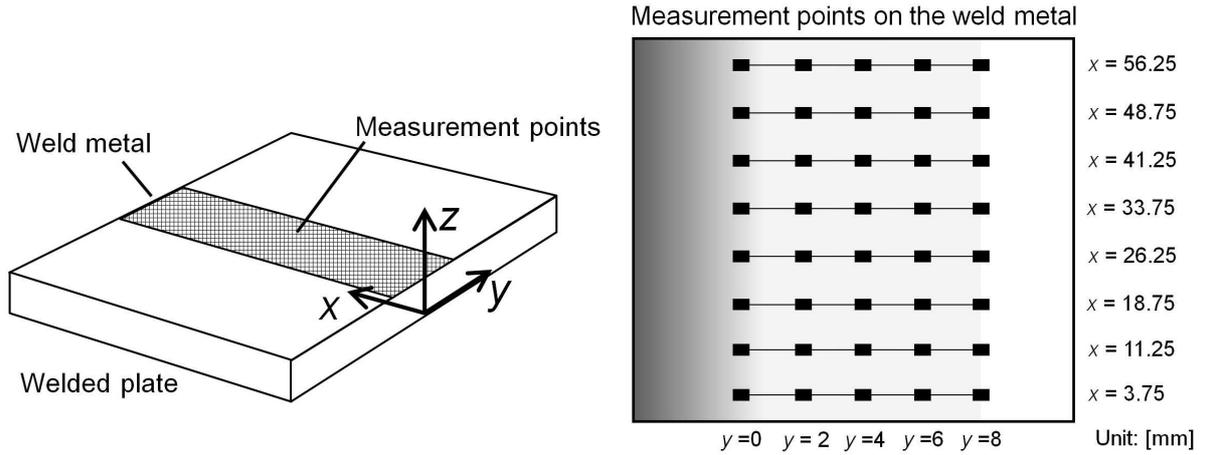
Figure 6. Measurement points on the weld metal

$$\{\varepsilon_s^*\}(y) = \sum_{i=1}^{4} \frac{\{a_{si}\}}{1 + \exp(p + q_i y)} \tag{14}$$

$$p = -5.0, \ q_1 = 0.60, \ q_2 = 0.40, \ q_3 = 0.30, \ q_4 = 0.25 \tag{15}$$

where the subscript s denotes x, y and z directions, respectively. Constant values p and q_i were determined in consideration of that welding eigen-strains were distributed less than 40mm in the y direction [Ueda et al. (1993)]. $\{a_{si}\}$ is a vector of unknown parameters. In addition, it was assumed that eigen-strains were constant in the welding and the thickness directions. Therefore, total number of unknown parameters of welding eigen-strains becomes twelve (4 functions × 3 directions). Furthermore, processing strains on the machined surface were considered as constant in the welding direction. The total number of unknown parameter becomes fifteen (5 points in each direction).

In order to stabilize solutions, the artificial noise method was used [Ogawa and Nakamura (2011b)]. When $[R_{abw}]$ in Eq. (12) is an $N \times M$ matrix with rank n, it can be decomposed as follows:

$$[R_{abw}]^+ = [U][B]^-[V]^T \tag{16}$$

$$[B]^- = \begin{bmatrix} B_n^- & 0 \\ 0 & 0 \end{bmatrix}, \ [B_n]^- = \begin{bmatrix} 1/\mu_1 & & & 0 \\ & 1/\mu_2 & & \\ & & \ddots & \\ 0 & & & 1/\mu_n \end{bmatrix}, \ \mu_1 \geq \mu_2 \geq \cdots \geq \mu_n \geq 0 \tag{17}$$

where $[U]$, $[B]$ and $[V]^T$ are $N \times N$, $N \times M$ and $M \times M$ matrixes, respectively. The values of μ_j ($1 \leqq j \leqq n$) are termed as singular values of $[B]$. Solutions become sensitive if singular values are smaller.

In the artificial noise method, $[B_n]^-$ matrix is replaced by $[B_{n\_\gamma}]^-$ as shown below:

$$[B_{n\_\gamma}]^- = ([B]^2 + \gamma[I])^-[B]^T \tag{18}$$

$$= \begin{bmatrix} \dfrac{\mu_1}{\mu_1^2 + \gamma} & & & 0 \\ & \dfrac{\mu_2}{\mu_2^2 + \gamma} & & \\ & & \ddots & \\ 0 & & & \dfrac{\mu_n}{\mu_n^2 + \gamma} \end{bmatrix} \tag{19}$$

where $[I]$ is unit matrix. Solutions can be stabilized by increasing the real parameter γ.

Results and Discussion

Figure 7 shows estimation accuracy of welding residual stresses on the bottom surface from residual strains on the base metal (Fig. 5) and on the base and weld metals (Figs. 5 and 6). In this analysis, it was assumed that observation error follows the normal distribution whose average was 0 and standard deviation was set as 500μ because measurement accuracy of X-ray diffraction for welded joints was about ±100MPa [Kurimura and Akiniwa (2009)]. As we can see in Fig. 7, welding residual stresses cannot be estimated accurately from residual strains just on base metal. On the other hand, stress evaluation accuracy can be improved successfully when measured strains on the weld metal are used as additional source of information. Here, the L-curve method [Hansen (1992)] was used to determine the value of the artificial noise.

(a) Residual stresses in the x direction (b) Residual stresses in the y direction
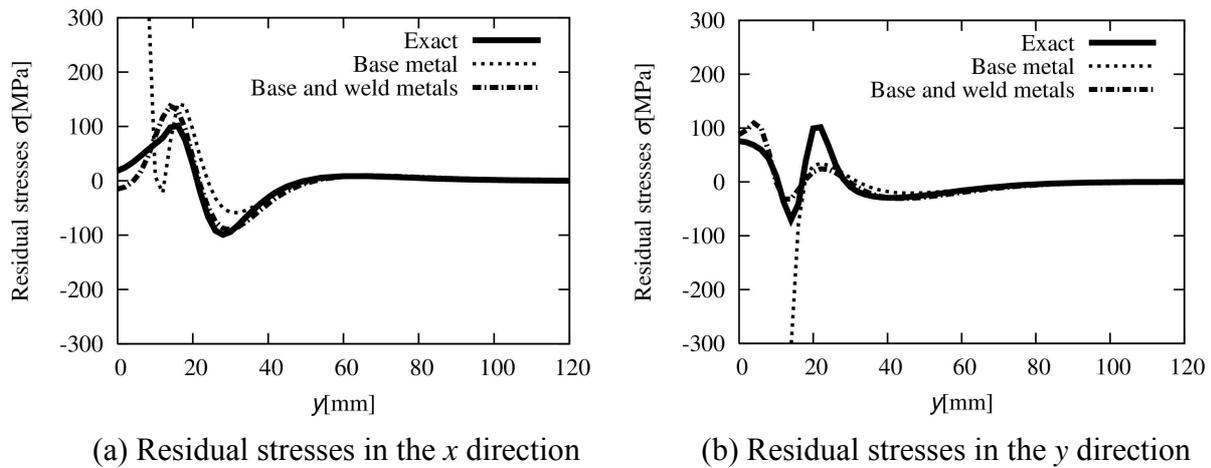
Figure 7. Estimated residual stresses on the bottom surface after the removal. The dotted and chain lines are estimated results from residual strains on the base metal and on the base and weld metals, respectively.

Conclusions

In the previous study, the author developed the bead flush method that is free from the influence of processing strains in machining. However, stress evaluation accuracy is relatively poor especially in the vicinity of the weld line when measurement errors by X-ray diffraction are considered. In this study, mathematical expressions to be able to use measured strains on weld metal as additional source of information was shown. And, numerical simulation for butt welded plate was carried out to prove the effectiveness of this method. It was clarified that estimation accuracy of residual stresses especially near the weld line could be improved successfully by using this method.

Acknowledgment

References

Suzuki, H. and Akita, K. (2009) New Developments of Neutron Stress Measurement (lecture), *Journal of the Society of Materials Science, Japan* **58**, 1051-1057. (in Japanese)

Yaghi, A. H., Hyde, T. H., Becker, A. A. and Sun, W. (2013) Finite element simulation of residual stresses induced by the dissimilar welding of a P92 steel pipe with weld metal IN625, *International Journal of Pressure Vessel and Piping* **111-112**, 173-186.

Mura, T. (1987) Micromechanics of Defects in Solids, *Martinus Nijhoff Publishers*, 1-15.

Kikuchi, M., Wada, Y., Shimizu, Y. and Suyama, H. (2009) Crack Growth Analysis in Weld-Heat Affected Zone Using S-Version FEM, *Journal of the Japan Society of Mechanical Engineers, Series A* **75**, 1381-1386. (in Japanese)

Ueda, Y., Fukuda, K., Nakacho, K. and Endo, S. (1975) A New Measuring Method of Residual Stresses with the Aid of Finite Element Method and Reliability of Estimated Values, *The Society of Naval Architects of Japan* **138**, 499-507. (in Japanese)

Ueda, Y., Fukuda, K. and Tanigawa, M. (1979) New Measuring Method of 3-Dimensional Residual Stresses Based on Theory of Inherent Strain, *The Society of Naval Architects of Japan* **145**, 203-211. (in Japanese)

Kubo, S. (1992) Inverse Problem, *Baifukan*. (in Japanese)

Masuda, K. and Nakamura, H. (2010a) Improvement of the Inverse Analysis Approaches for Assessment of Welding Deformations and Residual Stresses by Using Thermo Elasto-Plastic Welding Simulation (1st Report, Stress Analysis for the Bead Flush Method), *Journal of the Japan Society of Mechanical Engineers, Series A* **76**, 884-892. (in Japanese)

Masuda, K. and Nakamura, H. (2010b) Improvement of the Inverse Analysis Approaches for Assessment of Welding Deformations and Residual Stresses by Using Thermo Elasto-Plastic Welding Simulation (2nd Report, Deformation Analysis), *Journal of the Japan Society of Mechanical Engineers, Series A* **76**, 1186-1194. (in Japanese)

Nakamura, H., Naka, Y., Park, W. and Kobayashi, H. (1995) Computer Aided Non-Destructive Evaluation Method of Welding Residual Stresses by Removing Reinforcement of Weld, *Current Topics in Computational Mechanics, ASME Pressure Vessel and Piping* **305**, 49-56.

Kumagai, K., Nakamura, H. and Kobayashi, H. (2000) Computer Aided Nondestructive Evaluation Method of Welding Residual Stresses by Removing Reinforcement of Weld (Proposal of a New Concept and Its Verification), *Fatigue and Fracture Mechanics, ASME STP 1389* **31**, 128-142.

Ogawa, M. and Nakamura, H. (2011a) Proposal of an Estimation Method of Welding Residual Stresses in Welded Pipes for Risk-Analysis-Based Assurance of Structural Integrity (1st Report: Improvement of Accuracy by Pre-Treatment of Strain Data), *Journal of the Japan Society of Mechanical Engineers, Series A* **77**, 271-281. (in Japanese)

Ogawa, M. and Nakamura, H. (2011b) Proposal of an Estimation Method of Welding Residual Stresses in Welded Pipes for Risk-Analysis-Based Assurance of Structural Integrity (2nd Report: Application of the L-Curve and the Artificial Noise Methods), *Journal of the Japan Society of Mechanical Engineers, Series A* **77**, 282-292. (in Japanese)

Ogawa, M. (2013) Proposal of an Eigen-Strain Estimation for Determination of Residual Stresses Considering the Influence of Inelastic Strains by Machining or Heating, *Journal of the Japan Society of Mechanical Engineers, Series A* **79**, 1266-1277. (in Japanese)

Wilkinson, A. J. (1996) Measurement of elastic strains and small lattice rotation using electron back scatter diffraction, *Ultramicroscopy* **62,** 237-247.

Kumagai, K., Nakamura, H. and Kobayashi, H. (1999) Computer Aided Nondestructive Evaluation Method of Welding Residual Stresses by Removing Reinforcement of Weld (Proposal of a New Concept and Its Application to Butt Welded Plate), *Journal of the Japan Society of Mechanical Engineers, Series A* **65**, 133-140. (in Japanese)

Chen, J., Terasaki, T., Akiyama, T. and Kishitake, K. (1996) Effect of Layer Removal Methods on Residual Stress Estimated by Inherent Strain Method, *Quarterly journal of the Japan Welding Society* **14**, 762-767. (in Japanese)

Ueda, T., Ma, N. X. and Koki, R. (1993) TL_yL_z Method and T Method for Measurement of 3 Dimensional Residual Stresses in Bead-on-plate Welds (Report 2: Estimation and Measurement Methods of Residual Stresses Using Inherent Strain Distribution Described as Functions), *Japan Welding Society* **11**, 555-562. (in Japanese)

Kurimura, T. and Akiniwa, Y. (2009) Residual Stress of Weld Section and X-Ray Residual Stress Measurement (lecture), *Journal of the Society of Materials Science, Japan* **58**, 873-878. (in Japanese)

Hansen, P. C. (1992) Analysis of Discrete Ill-Posed Problems by Means of the L-Curve, *SIAM Review* **34**, 561-580.

An edge-based smoothed three-node Mindlin plate element (ES-MIN3) for static and free vibration analyses of plates

*Y.B. Chai[1,2], W. Li[1,2], and Z.X. Gong[1,2]

[1]School of Naval Architecture and Ocean Engineering, Huazhong University of Science and Technology, China
[2] Hubei Key Laboratory of Naval Architecture & Ocean Engineering Hydrodynamics (HUST), Huazhong University of Science and Technology, Wuhan City, Hubei 430074, China
*Corresponding author: cybhust@hust.edu.cn

Abstract

The smoothed finite element method (S-FEM) developed recently shows great efficiency in solving solid mechanics. This paper extends an edge-based smoothed finite element method for static and free vibration analyses of plates. The edge-based strain smoothing technique is combined with the three-node Mindlin plate element (MIN3) to give a so-called the edge-based smoothed MIN3(ES-MIN3). The system stiffness matrix is calculated by using the edge-based strain smoothing technique over the smoothing domains associated with the edges of elements. In each element the stabilized MIN3 is performed to avoid the transverse shear locking. Typical numerical examples demonstrate that the present ES-MIN3 is free of shear locking and can achieve the high accuracy compared to the exact solutions and others existing plate elements.

Keywords: Shear locking, Finite element method (FEM), Edge-based smoothed three-node Mindlin plate element (ES-MIN3), Strain smoothing technique

Introduction

Nowsdays, the plate structures have been used widely in many branches of structural engineering problems. Owing to limitations of the analytical methods, many different numerical methods, such as finite difference method, finite element method, boundary element method, meshfree method etc, have been proposed to analyze the plate structures. Among them the finite element method (FEM) is one of the most popular numerical methods to simulate the behaviors of plate structures. In the practical application, many plate elements based on the Reissner-Mindlin theory using the first-order shear deformation are preferred due to its simplicity and efficiency[Henry and Saigal (2000);Reddy (2006)]. These Reissner-Mindlin plate elements usually possess high accuracy and fast convergence speed for displacement, however, they also suffer from the 'shear locking' phenomenon which has the root of incorrect transverse forces under bending and induces over-stiffness as the plate become progressively thinner.

In order to eliminate shear locking and to increase the accuracy and stability of the solution, many new numerical techniques and effective modifications have been proposed, such as the mixed

formulation/hybrid elements[Lee and Wong (1982); Zienkiewicz and Lefebvre (1988); Miranda and Ubertini (2006); Auricchio and Taylor(1994); Lovadina (1998)] proposed by Lee et al and Miranda et al; the enhanced assumed strain method (EAS) [Simo and Rifai (1990); Simo et al. (1989);] proposed by Simo et al and the assumed natural strain (ANS) method[Tessler and Hughes (1985); Bathe and Dvorkin (1985); Batoz and Lardeur (1989)] proposed by Hughes et al. Recently. Bletcinger et al proposed the discrete shear gap method [Bletzinger et al. (2000)] to avoid transverse shear locking and to improve the accuracy of the present formulation. In fact, the DSG also can be classified as an ANS element. It is similar to the ANS methods in the terms of modifying the course of certain strains within the element, but is different in the aspect of removing of collocation points. The DSG can work well for different elements.

Also based on the ANS method, a three-node Mindlin plate element (MIN3), which avoids shear locking, was proposed by Tessler and Hhghes. In MIN3, a complete quadratic deflection field is constrained by continuous shear edge constraints. The numerical examples demonstrated that the MIN3 is free of shear locking and can achieve convergent solutions.

Recently, Liu et al have proposed a series of smoothed finite element method (S-FEM) by incorporating the strain smoothing technique[Chen et al. (2001)] of meshfree methods into the standard finite element method. In these S-FEM models, the compatible strain fields are smoothed based on the smoothing domains created from the entities of the element mesh such as cells (CS-FEM)[Liu et al. (2001); Nguyen (2008; 2012; 2013a;2013b); Wu and Wang (2013)], or nodes (NS-FEM) [Liu et al. (2009a; 2009b); Nguyen (2011)], or edges (ES-FEM)[Liu et al. (2009c); Nguyen (2009); Li et al. (2012; 2013)],or faces (FS-FEM)[Feng et al. (2013)], then the smoothed Galerkin weak forms are evaluated based on these smoothing domains. The S-FEM models can improve significantly the accuracy of solid mechanics owing to the strain smoothing technique on the smoothing domains.

In this paper, the edge-based strain smoothing technique is incorporated with the well-known three node Mindlin plate (MIN3) to give a so-called edge-based smoothed MIN3 (ES-MIN3). In the ES-MIN3 models, the calculation of the system stiffness matrix is performed using strain smoothing technique over the smoothing domains associated with the edges of elements. The numerical results show that present method is immune from shear locking and can achieve high accurate solutions in static and vibration analysis of the Reissner-Mindlin plate.

Governing equations and weak form for the Reissner-Mindlin plate

Consider a plate under bending deformation as shown in Figure.1. The middle (neutral) surface of plate oxy is chosen as the reference plane that occupies a domain $\Omega \subset R^2$. Let w be the deflection of

the plate and $\beta^T = (\beta_x, \beta_y)$ be the rotations of the normal to the middle surface of the plate around y-axis and x-axis, respectively. Then the unknown vector of three independent field variables at any point in the problem domain of the Reissner-Mindlin plates can be written as:

$$\mathbf{u}^T = [w \quad \beta_x \quad \beta_y] \tag{1}$$
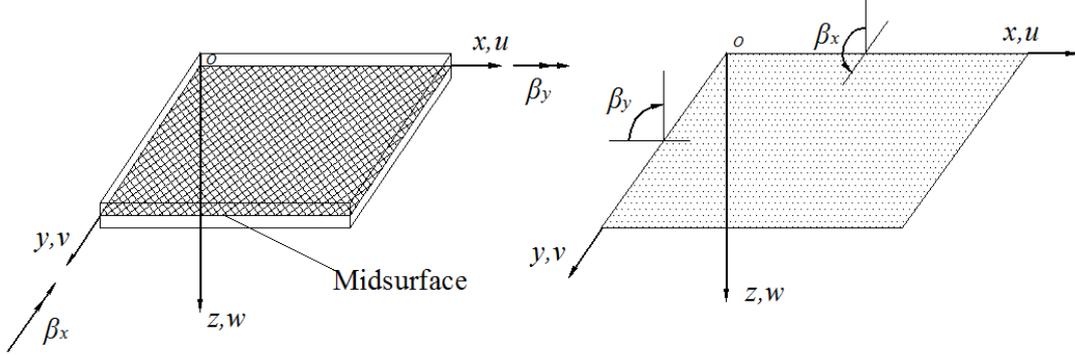
Figure.1. positive directions of displacement u, v, w and two rotation β_x, β_y for Reissner-Mindlin plate

Here we assume that the material is homogeneous and isotropic with Young's modulus E and Poisson's ratio v. The governing differential equations of the static Reissner-Mindlin plate can be expressed as:

$$\nabla \cdot \mathbf{D}_b \mathbf{\kappa}(\mathbf{\beta}) + Gkt\mathbf{\gamma} = 0 \quad \text{in } \Omega \tag{2}$$

$$Gkt\nabla \cdot \mathbf{\gamma} + p = 0 \quad \text{in } \Omega \tag{3}$$

$$w = \bar{w}, \quad \mathbf{\beta} = \bar{\mathbf{\beta}} \quad \text{on } \Gamma = \partial\Omega \tag{4}$$

in which t is the plate thickness and $p = p(x, y)$ is a distributed load per an area unit, G and $k = 5/6$ are shear modulus and shear correction factor, respectively, $\mathbf{D}_b$ is the bending stiffness constitutive, $\mathbf{\kappa}$ and $\mathbf{\gamma}$ are the bending and shear strains, respectively, defined by

$$\kappa = \mathbf{L}_d\mathbf{\beta}, \quad \mathbf{\gamma} = \nabla w + \mathbf{\beta} \tag{5}$$

where $\nabla = (\partial/\partial x, \partial/\partial y)$ is the gradient vector and $\mathbf{L}_d$ denotes a matrix of differential operators:

$$L_d = \begin{bmatrix} \dfrac{\partial}{\partial x} & 0 & \dfrac{\partial}{\partial y} \\ 0 & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial x} \end{bmatrix}^T \tag{6}$$

The standard Galerkin weakform of the static equilibrium equations for the Reissner-Mindlin plate is given by

$$\int_\Omega \delta\mathbf{\kappa}^T \mathbf{D}_b \mathbf{\kappa} d\Omega + \int_\Omega \delta\mathbf{\gamma}^T \mathbf{D}_s \mathbf{\gamma} d\Omega = \int_\Omega \delta w p d\Omega \tag{7}$$

where the bending stiffness constitutive coefficients $\mathbf{D}_b$ and the transverse shear stiffness constitutive coefficients $\mathbf{D}_s$ are defined as

$$\mathbf{D}_b = \frac{Et^3}{12(1-v^2)} \begin{bmatrix} 1 & v & 0 \\ v & 1 & 0 \\ 0 & 0 & 1-v/2 \end{bmatrix}, \qquad \mathbf{D}_s = ktG \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{8}$$

For the free vibration analysis of Reissner-Mindlin plates, the standard Galerkin weak form of the dynamic equilibrium can be written as

$$\int_\Omega \delta\boldsymbol{\kappa}^T \mathbf{D}_b \boldsymbol{\kappa} d\Omega + \int_\Omega \delta\boldsymbol{\gamma}^T \mathbf{D}_s \boldsymbol{\gamma} d\Omega + \int_\Omega \delta\mathbf{u}^T \mathbf{m}\ddot{\mathbf{u}} d\Omega = 0 \tag{9}$$

where $\mathbf{m}$ is the mass matrix of Reissner-Mindlin plate

$$\mathbf{m} = \rho \begin{vmatrix} t & 0 & 0 \\ 0 & \dfrac{t^3}{12} & 0 \\ 0 & 0 & \dfrac{t^3}{12} \end{vmatrix} \tag{10}$$

where ρ denotes the mass density of the material.

General FEM formulation for Reissner-Mindlin plate elements

In the process of the FEM formulation of the plate, the problem domain Ω is discretized into N_e finite elements such that $\Omega_1 \cup \Omega_2 \cup \Omega_3, \cdots, \Omega_{N_e} = \Omega$ and $\Omega_i \cap \Omega_j = \varnothing$, $i \neq j$, where N_e is the number of total elements. Then the finite element solution $\mathbf{u} = \lfloor w \quad \beta_x \quad \beta_y \rfloor$ of a displacement model for the Reissner-Mindlin plate can be expressed as

$$\mathbf{u} = \sum_{I=1}^{N_n} \begin{bmatrix} N_I(\mathbf{x}) & 0 & 0 \\ 0 & N_I(\mathbf{x}) & 0 \\ 0 & 0 & N_I(\mathbf{x}) \end{bmatrix} \mathbf{d}_I \tag{11}$$

where N_n is the number of total nodes of problem domain, $N_I(\mathbf{x})$ is the shape function at node I, $\mathbf{d}_I = \begin{bmatrix} w_I & \beta_{xI} & \beta_{yI} \end{bmatrix}^T$ is the nodal displacement vector associated to node I.

Then the bending and shear strains can be expressed as

$$\boldsymbol{\kappa} = \sum_I \mathbf{B}_{bI} \mathbf{d}_I$$
$$\boldsymbol{\gamma} = \sum_I \mathbf{B}_{sI} \mathbf{d}_I \tag{11}$$

where

$$\mathbf{B}_{bI} = \begin{bmatrix} 0 & \partial N_I / \partial x & 0 \\ 0 & 0 & \partial N_I / \partial y \\ 0 & \partial N_I / \partial y & \partial N_I / \partial x \end{bmatrix} \tag{12}$$

$$\mathbf{B}_{sI} = \begin{bmatrix} \partial N_I / \partial x & N_I & 0 \\ \partial N_I / \partial y & 0 & N_I \end{bmatrix} \tag{13}$$

The discretized system stiffness matrix, $\mathbf{K}$ can be expressed in terms of its bending, $\mathbf{K}_b$,and transverse shear, $\mathbf{K}_s$,components as

$$\mathbf{K} = \mathbf{K}_b + \mathbf{K}_s = \int_\Omega \mathbf{B}_b^T \mathbf{D}_b \mathbf{B}_b d\Omega + \int_\Omega \mathbf{B}_s^T \mathbf{D}_s \mathbf{B}_s d\Omega$$
$$= \sum_{i=1}^{N_e} \int_{\Omega_i^e} \mathbf{B}_b^T \mathbf{D}_b \mathbf{B}_b d\Omega + \sum_{i=1}^{N_e} \int_{\Omega_i^e} \mathbf{B}_b^T \mathbf{D}_b \mathbf{B}_b d\Omega \tag{14}$$

For static analysis, the discretized system equations of the Reissner-Mindlin plate can be expressed as

$$\mathbf{Kd} = \mathbf{F} \tag{15}$$

where $\mathbf{F}$ is the load vector and has the form of

$$\mathbf{F} = \int_\Omega \mathbf{N}^T p d\Omega + f_b = \sum_{i=1}^{N_e} \int_{\Omega_i^e} \mathbf{N}^T p d\Omega + f_b \tag{16}$$

in which f_b relates to the prescribed boundary loads.

For the free vibration, the force form vanishes and we shall have

$$\left(\mathbf{K} - \omega^2 \mathbf{M} \right) \mathbf{d} = \mathbf{0} \tag{17}$$

where ω is the natural frequency of the free vibration and $\mathbf{M}$ is the global mass matrix

$$\mathbf{M} = \int_\Omega \mathbf{N}^T \mathbf{m} \mathbf{N} d\Omega = \sum_{i=1}^{N_e} \int_{\Omega_i^e} \mathbf{N}^T \mathbf{m} \mathbf{N} d\Omega \tag{18}$$

Formulation of the MIN3

The main assumption of MIN3 is that the rotations are linear through the rotational DOFs at three nodes of the elements and deflection is quadratic through the deflection DOFs at six nodes (three nodes of the elements and three mid-edge points). The deflection DOFs at three mid-edge points can be removed by enforcing continuous shear constraints at every element edge, and then the deflection is approximated only by vertex DOFs at three nodes of the elements. Numerical examples demonstrated that the MIN3 element can overcome shear-locking-free and produces convergent solutions [Tessler and Hughes (1985)].

As shown in Figure.2, using the three-node triangular element mesh, the linear rotations β_x and β_y can be expressed as

$$\beta_x = \sum_{I=1}^{3} N_I(\mathbf{x})\beta_{xI} = \mathbf{N}\boldsymbol{\beta}_x$$

$$\beta_y = \sum_{I=1}^{3} N_I(\mathbf{x})\beta_{yI} = \mathbf{N}\boldsymbol{\beta}_y$$

(19)
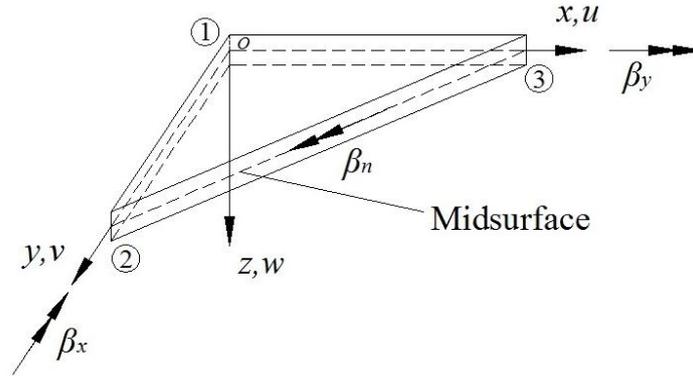
Figure. 2 Three-node triangular element

And the initial quadratic deflection w can be expressed as

$$w = \sum_{I=1}^{6} R_I w_I = \mathbf{R}w_{ini}$$

(20)

where $\mathbf{N} = \begin{bmatrix} N_1(x) & N_2(x) & N_3(x) \end{bmatrix}$ are the linear shape functions at node I. $\boldsymbol{\beta}_x^T = \begin{bmatrix} \beta_{x1} & \beta_{x2} & \beta_{x3} \end{bmatrix}$ and $\boldsymbol{\beta}_y^T = \begin{bmatrix} \beta_{y1} & \beta_{y2} & \beta_{y3} \end{bmatrix}$ are the rotational DOFs at three nodes of the element; $\mathbf{w}_{ini}^T = \begin{bmatrix} w_1 & w_2 & w_3 & w_4 & w_5 & w_6 \end{bmatrix}$ is the deflection DOFs at six nodes (three nodes of the elements and three mid-edge points as shown in Table. 1), and R is the row vector of quadratic shape functions given by

$$R_i = N_i(2N_i - 1), \; R_{i+3} = 4N_iN_k \; (i = 1, 2, 3; k = 2, 3, 1)$$

(21)

Table 1 Nodal configuration for initial (unconstrained) and constrained displacement

Shape functions		Initial nodal configuration	Continuous shear edge constraints $(w_{,s} + \beta_n)_{,s}\big|_{edges} = 0$	Constrained nodal configuration
w	β_x, β_y			
Quadratic	Linear		Three edge constraints	

Equations (19) and (20) can be directly used in formulating element matrices. However, it may be advantageous from the standpoint of nodal simplicity to condense out the mid-edge deflection

DOFs, w_4, w_5 and w_6 in w. This can be accomplished by enforcing continuous shear constraints at every element edge as given by the following differential relation

$$(w_{,s} + \beta_n)_{,s}\big|_{\text{edges}} = 0 \tag{22}$$

where s denotes the edge coordinate and β_n is the tangential edge rotation as shown in Figure.2. The enforcement of constraint (21) at three element edges yields

$$w_{i+3} = \frac{1}{2}(w_i + w_j) + \frac{1}{8}\Big[b_k(\beta_{xi} - \beta_{xj}) + a_k(\beta_{yj} - \beta_{yi})\Big] \tag{23}$$
$$(i = 1,2,3; j = 2,3,1; k = 3,1,2)$$

where $a_1 = x_3 - x_2, a_2 = x_1 - x_3, a_3 = x_2 - x_1, b_1 = y_2 - y_3, b_2 = y_3 - y_1, b_3 = y_1 - y_2$ as shown in Figure.3.

By substituting (23) into (20), there results a constrained deflection field exclusively in terms of vertex DOFs.

$$w = \sum_{I=1}^{3} N_I w_I + \sum_{I=1}^{3} H_I \beta_{xI} + \sum_{I=1}^{3} L_I \beta_{yI} = \mathbf{N}\mathbf{w} + \mathbf{L}\boldsymbol{\beta}_x + \mathbf{H}\boldsymbol{\beta}_y \tag{24}$$

where $\mathbf{w}^T = \begin{bmatrix} w_1 & w_2 & w_3 \end{bmatrix}, \mathbf{H} = \begin{bmatrix} H_1 & H_2 & H_3 \end{bmatrix}, \mathbf{L} = \begin{bmatrix} L_1 & L_2 & L_3 \end{bmatrix}$ are the vectors of shape functions, with, $I = 1,2,3$, given by

$$H_1 = \frac{1}{2}(a_2 N_3 N_1 - a_3 N_1 N_2)$$
$$H_2 = \frac{1}{2}(a_3 N_1 N_2 - a_1 N_2 N_3)$$
$$H_3 = \frac{1}{2}(a_1 N_2 N_3 - a_2 N_3 N_1)$$
$$L_1 = \frac{1}{2}(b_3 N_1 N_2 - b_2 N_3 N_1) \tag{25}$$
$$L_2 = \frac{1}{2}(b_1 N_2 N_3 - b_3 N_1 N_2)$$
$$L_3 = \frac{1}{2}(b_2 N_3 N_1 - b_1 N_2 N_3)$$

Then the element stiffness matrix can be finally obtained and written in the following form:

$$\mathbf{K}_e^{\text{MIN3}} = \int_{\Omega_e} \mathbf{B}_b^T \mathbf{D}_b \mathbf{B}_b d\Omega + \int_{\Omega_e} \mathbf{B}_s^T \mathbf{D}_s \mathbf{B}_s d\Omega \tag{26}$$

where

$$\mathbf{B}_b = \begin{bmatrix} 0 & \partial N_I/\partial x & 0 \\ 0 & 0 & \partial N_I/\partial y \\ 0 & \partial N_I/\partial y & \partial N_I/\partial x \end{bmatrix} \tag{27}$$

$$\mathbf{B}_s = \begin{bmatrix} \partial N_I/\partial x & \partial L_I/\partial x & \partial H_I/\partial x + N_I \\ \partial N_I/\partial y & \partial L_I/\partial x + N_I & \partial H_I/\partial y \end{bmatrix} \tag{28}$$

$a_1 = x_3 - x_2 \quad a_2 = x_1 - x_3 \quad a_3 = x_2 - x_1$
$b_1 = y_2 - y_3 \quad b_2 = y_3 - y_1 \quad b_3 = y_1 - y_2$
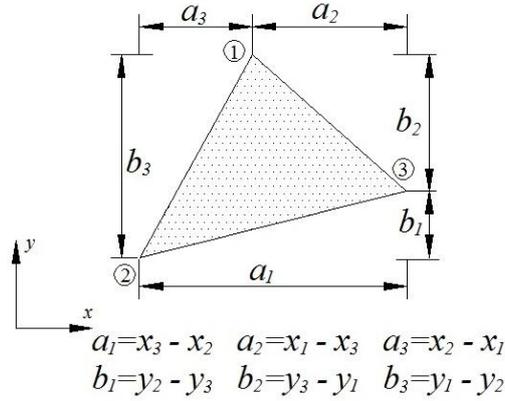
Figure.3 Three-node triangular element coordinate description

In order to further improve the accuracy of approximate solutions and to stabilize shear force oscillations. It was suggested that $\mathbf{D}_s$ in (8) should be replaced by $\hat{\mathbf{D}}_s$

$$\hat{\mathbf{D}}_s = \frac{Gkt}{t^2 + \alpha h_e^2} \tag{29}$$

In which h_e is the longest length of the edges of the element and α is a positive constant [Lyly and Stenberg (1993)].

Formulation of ES-MIN3

In this section, a new triangular element named an edge-based smoothed triangular element is established by combining the edge-based strain smoothing technique with the MIN3 (ES-MIN3). In this work, we incorporate the ES-FEM with the MIN3 to give a so-called ES-MIN3 for the plate elements. In the ES-MIN3, we do not use the compatible strain fields as in (11) but the smoothed strain fields over local smoothing domains associated with the edges of elements. Naturally the numerical integrations in (14) for the stiffness matrix are no longer based on elements as in standard FEM but on the edge-based smoothing domain Ω_k ($k = 1, 2, \cdots, N$), where N is the total number of edges in the 2D problem domains, for triangular elements, the smoothing domain for edge k is created by sequentially connecting two end points of the edge and centroids of its surrounding elements. As shown in Figure.4, for interior edges, the smoothing domain Ω_k for edge k is formed by assembling two sub-domains of two neighboring elements; while for global boundary edge, the smoothing domain Ω_k of edge k is a single sub-domain, in this case, the strain and stain matrix can be calculated as same as those in FEM.

In the present method, smoothing operation is applied over each smoothing domain, so the smoothed bending strain $\bar{\kappa}$ and smoothed shear stain $\bar{\gamma}$ can be calculated by

$$\overline{\kappa}_e = \int_{\Omega_e} \kappa W(x) d\Omega$$

$$\overline{\gamma}_e = \int_{\Omega_e} \gamma W(x) d\Omega$$
(30)

where $W(x)$ is a given smoothing function that satisfies at least unity property

$$\int_{\Omega_k} W_k(\mathbf{x}) d\Omega = 1$$
(31)
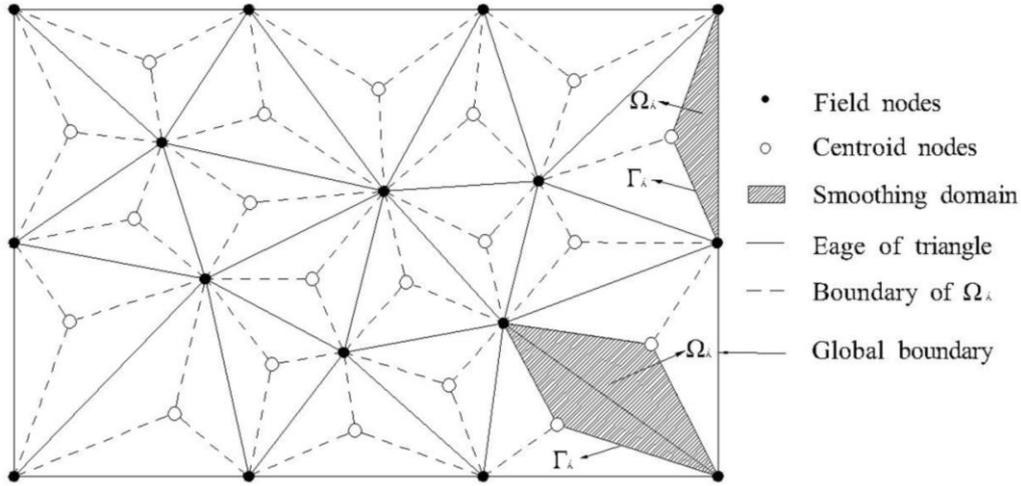
Figure. 4. Edge-based smoothing domains in 2D problem created by sequentially connecting the centroids of the adjacent triangles with the end-points of the edge.

In this study, the following simplest form of the smoothing function is used

$$W_k(\mathbf{x}) = \begin{cases} 1/A_k & x \in \Omega_k \\ 0 & x \notin \Omega_k \end{cases}$$
(32)

where A_k is the area of the smoothing domain of the kth edge and is computed by

$$A_k = \int_{\Omega_k} d\Omega = \frac{1}{3} \sum_{j=1}^{n_k^e} A_j^e$$
(33)

where n_k^e is the number of elements around the edge k ($n_k^e = 1$ for the boundary edges and $n_k^e = 2$ for inner edges, as shown in Figure. 4), A_j^e is the area of the jth element around the edge k.

By using the edge-based strain smoothing operation, the smoothed strain of the smoothed strain of the smoothing domain Ω_k^s in (30) can be expressed as follows

$$\overline{\kappa}_k = \frac{1}{A_k^s} \int_{\Omega_k^s} \kappa d\Omega = \frac{1}{A_k^s} \sum_{q=1}^{n_s} A_{k,q} \cdot \kappa_{k,q}$$

$$\overline{\gamma}_k = \frac{1}{A_k^s} \int_{\Omega_k^s} \gamma d\Omega = \frac{1}{A_k^s} \sum_{q=1}^{n_s} A_{k,q} \cdot \gamma_{k,q}$$
(34)

where $A_{k,q}$ denotes the area of the sub-smoothing domain associated with inner edge k, $\kappa_{k,q}$ and $\gamma_{k,q}$ are bending strain and shear strain of the qth sub-smoothing domain, respectively. With the above formulation, the smoothed strains for the smoothing domain of edge k can be expressed in the following forms:

$$\bar{\kappa}_k = \sum_{I=1}^{M_k} \bar{\mathbf{B}}_{b,I}(x_k) \cdot d_I$$

$$\bar{\gamma}_k = \sum_{I=1}^{M_k} \bar{\mathbf{B}}_{s,I}(x_k) \cdot d_I$$

(35)

where M_k is the total number of nodes in the influence domain of edge k, $\bar{B}_{b,I}(x_k)$ and $\bar{B}_{s,I}(x_k)$ are termed as the smoothed strain matrix that can be calculated as

$$\bar{\mathbf{B}}_{b,I}(x_k) = \frac{1}{A_k} \sum_{i=1}^{M_k} \frac{1}{3} A_i \mathbf{B}_{bi}(x_k)$$

$$\bar{\mathbf{B}}_{s,I}(x_k) = \frac{1}{A_k} \sum_{i=1}^{M_k} \frac{1}{3} A_i \mathbf{B}_{si}(x_k)$$

(36)

Therefore the global stiffness matrices of the ES-MIN3 element can be assembled by

$$\bar{\mathbf{K}}_k = \sum_{i=1}^{n_k} \mathbf{K}_k$$

(37)

where K_k is the smoothed element stiffness matrix given by

$$\bar{\mathbf{K}}_k = \int_{\Omega_k} \bar{\mathbf{B}}_b^T \mathbf{D}_b \bar{\mathbf{B}}_b d\Omega + \int_{\Omega_k} \bar{\mathbf{B}}_s^T \mathbf{D}_s \bar{\mathbf{B}}_s d\Omega = \bar{\mathbf{B}}_b^T \mathbf{D}_b \bar{\mathbf{B}}_b A_k + \bar{\mathbf{B}}_s^T \mathbf{D}_s \bar{\mathbf{B}}_s A_k$$

(38)

The procedure of assembling the global stiffness matrix in the ES-MIN3 is exactly the same as the practice in the standard FEM. It can be easily seen from (37) that the resultant linear system is symmetric and banded (due to the compact supports of FEM shape functions), which implies that the system equation can be solved efficiently.

Numerical results

Static analysis

Consider a flexible rectangular plate (0.314m × 0.414m) which is made of aluminum (ρ =2700kg/m^3, v =0.3, and E =71GPa). The thickness of the plate is 0.001m. The plate is subjected to a uniform load of $q(x,y)$=1Pa, and is given for clamp boundary condition. Uniform meshes of $2 \times N \times N$ three-node triangular plate elements shown in Figure.5 is used in the computation, where N denotes the number of elements per edge.
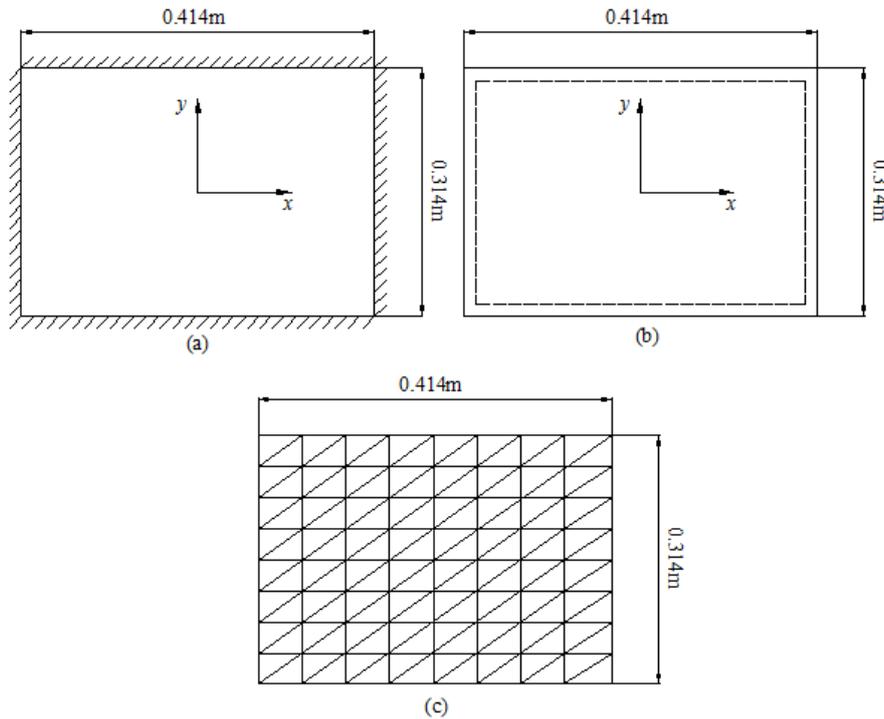
Figure 5 Two rectangular plate models and the representative meshes: (a) clamped plate; (b) simply supported plate; (c) regular mesh using three-node triangular elements

For static analysis, the deflection at the center point of the plate is computed; the result is plotted against the mesh density in terms of number of elements per edge N, as shown in Figure.6. It is seen that the ES-MIN3 achieves the higher accuracy compared to the DSG and MIN3 elements.
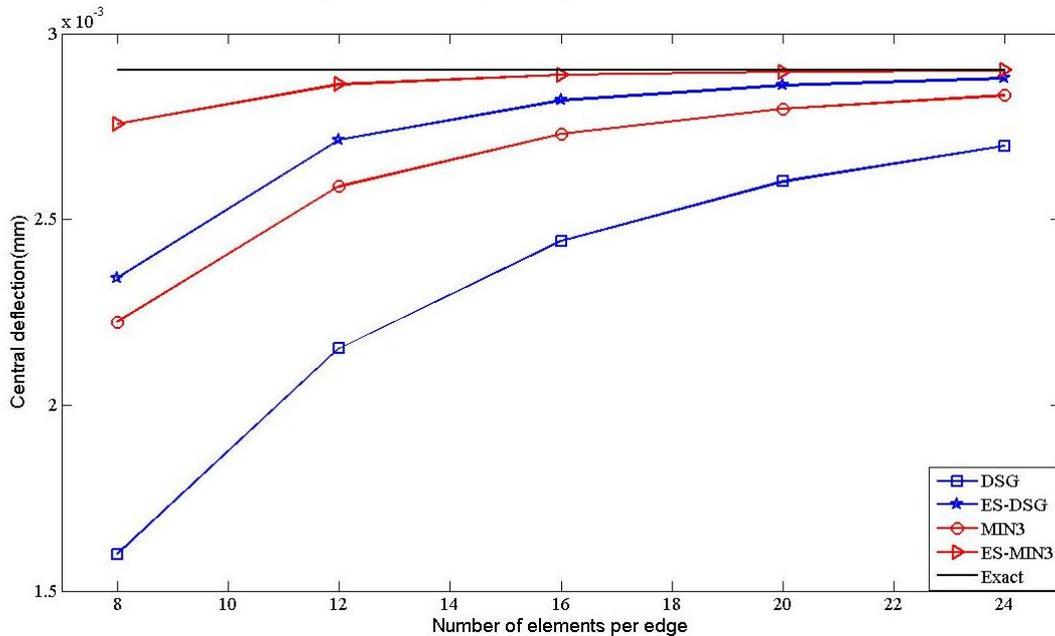
Figure 6 Convergence of deflection of the plate at the center against the mesh density

Free vibration analysis of plates

In this section, we investigate the performance of the ES-MIN3 used for computing the natural

frequencies of plates. The geometry and material parameters of the plate are the same as in last section. All the edges of the plate are simply-supported and five uniform meshes of $2 \times N \times N$ three-node triangular plate elements with N=8, 12, 16, 20, 24 are used in the computation. The first six natural frequencies of the plate obtained from ES-MIN3 are listed in Table 2, for comparison, the analytical solutions and some other numerical results are also listed in the table. As indicated in the table, all the numerical results are in good agreement with the analytical results in low frequency range. The errors of the results for all these numerical methods become larger with the increasing of mode order. However, the results obtained using ES-MIN3 is much more accurate and converged much faster than those obtained using other methods. It is confirmed that the ES-MIN3 is efficient and can give high accurate solutions in free vibration analysis. In particular, the ES-MIN3 can achieve accurately the values of high frequencies of plate by using only coarse meshes.

Table 2 Convergence of the first six natural frequencies (Hz) of the plate

Meshing	Methods	Mode sequence number					
		1	2	3	4	5	6
8	DSG	43.87	99.74	138.23	190.09	243.67	313.86
	ES-DSG	39.93	88.57	122.37	172.39	191.12	277.93
	MIN3	41.95	96.03	130.66	186.66	220.30	301.90
	ES-MIN3	39.86	87.09	120.98	171.40	178.50	271.09
12	DSG	41.71	92.82	125.83	173.07	204.74	272.24
	ES-DSG	39.32	84.41	116.79	161.88	168.01	251.05
	MIN3	40.27	88.14	120.48	168.61	182.61	263.15
	ES-MIN3	39.33	83.88	116.36	161.27	163.64	244.26
16	DSG	40.63	88.97	120.61	165.99	186.77	257.54
	ES-DSG	39.13	83.05	115.06	158.00	161.48	239.85
	MIN3	39.67	85.25	117.14	162.07	169.95	251.07
	ES-MIN3	39.15	82.84	114.88	157.55	159.77	236.19
20	DSG	40.04	86.62	117.94	162.08	176.45	250.25
	ES-DSG	39.05	82.47	114.31	156.07	159.02	234.67
	MIN3	39.40	83.89	115.63	158.91	164.29	243.79
	ES-MIN3	39.07	82.38	114.21	155.77	158.21	232.69
24	DSG	39.69	85.15	116.42	159.65	170.07	246.09
	ES-DSG	39.00	82.17	113.91	155.00	157.85	232.00
	MIN3	39.25	83.15	114.82	157.10	161.34	238.58
	ES-MIN3	39.02	82.13	113.86	154.81	157.42	230.83
	Analytical	38.95	81.61	113.11	152.72	155.78	226.89

Conclusions

In this work, the edge-based smoothed finite element method is combined with the well-known MIN3 to give a so-called the ES-MIN3 for static and free vibration analyses of plates. The smoothed Galerkin weak form is adopted to formulate the discretized system equations. The numerical integration is performed over the smoothing domains associated with edges of mesh. Through the formulation and the numerical examples, some concluding remarks can be drawn as follows:

1) The ES-MIN3 is straightforward and the implementation is as easy as MIN3 for the static and free vibration analyses of plates.

2) The shear locking of the triangular plate elements has been successfully alleviated with ES-MIN3 and the ES-MIN3 elements have only three DOFs at each vertex node without additional degrees of freedom, in addition, the ES-MIN3 only use the triangular elements which is a clear adbantage compared to quadrilateral elements when the geometry domain of plate is complicated.

3) For both static and free vibration analyses, the results of the ES-MIN3 agree well with other methods. The ES-MIN3 gives much more accurate results than the DSG, MIN3 and is a good competitor to the ES-DSG.

4) The ES-MIN3 works very well with triangular meshes and it is thus very promising to solve real engineering problems which usually are of complicated geometries with very accurate results.

References

Henry, T. Y., Saigal. S., Masud, A., Kapania, R. K. (2000) A survey of recent shell finite elements, *International Journal for Numerical Methods in Engineering* **47**, 101–127.

Reddy, J. N, (2006) Theory and analysis of elastic plates and shells.CRC Press, Taylor and Francis Group, New York

Lee, S. W., Wong, C. (1982) Mixed formulation finite elements for Mindlin theory platebending, *International Journal for Numerical Methods in Engineering* **18**, 1297–1311.

Zienkiewicz, O. C., Lefebvre, D. (1988) A robust triangular plate bending element of the Reissner–Mindlin type, *International Journal for Numerical Methods in Engineering* **26**, 1169–1184.

Miranda, S. D., Ubertini, F.(2006) A simple hybrid stress element for shear deformable plates, *International Journal for Numerical Methods in Engineering* **65**, 808–833.

Auricchio, F., Taylor, R. L. (1994) A shear deformable plate element with an exact thin limit, *Computer Methods in Applied Mechanics and Engineering* **65**, 393–412.

Lovadina, C. (1998) Analysis of a mixed finite element method for the Reissner–Mindlin plate problems, *Computer Methods in Applied Mechanics and Engineering* **163**, 71–85.

Simo, J. C., Rifai, M. S. (1990) A class of mixed assumed strain methods and the method of incompatible modes, *International Journal for Numerical Methods in Engineering* **29**, 1595–1638.

Simo, J. C., Fox, D.D., Rifai, M. S. (1989) On a stress resultant geometrically exact shell model. Part II: The linear theory, computational aspects, *Computer Methods in Applied Mechanics and Engineering* **73**, 53–92.

Tessler, A., Hughes, T. J. R. (1985) A three-node Mindlin plate element with improved transverse shear, *Computer Methods in Applied Mechanics and Engineering* **50**, 71–101.

Bathe, K. J., Dvorkin, E. N. (1985) A four-node plate bending element based on Mindlin/Reissener plate theory and a mixed interpolation, *International Journal for Numerical Methods in Engineering* **21**, 367–383.

Batoz, J. L., Lardeur, P. (1989) A discrete shear triangular nine d.o.f. element for the analysis of thick to very thin plates, *International Journal for Numerical Methods in Engineering* **29**, 533–560.

Bletzinger, K. U., Bischoff, M., Ramm, E. (2000) A unified approach for shear-locking free triangular and rectangular shell finite elements, *Computers and Structures* **75**, 321–334.

Chen, J. S., Wu, C. T., Yoon, S. You, Y. (2001) A stabilized conforming nodal integration forGalerkin mesh-free methods, *International Journal for Numerical Methods in Engineering* **50**, 435–466.

Liu, G. R., Dai, K. Y. and Nguyen, T. T. (2007) A smoothed finite element method for mechanics problems, *Computational Mechanics* **39**, 859–877.

Nguyen, T. T., Rabczuk, T., Nguyen, X. H., Bordas, S. (2008) A smoothed finite element method for shell analysis, *Computer Methods in Applied Mechanics and Engineering* **198**, 165–177.

Nguyen, T. T., Phung, V. P., Thai, C. H., Nguyen, X. H. (2012) A cell-based smoothed discrete shear gap method using triangular elements for static and free vibration analyses of Reissner–Mindlin plates, *International Journal for Numerical Methods in Engineering* **91**, 705–741.

Nguyen, T. T., Phung, V. P., Thai, C. H., Nguyen, X. H. (2013a) A cell-based smoothed discrete shear gap method (CS-DSG3) using triangular elements for static and free vibration analyses of shell structures, *International Journal of Mechanical Sciences* **74**, 32–45.

Nguyen, T. T., Phung, V. P., H. Luong, V. H., Nguyen, V. H., Nguyen, X. H. (2013b) A cell-based smoothed three-node Mindlin plate element (CS-MIN3) for static and free vibration analyses of plates, *Computational Mechanics*, **51**, 65–81.

Wu, C. T., Wang, H. P. (2013) An enhanced cell-based smoothed finite element method for the analysis of Reissner–Mindlin plate bending problems involving distorted mesh, *International Journal for Numerical Methods in Engineering* **95**, 288–312.

Liu, G. R., Nguyen, T. T., and Zeng, K. (2009a). A novel singular node based smoothed finite element method (NS-FEM) for upper bound solutions of fracture problems, *International Journal for Numerical Methods in Engineering,* **83**, 1466–1497.

Liu, G. R., Nguyen, T. T., Nguyen, X. H. and Lam, K. Y. (2009b) A node-based smoothed finite element method (NS-FEM) for upper bound solutions to solid mechanics problems, *Computers and Structures* **87**, 14–26.

Nguyen, T. T., Liu, G. R., Nguyen, X. H. and Nguyen, X. C. (2011) Adaptive analysis using the node-based smoothed finite element method (NS-FEM) , *International Journal for Numerical Methods in Engineering,* **27**, 198–218.

Liu, G. R., Nguyen, T. T., and Lam, K. Y. (2009c) An edge-based smoothed finite element method (ES-FEM) for static, free and forced vibration analysis of solids, *Journal of Sound and Vibration*, **320**, 1100–1130.

Nguyen, X. H., Liu, G. R., Thai, H. C., and Nguyen, T. T. (2009) An edge based smoothed finite element method with stabilized discrete shear gap technique for analysis of Reissner–Mindlin plates, *Computer Methods in Applied Mechanics and Engineering* **199**, 471–489.

Li, W., Chai, Y. B., Lei, M. and Liu, G. R. (2013) Analysis of Coupled Structural-Acoustic Problems Based on the Smoothed Finite Element Method (S-FEM), 4th International Conference on Computational Methods, Golded Coast, Australia.

Li, W., Lei, M., Chai, Y. B. and Liu, G. R. (2013) Analysis of Exterior Acoustics using the Edge-based Smoothed Finite Element Method (ES-FEM), APCOM & ISCM 2013，Singapore.

Feng, S. Z., Cui, X. Y., Liu, G. R. (2013) Transient thermal mechanical analyses using a face-based smoothed finite element method (FS-FEM), *International Journal of Thermal Sciences*, **74**, 95–103.

Lyly, M., Stenberg, R.,Vihinen, T. (1993) A stable bilinear element for the Reissner–Mindlin plate model, *Computer Methods in Applied Mechanics and Engineering* **110**, 343–357.

Computational methods on tool forces in friction stir welding

*Z. Zhang, Q. Wu, and H.W. Zhang

State Key Laboratory of Structural Analysis for Industrial Equipment, Department of Engineering Mechanics, Faculty of Vehicle Engineering and Mechanics, Dalian University of Technology, Dalian 116024, China

*Corresponding author: zhangz@dlut.edu.cn

Abstract

The commercial applications of friction stir welding (FSW) to hard materials are limited by tool breakages. But the tool forces and the relations to tool geometries remain unknown. So, here we established a model on calculation of tool forces in FSW and examined how the tool design affects the temperature fields and the tool forces along the welding direction. Results indicate that temperature rises are very important for the tool forces in FSW. Both the increase of the shoulder size and the increase of the rotating speed can lead to the increase of the welding temperatures in FSW and then decrease the tool forces in the welding direction. Larger shoulder or higher rotating speed can increase the tool life.

Keywords: Friction stir welding, Finite element method, Tool force, Temperature

Introduction

Friction stir welding (FSW) has been invented for more than 15 years. Due to its solid joining nature, FSW has many advantages over the traditional fusion welding techniques, such as low distortion, low welding defects, fine grains in welding zone, etc., which makes it being successfully applied to aerospace, automobile, ship industries, etc. In FSW, a rotating tool is inserted into the butt of two welding plates and then translates along the welding line [Thomas et al. (1991); Mishra and Ma (2005)]. Based on the principles for FSW, friction stir processing (FSP) was developed by [Berbon et al. (2001)] as a genetic tool for microstructural modifications. FSW has been used for the joining of aluminum alloys [Ahmed et al. (2008); Altenkirch et al. (2008); Nielsen (2008); Fonda et al. (2008); Cabibbo et al. (2007)], magnesium alloys [Afrin et al. (2008) ; Gharacheh et al. (2006); Park et al. (2003)], stainless steels [Reynolds et al. (2003); Saeid et al. (2008)], titanium alloys [Mironov et al. (2008); Lee et al. (2005)], copper alloys [Park et al. (2004)], composite materials [Feng et al. (2008); Fernandez et al. (2004)] and even the joining of dissimilar materials [Kwon et al. (2008); Ouyang et al. (2006); Cavaliere et al. (2009)]. During the welding process, welding tool is believed to be the key component for a successful FSW [Elangovan et al. (2008); Zhang et al. (2009); Kumar and Kailas (2008)], especially for FSW of strong material [Bhadeshia and DebRoy (2009)]. Although the tool force in welding direction can be measured in experiments [Yan et al. (2005)], the theories for determination of the tool forces in FSW should be needed for the development of reliable, lasting and cost effective welding tools and even for the optimization of welding tools with lower costs. So, here we presented a method for calculation of tool forces. Temperature rises are believed to be one of the key factors to affect the plasticized material flow near the welding tool [Zhang and Zhang (2009); Nandan et al. (2007)]. So, the calculated tool forces, the temperature rises and the different tools are considered together for the examination on how the tool design affects the temperature fields and the tool forces along the welding direction.

Model description

ABAQUS was used with the combination of the user subroutine which was compiled by FORTRAN code for the description of a modified coulomb friction law [Zhang (2008)]. Eight node

thermo-mechanical brick elements are used for the mesh generation of the workpiece. For convenience of mesh generation with brick elements, a circular workpiece with the radii of 30mm is considered. The model has been validated for the temperature and material flow during FSW of AA6061 [Zhang and Zhang (2008); Zhang et al. (2011); Zhang and Zhang (2007); Zhang and Zhang (2009)]. The applied axial pressure is selected to be 70MPa, the pre-heating time 1.5s and the traverse speed 140mm/min for the current computations. The inflow temperature is set to be the room temperature (25°C). The boundary of the welding plate is treated as Eulerian type, on which the motion of the material points can be independent of the meshes. Arbitrary Lagrangian— Eulerian (ALE) method [Belytschko et al. (2000)] is combined with the adaptive meshing to avoid excessive element distortions. Four different tool profiles are used for the numerical comparisons, as shown in Fig.1. For the tool with a shoulder diameter of 20mm, three rotating speeds, i.e. 500rpm, 550rpm and 600rpm are used to study the effect of rotating speed on tool force. For other cases, the rotating speed is set to be 500rpm.

(a) Tool I   (b) Tool II

(c) Tool III   (d) Tool IV
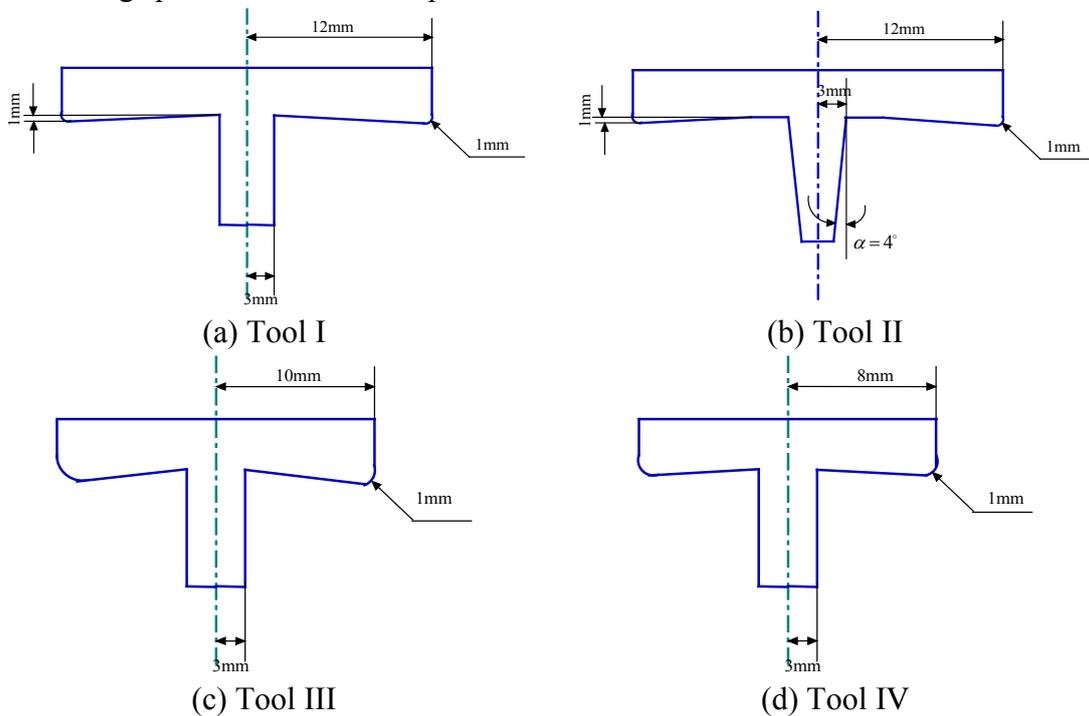
Figure 1 Schematics of welding tools

As shown in Fig.2, the tool forces in FSW can be determined by the hydraulic pressure and the frictional stresses on contact surfaces,

$$F_{Long} = \int_{S_1} p \sin\theta \, dS_1 + \int_{S_1} p_t \cos\theta \, dS_1 + \int_{S_2} p_t \, dS_2 \tag{1}$$

where p is the hydraulic pressure and p_t frictional stress. S_1 and S_2 are the pin-plate contact area and the shoulder-plate contact area, respectively.

(a) Pin

(b) Shoulder
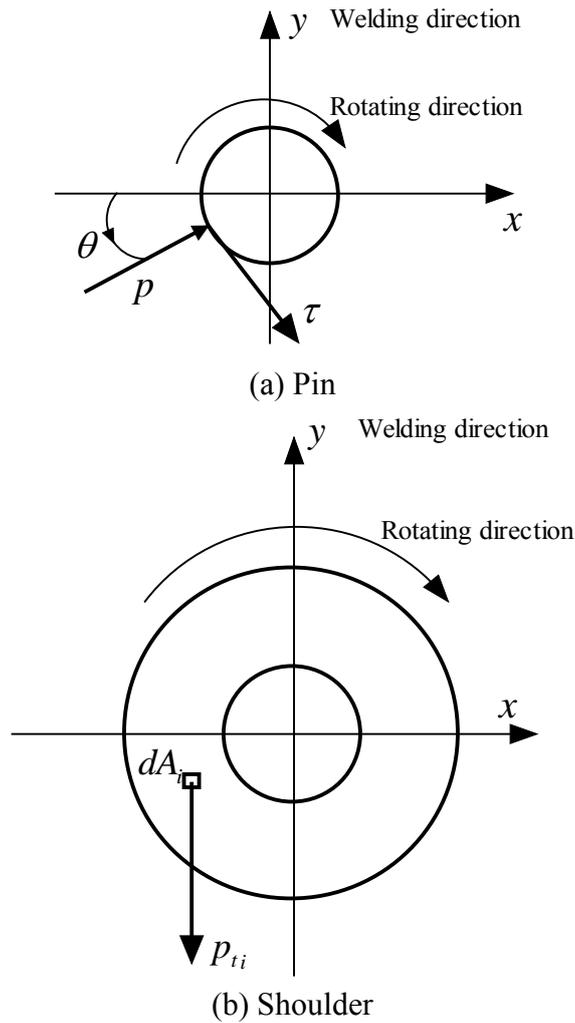
Figure 2 Calculation of tool forces in FSW

The hydraulic pressure can be obtained from the trace of the stress tensor in FSW,

$$p = -\frac{1}{3}\text{trace}(\sigma_{ij}) \tag{2}$$

where σ_{ij} is the Cauchy stress tensor which can be computed using the constitutive equation,

$$\sigma_{ij} = C_{ijkl}\varepsilon_{kl}^e \tag{3}$$

where C_{ijkl} is the elasticity tensor and ε_{kl}^e is the elastic part of the strain($\varepsilon_{kl} - \varepsilon_{kl}^p$). The total strain can be computed using the strain displacement equation,

$$\varepsilon_{ij} = \frac{1}{2}\left(u_{i,j} + u_{j,i}\right) \tag{4}$$

where u_i is the displacement. The boundary condition used for the inflow and outflow regions are u_i=0 for i=2, 3 and $u_{i,t}$=140mm/s for i=1.

The predictor-corrector method is used for the calculations of the plastic strain,

$$\dot{\varepsilon}^p = \dot{\lambda}\frac{3\sigma'}{2\bar{\sigma}} \tag{5}$$

where λ is the plastic rate parameter, σ' is the deviatoric stress, and $\bar{\sigma}$ is the von Mises effective stre4ss. The yield function can be defined as,

$$f = \bar{\sigma} - \sigma_s(T) = 0 \tag{6}$$

where σ_s is the yield stress which is the function of temperature.
The temperature is determined by solving the heat transfer equation

$$\rho c \frac{\partial T}{\partial t} - \frac{\partial}{\partial x}\left(k\frac{\partial T}{\partial x}\right) - \frac{\partial}{\partial y}\left(k\frac{\partial T}{\partial y}\right) - \frac{\partial}{\partial z}\left(k\frac{\partial T}{\partial z}\right) - \rho Q = 0 \tag{7}$$

$$k\frac{\partial T}{\partial x}n_x + k\frac{\partial T}{\partial y}n_y + k\frac{\partial T}{\partial z}n_z = q \tag{8}$$

where ρ is the density, c is the specific heat, k is the thermal conductivity, q is the heat flux on the contact area and n_x, n_y, n_z represent the directions. The temperature dependent values of c and k can be found in [An and Liu (1998)]. Q is the heat generated by the plastic deformations,

$$Q = \sigma_{ij}\dot{\varepsilon}_{ij}^p \tag{9}$$

where $\dot{\varepsilon}_{ij}^p$ is the strain rate.
The heat flux on the contact area q is,

$$q = \eta p_t \dot{\gamma} \tag{10}$$

where η is the fraction of frictional heat entering the workpiece (90% in current work), $\dot{\gamma}$ is the relative velocity between the tool and workpiece.
The general finite element form for the heat transfer equation can be obtained by the spatial discretization,

$$C\dot{T} + KT = P \tag{11}$$

where C is the heat capacity matrix, K the thermal conductivity matrix and P is the thermal load matrix which is determined by the mentioned internal heat source, the heat flux on contact surface and the boundary conditions. Explicit forward difference integration method is used to solve this equation.
The displacement required to compute total strain can be determined solving the equation of motion,

$$\sigma_{ij,j} + F_i = \rho u_{i,tt} \tag{12}$$

where F_i is the body force per unit volume and $u_{i,tt}$ is the acceleration.
The classical finite element form of the above equations can be obtained by spatial discretization,

$$M\ddot{u} = P' - P'_{int} \tag{13}$$

where M is the mass matrix, P' load matrix and P'_{int} internal nodal forces. Explicit central difference integration method is used for the computation of displacements and accelerations of nodes. The load matrix consists of the contact forces, normal (p_n) and tangent (p_t). The contact forces can be expressed in terms of displacements of the contact points by the penalty algorithm [Zhang et al. (2005)] with consideration of the shear failure criterion,

$$p_n = -E_n\left(du_n^s - du_n^m - \delta^*\right) \text{ for both sliding and sticking states} \tag{14}$$

$$p_t = -\min\left(\mu p_n sign(du_t^s - du_t^m), \sigma_s / \sqrt{3}\right) \text{ for sliding state} \tag{15}$$

where E_n is penalty factor for normal contact, which can be taken from 1 to 100 times of the element stiffness of the representative underlying welding material according to the overclosures in calculations. δ^* is the gap. du_n is the normal displacement. The superscripts m and s represent the mater (tool) and slave (welding plate) contact surfaces. The displacement for master is considered to be zero in this calculation.

Results and discussions

The computed temperature fields around the tool are shown in Fig. 3. The maximum temperature for Tool I in 500rpm and 140mm/min is $430\,^{\circ}\mathrm{C}$, as shown in Fig. 3(a). The experimental measured temperature is about $440\,^{\circ}\mathrm{C}$ under the same welding conditions and tool sizes [Chen and Kovacevic (2003)], which can verify the developed model for heat transfer. When conical pin is used, the maximum temperature is decreased slightly to $426\,^{\circ}\mathrm{C}$, as shown in Fig. 3(b). Compared with the variation of pin shape, the effect of shoulder size on temperature is more obvious. When the shoulder radius is changed to 10mm, the maximum temperature is decreased to $384\,^{\circ}\mathrm{C}$, as shown in Fig. 3(c). With the further decrease of the shoulder radius to 8mm, the maximum temperature can be further decreased to $344\,^{\circ}\mathrm{C}$, as shown in Fig. 3(d). The observation on effect of shoulder size on temperature is consistent to previous studies [Zhang et al. (2009)]. Usually, higher rotating speed is used for smaller shoulder. So, two new cases for Tool III (Fig. 1) with higher rotating speeds are adopted for comparisons. When the rotating speed is increased to 550rpm, the maximum temperature can be increased to $393\,^{\circ}\mathrm{C}$, as shown in Fig. 3(e). With the further increase for the rotating speed to 600rpm, the maximum temperature can be increased to $400\,^{\circ}\mathrm{C}$, as shown in Fig. 3(f).
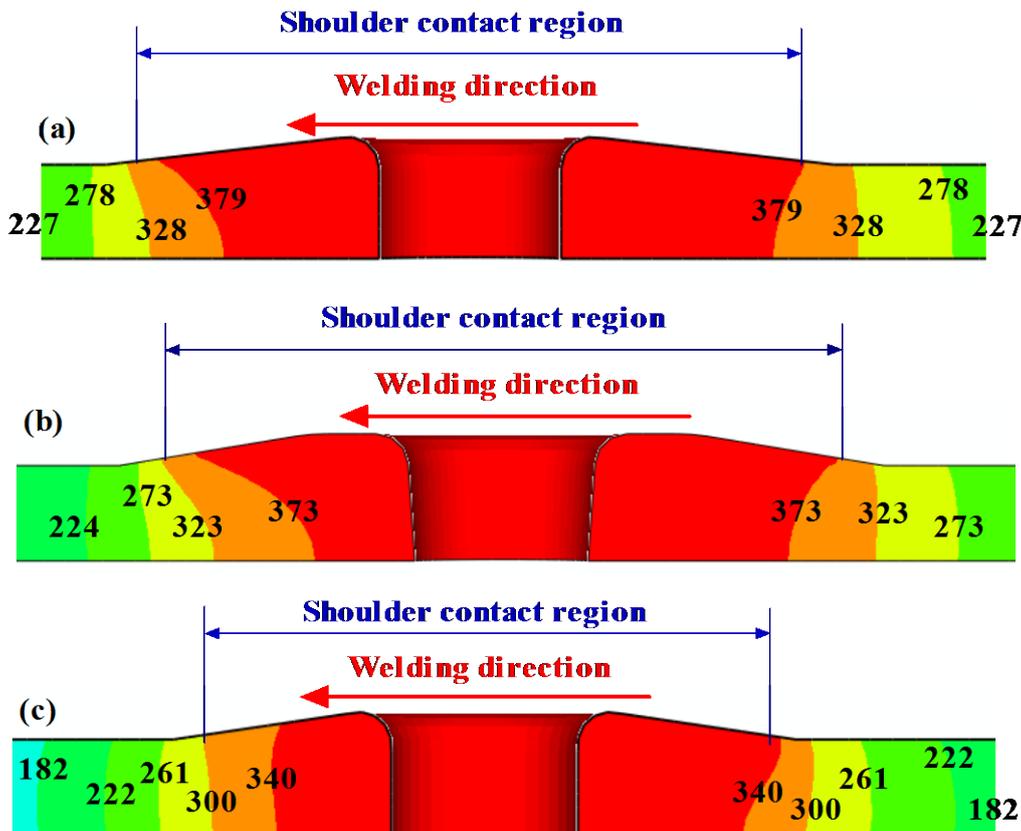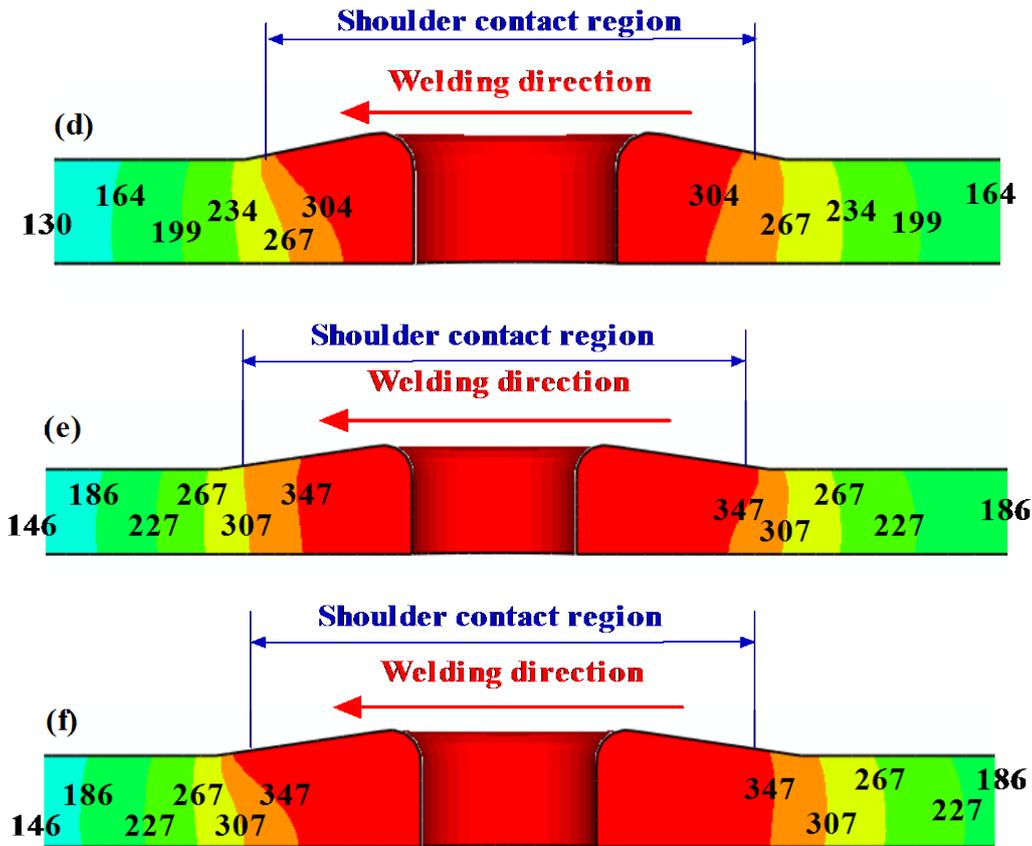
Figure 3 Temperatures in different cases: (a) Tool I in 500rpm ($T_{\max} = 430°C$); (b) Tool II in 500rpm ($T_{\max} = 426°C$); (c) Tool III in 500rpm ($T_{\max} = 384°C$); (d) Tool IV in 500rpm ($T_{\max} = 344°C$); (e) Tool III in 550rpm ($T_{\max} = 393°C$); (f) Tool III in 600rpm ($T_{\max} = 400°C$)

Temperature fields are important for the plasticization of the material near the welding tool. So, it can significantly affect the frictional force calculation and even the tool forces. Frictional stresses along selected paths A=>B=>C are shown in Fig.4 for different cases. The frictional stress on the selected path is very similar for Tool I and Tool II. When the conical pin is adopted, the friction stress on the contact surface is increased slightly. The friction stress can be increased due to the decrease of the shoulder diameter and the decrease of the rotating speed. With consideration of the temperature fields shown in Fig.3, the frictional stress can be increased with the decrease of temperature in FSW. Moreover, the frictional stress on shoulder-plate interface is lower than the one on the pin-plate interface. But with the increase of the temperature due to the increase of rotating speed or increase of shoulder size, the friction stresses on pin-plate and shoulder-plate interfaces become similar.
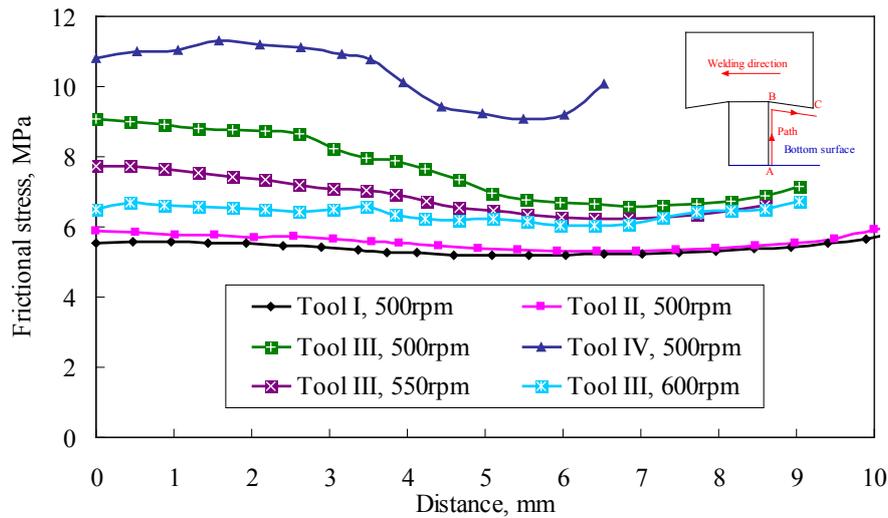
Figure 4 Comparison of frictional stress in different cases

The calculated forces for different cases are shown in Fig. 5. When Tool I with 500rpm is used, the calculated force in welding direction is 3.4kN. But when conical pin is adopted under the same welding conditions, the tool force in welding direction is increased to 4kN due to the decrease of temperature. When smaller shoulders are used for Tool III and Tool IV, the tool forces in welding direction is obviously increased due to the obvious decrease of temperatures. This is the reason that the larger shoulder can increase the tool life in manufacture. The experimentally measured transverse force is about 8kN under the rotating speed of 500rpm and the shoulder diameter of 20.3mm for FSW of AA2524 [Yan et al. (2005)]. In fact, AA6061 can be believed to be softer than AA2524 due to the smaller yield stress. So, it can be deduced that the transverse force for FSW of AA6061 should be smaller than the one in FSW of AA2524 under the similar welding conditions. For Tool III in 500rpm, the transverse force in FSW of AA6061 is 6.5kN. The comparison with Ref. [Yan et al. (2005)] shows that the computed transverse force in current model is reasonable. For smaller shoulder (Tool III), the increase of the rotating speed can decrease the tool force in welding direction apparently. This means that higher rotating speed should be used in manufacture for smaller shoulder in FSW, which can lead to more temperature rises. With the increase of the temperature, the material becomes softer and then the tool force in welding direction can be decreased.
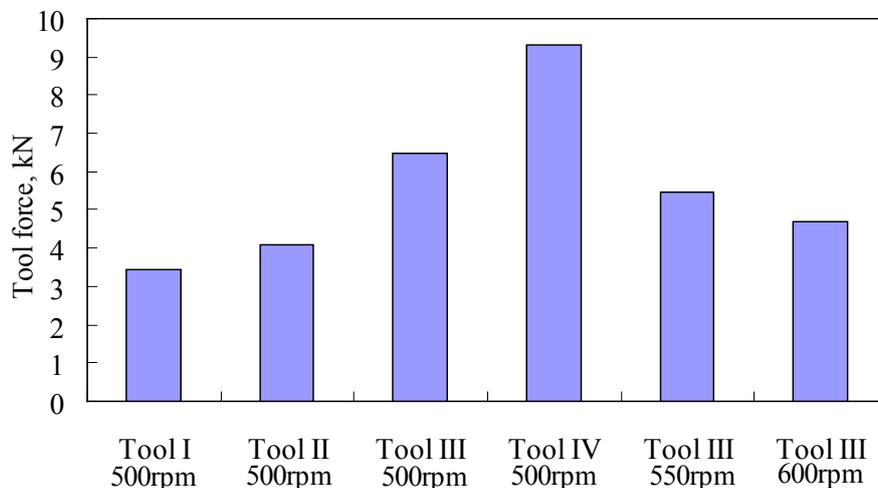
Figure 5 Tool forces in welding direction in different cases

Conclusions

1) Temperature rise is very important for the tool force in FSW. Higher temperature can lead to softer material near the welding tool and the decrease the tool forces in welding direction.
2) Both the increase of the shoulder size and the increase of the rotating speed can lead to the increase of the temperatures in FSW and then decrease the tool forces in welding direction.
3) Larger shoulder or higher rotating speed can increase the tool life.

Acknowledgement

References

W.M. Thomas, E.D. Nicholas, J.C. Needham, M.G. Murch, P. Templesmith, C.J. Dawes. Friction stir welding, International Patent Application No. PCT/GB92102203 and Great Britain Patent Application No. 9125978.8, 1991.

R.S. Mishra, Z.Y. Ma. Friction stir welding and processing, Materials Science and Engineering R, 2005, 50:1-78.

P.B. Berbon, W.H. Bingel, R.S. Mishra, C.C. B.M.W. Mahoney. Friction stir processing: a tool to homogenize nanocomposite aluminum alloys, Scripta Materialia, 2001, 44: 61-66.

M.M.Z. Ahmed, B.P. Wynne, W.M. Rainforth, P.L. Threadgill. Quantifying crystallographic texture in the probe-dominated region of thick-section friction-stir-welded aluminium, Scripta Materialia, 2008, 59: 507-510.

J. Altenkirch, A. Steuwer, M. Peel, D.G. Richards, P.J. Withers. The effect of tensioning and sectioning on residual stresses in aluminium AA7749 friction stir welds, Materials Science and Engineering A, 2008, 488: 16-24.

K.L. Nielsen. Ductile damage development in friction stir welded aluminum (AA2024) joints, Engineering Fracture Mechanics, 2008, 75: 2795-2811.

R.W. Fonda, K.E. Knipling, J.F. Bingert. Microstructural evolution ahead of the tool in aluminum friction stir welds, Scripta Materialia, 2008, 58: 343-348.

M. Cabibbo, H.J. McQueen, E. Evangelista, S. Spigarelli, M. Di Paola, A. Falchero. Microstructure and mechanical property studies of AA6056 friction stir welded plate, Materials Science and Engineering: A, 2007, 460-461: 86-94.

N. Afrin, D.L. Chen, X. Cao, M. Jahazi. Microstructure and tensile properties of friction stir welded AZ31B magnesium alloy, Materials Science and Engineering A, 2008, 47: 179-186.

M. Abbasi Gharacheh, A.H. Kokabi, G.H. Daneshi, B. Shalchi, R. Sarrafi. The influence of the ratio of "rotational speed/traverse speed" (ω/v) on mechanical properties of AZ31 friction stir welds, International Journal of Machine Tools and Manufacture, 2006, 46: 1983-1987.

S.H.C. Park, Y.S. Sato, H. Kokawa. Effect of micro-texture on fracture location in friction stir weld of Mg alloy AZ61 during tensile test, Scripta Materialia, 2003 49: 161-166.

A. P. Reynolds, W. Tang, T. Gnaupel-Herold, H. Prask. Structure, properties, and residual stress of 304L stainless steel friction stir welds, Scripta Materialia, 2003, 48: 1289-1294.

T. Saeid, A.A.H. Assadi, F.M. Ghaini. Effect of friction stir welding speed on the microstructure and mechanical properties of a duplex stainless steel, Materials Science and Engineering A, 2008, 496: 262-268.

S. Mironov, Y. Zhang, Y.S. Sato, H. Kokawa. Crystallography of transformed β microstructure in friction stir welded Ti–6Al–4V alloy, Scripta Materialia, 2008, 59: 511-514.

W.B. Lee, C.Y. Lee, W.S. Chang, Y.M. Yeon, S.B. Jung. Microstructural investigation of friction stir welded pure titanium, Materials Letters, 2005, 59: 3315-3318.

H.S. Park, T. Kimura, T. Murakami, Y. Nagano, K. Nakata, M. Ushio. Microstructures and mechanical properties of friction stir welds of 60% Cu–40% Zn copper alloy, Materials Science and Engineering A, 2004, 371: 160-169.

A.H. Feng, B.L. Xiao, Z.Y. Ma. Grain boundary misorientation and texture development in friction stir welded SiCp/Al–Cu–Mg composite, Materials Science and Engineering A, 2008, 497: 515-518.

G.J. Fernandez, L.E. Murr. Characterization of tool wear and weld optimization in the friction-stir welding of cast aluminum 359+20% SiC metal-matrix composite, Materials Characterization, 2004, 52: 65-75.

Y.J. Kwon, I. Shigematsu, N. Saito. Dissimilar friction stir welding between magnesium and aluminum alloys, Materials Letters, 2008, 62: 3827-3829.

J. Ouyang, E. Yarrapareddy, R. Kovacevic. Microstructural evolution in the friction stir welded 6061 aluminum alloy (T6-temper condition) to copper, Journal of Materials Processing Technology, 2006, 172: 110-122.

P. Cavaliere, A. De Santis, F. Panella, A. Squillace. Effect of welding parameters on mechanical and microstructural properties of dissimilar AA6082–AA2024 joints produced by friction stir welding, Materials & Design, 2009, 30: 609-616.

K. Elangovan, V. Balasubramanian. Influences of tool pin profile and welding speed on the formation of friction stir processing zone in AA2219 aluminium alloy, Journal of Materials Processing Technology, 2008, 200: 163-175.

Z. Zhang, Y.L. Liu, J.T. Chen. Effect of shoulder size on the temperature rise and the material deformation in friction stir welding, International Journal of Advanced Manufacturing Technology, 2009, 45: 889-895.

K. Kumar, S.V. Kailas. The role of friction stir welding tool on material flow and weld formation, Materials Science and Engineering A, 2008, 485: 367-374.

H.K.D.H. Bhadeshia, T. DebRoy. Critical assessment: friction stir welding of steels, Science and Technology of Welding and Joining, 2009, 14(3): 193-196.

J.H. Yan, M.A. Sutton, A.P. Reynolds. Process—structure—property relationships for nugget and heat affected zone regions of AA2524—T351 friction stir welds, Science and Technology of Welding and Joining, 2005, 10(6): 725-736.

Z. Zhang, H.W. Zhang. Numerical studies on controlling of process parameters in friction stir welding, Journal of Materials Processing Technology, 2009, 209(1): 241-270.

R. Nandan, G.G. Roy, T.J. Lienert, T. Debroy. Three-dimensional heat and material flow during friction stir welding of mild steel, Acta Materialia, 2007, 55: 883-895.

Z. Zhang. Comparison of two contact models in the simulation of friction stir welding process, Journal of Materials Science, 2008, 43: 5867-5877.

Z. Zhang, H.W. Zhang. A fully coupled thermo-mechanical model of friction stir welding, International Journal of Advanced Manufacturing Technology, 2008, 37: 279-293.

Z. Zhang, J.T. Chen, Z.W. Zhang, H.W. Zhang. Coupled thermo-mechanical model based comparison of friction stir welding processes of AA2024-T3 in different thicknesses, Journal of Materials Science, 2011, 46: 5815-5821.

Z. Zhang, H.W. Zhang. Numerical studies of pre-heating time effect on temperature and material behaviors in friction stir welding process, Science and Technology of Welding and Joining, 2007, 12(5): 436-448.

Z. Zhang, H.W. Zhang. Numerical studies on the effect of transverse speed in friction stir welding, Materials & Design, 2009, 30: 900-907.

T. Belytschko, W.K. Liu, B. Moran. Nonlinear finite elements for continua and structures, John Wiley & Sons, Ltd, New York, 2000.

J.R. An, Y.H. Liu. Handbook of Chinese and foreign metals materials most in use, Xi'an: Shanxi Science and Technology Press, 1998. (in Chinese)

H.W. Zhang, H. Wang, P. Wriggers, B.A. Schrefler. A finite element model for contact analysis of multiple Cosserat bodies, Computational Mechanics, 2005, 36(6): 444-458.

C.M. Chen, R. Kovacevic. Finite element modeling of friction stir welding—thermal and thermomechanical analysis, International Journal of Machine Tools and Manufacture, 2003, 43: 1319-1326.

Modelling compressive crush of composite tube reinforced foam sandwiches

∗ Jin Zhou[1], Zhongwei Guan[1] and Wesley Cantwell[2]

[1]School of Engineering, University of Liverpool, Liverpool L69 3GQ, UK
[2]Aerospace Research and Innovation Center (ARIC), Khalifa University of Science , Technology and Research, Abu Dhabi, UAE

*Jin Zhou: jinzhou@LIV.AC.UK

Abstract

The finite element (FE) models were developed to predict the compressive response and energy-absorbing capability of composite tube reinforced PVC foams. A vectorized user material subroutine (VUMAT) was employed to define Hashin's 3D damage criteria for the composite tube to model the corresponding deformation and failure mechanisms. Good agreement was obtained in terms of the load–displacement traces, the deformation and failure modes. Using validated models, parametric studies were further carried out to investigate the crushing characteristics of composite tube reinforced foam to optimize the composite tube configurations within the foam.

It has been shown that reinforcing the foams with composite tubes results in a significant increase in both their compressive strength and energy absorption relative to their plain counterparts. It has also been shown that the 12.5 mm carbon fibre tube reinforced foam out-performs the 10 and 8 mm tube reinforced foam based sandwich panel in term of energy absorption. The energy absorption increased with increasing of both foam density and tube diameters.

Keywords: Finite Element, Hashin 3-D Criteria, User-defined subroutine, Carbon tube reinforced foam, Energy absorption.

Introduction

As a result of their superior specific strength and stiffness characteristics, excellent fatigue properties and impressive corrosion resistance, composite materials, such as carbon fibre reinforced plastic (CFRP) are currently finding widespread use in a wide range of high-performance engineering structures. An additional attractive feature of these lightweight materials is their ability to absorb significant energy under certain well-defined loading conditions. Extensive testing has shown that composites, when produced in a tubular form and loaded in compression, are capable of absorbing significant energy through a range of failure mechanisms including fibre fracture, matrix cracking, debonding and delamination [Farley and Jones (1992)]. Over the years, this impressive energy-absorbing capability has attracted the interest of many vehicle manufacturers, including Chrysler and Ford. Indeed, [Jacob et al. (2002)] calculated that only 600 grams of composite is required to absorb the energy of a medium-sized car travelling at 35 mph. Alia investigated the energy-absorbing characteristics of polymer foams reinforced with small carbon fibre reinforced epoxy tubes [Alia et.al 2014]. Figure 1a highlights the extraordinary failure characteristics associated with composite materials, where a 10 mm diameter CFRP tube is being crushed at a low rate of strain. These images give evidences that these failure modes of carbon tube are typical ones of those observed in larger diameter tubes, with extensive splaying, fibre fracture and matrix cracking. Zhou et al. investigated the failure mechanism and energy absorption capacity of both carbon and glass rods experimentally. The progress failure of carbon and glass fibre rods subjected to compression was examined and shown in Figure 1b that the carbon fibre rod under compression demonstrates more ductile failure than the glass fibre one [Zhou et al. (2013)]. Clearly, the failure

pattern of the rod influences its energy absorbing capacity. If buckling failure can be avoided the energy absorption will be maximized. Therefore, it is necessary to introduce constraints to composite tubes or rods. An effective way to apply such constraints is to embed them into PVC foam, so that a progressive crushing of composite tube or rod can be realised. The study exhibits failed PVC foam core with embedded carbon fibre and glass fibre pins. Both carbon and glass fibre rods turn into dust, which indicate tests with successful constraints offered by PVC foam. The energy-absorbing capacity of a composite tube or rod is most frequently evaluated by determining its specific energy absorption (SEA) capability in J/kg. SEA values can vary greatly, for example, from 20 kJ/kg for a pultruded glass fibre/epoxy [Jacob et al. (2002)] to values well in excess of 100 kJ/kg for carbon fibre-based systems [Hamada (1993)]. The precise value depends on a number of parameters, including the geometry of the tube, its fibre architecture, as well as the mechanical properties of the matrix phase. For example, Hamada and co-workers showed that the energy-absorbing capacity of a 55 mm diameter CFRP tube decreased by fifteen percent in passing from a unidirectional tube to one with its fibres oriented at +/-25°. A number of researchers have studied the influence of specimen geometry on the energy-absorbing capability of composite tubes. [Thornton et al. (1979); Thornton and Edwards (1982)] investigated geometrical effects in the energy-absorbing response of tubes based on circular, square and rectangular cross-sections and showed that the former out-performed both their square and rectangular tubular counterparts. [Farley et al. (1986)] conducted tests on carbon and Kevlar fibre reinforced tubes, with ply orientations typical of those used in sub-floor beam structures and showed that the tube diameter to thickness ratio played a significant role in determining its subsequent strain energy-absorbing capacity. Similar trends have been observed by [Alia et al. (2014)] following tests on circular composite tubes, with values increasing by over fifty percent as the D/t ratio is reduced from approximately 42 to 6. This evidence suggests that the use of very low values of D/t can lead to greatly enhanced energy absorption in tubular structures. Following these initial tests on small diameter reinforcements, individual tubes were embedded in a polymer foam and crushed at quasi-static rates of strain

Figure 1 Failure modes of the composite tube and rods subject to crushing load (a. tube, b. rods)

Composite sandwich structures are increasingly finding use in a wide range of lightweight load-bearing engineering structures. Sandwich structures, such as those used in high-performance aerospace components, are typically based on thin composite. The variation of the specific energy absorption of circular CFRP tubes with diameter/thickness ratio by [Alia et al. (2014)]. The tube embedded in a polymer foam and skins bonded to a low density foam or honeycomb core. The skins are usually thin, often rendering these lightweight panels highly susceptible to damage by a hard projectile, such as that associated with runway debris or hail. A number of investigations have focused on the potential hazard resulting from an uncontained turbine engine failure on outer parts of an aircraft [Shockey (1997), Rouse et.al (1997). In such sandwich structures, the skin sheets carry bending loads, whilst the core resists transverse shear and through-thickness indentation forces. Therefore, to enhance the load carrying capacity it is desirable to maximize the through-thickness stiffness and strength of the core. One approach to achieve this goal is to add reinforcing pins to the core, with the ends of the pins embedded in woven carbon fibre skin sheets. [Cartie and Fleck (2003)] undertook the theoretical analysis and revealed that the through-thickness stiffness

and strength are relatively insensitive to the pin arrangements in pyramidal, tetrahedral and random patterns.

Since experimental trials are usually time-consuming and costly, it is evident that modelling the crushing behavior to investigate the energy-absorbing characteristics using commercial finite element software would be great interest. Once these models are verified, they can be used to predict the response of rods and tube reinforced foam based on different configurations, loading and boundary conditions without undertaking experimental tests. A number of numerical work has been carry out to modeling the response of composite tubes. Carla McGregora developed a model to predict the damage propagation, failure mode and energy absorption in triaxially braided composite tubes under axial compression using LS-DYNA [Carla McGregora et.al (2010)]. The two-ply and four-ply square tubes were modeled to predict energy absorbing of front rail structures components on vehicular under axial crushing. A micromechanical model incorporated as a subroutine coded into the ABAQUS implicit by [Beard and Chang (2002)] to simulate the complete crushing process of plug-initiated triaxially braided composite tubes with promising initial results. The developed model was incorporated into ABAQUS/Explicit to model dynamic response of tube under crushing load by [Flesher (2006)] Another damage model on composite (Mat_58) in LS-DYNA with a lower accuracy that the predicted SEA values of un-initiated tubes were 30-40% lower than experimental results [Xiao et al. (2009)]. A developed model for composite tube was only able to capture the axial crushing features of plug-initiated braided composite tubes accurately [C.J. McGregor et al. (2007)], however the simulation of the this model on un-initiated tubes was not successfully on failure modes due to model instability, there was a discrepancies between predicted and observed failure modes.

It is a challenging task to develop a model that is able to capture both the energy absorption and failure mode. A few researchers have attempted to model crushing of composite tubes and to simulate a similar splaying mode of failure. Mamalis et al. developed a finite element modeling to simulate axial collapse of CFRP square tubes under static and dynamic load. The model introduced a third layer to model the resin layer into pulverized debris during axial crushing. The deletion of the failure elements in the middle layer resulted a low energy absorption [Mamalis et al. (2006)]. A splaying mode of failure on glass/polyester tubes has been developed by Silcock et al using LS-DYNA [Silcock et al. (2006)]. The model employed a spotweld approach and pre-defining a debris wedge to simulate a delamination, initiation and propagation of the splaying failure mode. Although the failure modes were simulated successful, the correlation between measured and predicted load-displacement profiles was lower. The most successful model to simulate the splaying mode of failure on tube was developed by Pinho et al. using a decohesion element incorporated into ABAQUS. Both the delamination and the transverse tearing through thickness between the composite layers fronds were simulated. Both the load and failure modes was captures reasonably, however only a portion of the propagation during the crushing process was modeled. [Pinho et al. (2004)]

Although a few numerical modeling developed to simulate the response of composite using commercial software LS-DYNA and Abaqus. However, those such as ABAQUS only has a number of failure criteria for composite materials modeled using 2D elements, such as plane stress and continuum shell elements [Carla McGregor (2010)]. Further, none of these criteria consider strain-rate effects in composite materials, which is clearly important in dynamic studies. The 2D elements, with the existing failure criteria, are not capable of taking large through-the-thickness rate-dependent deformations into account. Therefore, it is necessary to develop a constitutive model with associated failure criteria suitable for simulating a composite material using 3D solid elements. Limited numerical modeling was developed to investigate the structural response of composite using three-dimensional 3D solid elements. Recently, Thuc et al developed a FE models which were

validated using experimental data from tests on FMLs based on a 2024-O aluminium alloy and a woven glass–fibre/polypropylene composite. The rate-dependent failure criteria for a unidirectional composite were used, which were based on the modified Hashin's 3D failure criteria [Thuc et al. (2013]. The constitutive model and failure criteria were then implemented in ABAQUS/Explicit using the VUMAT subroutine. Based on the previous research [Thuc et al. (2013], A further parametric studies were carried out to investigate the influence of the properties of the aluminium alloy on the blast resistance of FMLs for aerospace applications. A vectorized user material subroutine (VUMAT) was employed to define Hashin's 3D rate-dependant damage constitutive model of the GFPP. [Thuc et al. (2014)] Sandwich panels based on three-dimensional woven S-glass/epoxy skins and a crosslinked PVC core were modelled using finite element techniques to investigate the effect of through-the-thickness stitching on the blast resistance of the panels by [Guan et al. (2014)]. The finite element model accurately predicted the failure modes and deformed shapes of the sandwich panels over the range of impulsive loading conditions.

The superior mechanical properties provided by a roll wrapped composite tube manufactured predominantly using high modulus (T700) unidirectional pre-preg carbon fibre oriented to provide maximum strength in the lateral (length-ways) axis, also the use of pre-preg reinforcement oriented at 90° to ensures that the tube has good crush/burst strength around the section of the tube. Their superior mechanical properties offers special energy absorption which mean that tubes of the same weight as an aluminium or steel tube can be much stronger, or that tubes of the same strength can be much lighter, contribute more energy absorption subject to compressive crushing.

This paper presents numerical modeling of compressive structural behavior of PVC foam core panels reinforced by CFRP tubes. Here, the foam was modeled as a crushable foam material with strain hardening. A vectorized user material subroutine (VUMAT) was employed to define Hashin's 3D damage criteria for the composite tube to model the corresponding deformation and failure mechanisms. Energy absorption of the sandwich panels made with different densities of the cores was also investigated. Modeling results were compared with the experimental results, in terms of load-displacement relationships, deformation and failure modes. Reasonably correlation was obtained.

2 Finite element modeling

2.1 PVC foam

Numerical models were developed to simulate the mechanical response of the tube reinforced foam subjected to comparison. The PVC foam core in the structure was modeled as a crushable foam subjected to compressive loading with rate-dependent strain hardening and both shear and ductile failure criteria. It was assumed that the Poisson's ratio of all of the foams was 0.32. The phenomenological yield surface proposed by [Deshpande and Fleck (2001)] for a closed-cell foam material, given by:

$$\phi \equiv \frac{1}{\left[1+(\alpha/3)^2\right]}\left[q^2+\alpha^2\sigma_m^2\right]-\sigma_y^2 \leq 0 \tag{1}$$

where σ_y is the uniaxial yield strength of the foam in tension or compression, q is the Von Mises stress, and σ_m is the mean stress. The term α defines the shape of the yield surface, which is given by

$$\alpha = \frac{3k}{\sqrt{(3k_t+k)(3-k)}} \tag{2}$$

203

where k and k_t are related to the ratios of the initial uniaxial yield stress σ_c^o and the hydrostatic tensile yield stress p_t to the hydrostatic compressive yield stress p_c^o, respectively.

The yield stress in hydrostatic compression, p_c describes the development of the size of the yield surface and is given as:

$$p_c(\varepsilon_{pl}^{vol}) = \frac{\sigma_c(\varepsilon_{pl}^{vol})\left[\sigma_c(\varepsilon_{pl}^{vol})\left(\frac{1}{\alpha^2}+\frac{1}{9}\right)+\frac{p_t}{3}\right]}{p_t+\frac{\sigma_c(\varepsilon_{pl}^{vol})}{3}} \tag{3}$$

where ε_{pl}^{vol} is defined as the plastic volumetric strain in the volumetric hardening model, and is set equal to ε_{pl}^{axial} the compressive plastic strain. The term, p_c can therefore be deterimed from a compression test on the foam. Mechanical properties of the foams investigated are shown in Table 1.

Here, it is assumed that the response of a rate-dependent solid obeys the uniaxial flow rate definition, which is given as:

$$\dot{\bar{\varepsilon}}_{pl} = h(q, \bar{\varepsilon}_{pl}, \theta) \tag{4}$$

in which the term h is a strain-hardening function, $\bar{\varepsilon}_{pl}$ is defined as the equivalent plastic strain, and the parameter θ is the temperature. The rate-dependent hardening curves can therefore be expressed as:

$$\bar{\varepsilon}(\bar{\varepsilon}_{pl}, \dot{\bar{\varepsilon}}_{pl}) = \sigma_y(\bar{\varepsilon}_{pl})R(\dot{\bar{\varepsilon}}_{pl}) \tag{5}$$

in which $\dot{\bar{\varepsilon}}_{pl}$ and R are defined as the equivalent plastic strain-rate and the stress ratio ($= \bar{\sigma}/\sigma_y$) respectively, which are given as:

$$\bar{\varepsilon}_{pl} = \int_0^t \sqrt{\frac{2}{3}\dot{\bar{\varepsilon}}_{pl}:\dot{\bar{\varepsilon}}_{pl}}dt \text{ and } R = \bar{\sigma}/\sigma_y \tag{6}$$

Damage initiation in the PVC foam was modelled by applying a ductile damage criterion combined with a shear damage criterion. The former assumes that the equivalent plastic strain at the onset of damage is a function of the stress triaxiality (ratio of the pressure stress to the effective stress) and strain-rate. The latter criterion assumes that the equivalent plastic strain at the onset of damage is a function of the shear stress ratio and strain-rate. The fracture strains corresponding to the initiation of ductile damage and shear damage and the related strain-rate need to be specified.

Table 1 Mechanical properties of the foams investigated used in this study [Zhou et al. (2012)].

	C40	C80	C130	C200
Density (kg/m^3)	40	80	130	200
Compressive modulus (MPa)	37	97	160	280
Compressive strength (MPa)	0.45	1.3	2.6	4.8
Compressive fracture strain	0.65	0.7	0.7	0.7
Tensile modulus (MPa)	28	66	110	175
Tensile strength (MPa)	0.7	2.0	3.8	6.0
Shear modulus (MPa)	13	30	47	75
Shear strength (MPa)	0.5	1.2	2.3	3.5
Shear fracture strain	0.08	0.23	0.30	0.30
Work of fracture in tension (kJ/m^2)	0.21	0.44	0.76	1.33
Work of fracture in shear (kJ/m^2)	4.5	12.6	27.6	44.2

Poisson's ratio	0.32	0.32	0.32	0.32

2.2 CFRP Tube

2.2.1 Carbon fibre woven composites

The superior mechanical properties provided by a composite tube manufactured predominantly using high modulus (T700) unidirectional pre-preg carbon fibre oriented to provide maximum strength in the lateral (length-ways) axis, also the use of pre-preg reinforcement oriented at 90° to ensures that the tube has good crush/burst strength around the section of the tube. Given that a roll wrapped carbon fibre composite tubes tubes are manufactured from special high-modulus Toray T700 unidirectional pre-preg carbon fibre oriented at 0° (down the length of the tube) and unidirectional E-Glass oriented at 90° (around the section of the tube) by placing fibres in a [0,90,0,90,0] pattern. The overall strength of tube equal 50% CF at 0° and 50% CF at 90° direction. A constitutive model and failure criteria suitable for simulating the solid geometry composite using 3D solid elements was employed to summate the failure mechanism of carbon fire tubes.

Failure criteria for laminated composites are available in ABAQUS, which can be applied for panel coordinate and continuum shell elements only. However, none of these existing criteria consider the third direction through-the-thickness and strain-rate effects in the composite material in a cylindrical coordinate system using 3D solid elements. In order to develop a constitutive model and failure criteria suitable for simulating the composite tube using 3D solid elements, a 3D rate-dependent failure criteria for a anisotropic composite is developed by modifying Hashin's 3D failure criteria [Hashin (1980), Thuc et.al (2012)], to include rate-dependent elastic moduli and strength properties. The failure criteria, with the related constitutive model, are implemented into ABAQUS/Explicit using a VUMAT subroutine provided by ABAQUS [ABAQUS Theory Manual. 6.11(2011)].

The material behaviour within the cross section is same in the lateral axis and roll directions according to the material test data provided by the manufacturer. Therefore, the developed Hashin's 3D failure criteria [Thuc et.al (2013] are able to simulate overall response of a roll wrapped composite layer in a cylindrical coordinate. The failure functions may be expressed as follows:

Fibre tension: $\quad (\sigma_{11} \geq 0): F_f^t = \left(\dfrac{\sigma_{11}}{X_{1t}}\right)^2 + \left(\dfrac{\sigma_{12}}{S_{12}}\right)^2 + \left(\dfrac{\sigma_{13}}{S_{13}}\right)^2, d_{ft} = 1$ $\qquad$ (7)

Fibre compression: $\quad (\sigma_{11} < 0): F_f^c = \dfrac{|\sigma_{11}|}{X_{1t}}, d_{fc} = 1$ $\qquad$ (8)

Matrix tension: $\quad (\sigma_{22} + \sigma_{33} \geq 0): F_m^t = \dfrac{(\sigma_{22} + \sigma_{33})^2}{X_{2t}^2} + \dfrac{\sigma_{23}^2 - \sigma_{22}\sigma_{33}}{X_{23}^2} + \dfrac{\sigma_{12}^2 + \sigma_{13}^2}{X_{12}^2}, d_{mt} = 1$ $\qquad$ (9)

Matrix compression: $(\sigma_{22} + \sigma_{33} \geq 0)$:

$$F_m^c = \left[\left(\frac{X_{2c}}{2S_{23}}\right)^2 - 1\right]\frac{(\sigma_{22} + \sigma_{33})}{X_{2c}^2} + \frac{(\sigma_{22} + \sigma_{33})^2}{4S_{23}^2} + \frac{\sigma_{23}^2 - \sigma_{22}\sigma_{33}}{X_{23}^2} + \frac{\sigma_{12}^2 + \sigma_{13}^2}{X_{12}^2}, d_{mc} = 1 \qquad (10)$$

where X_{1t}, X_{1c}, X_{2t}, X_{2c}, S_{12}, S_{13} and S_{23} are the various strength components [18] and d_{ft}, d_{fc}, d_{mt} and d_{mc} are the damage variables associated with the four failure modes.

The response of the material after damage initiation (which describes the rate of degradation of the material stiffness once the initiation criterion is satisfied) is defined by the following equation:

$$\sigma = C(d) \cdot \varepsilon \;,\; \sigma_{ij} = C_{ij} \cdot \varepsilon_{ij} \tag{11}$$

$$\begin{bmatrix} \sigma_{11} \\ \sigma_{22} \\ \sigma_{33} \\ \sigma_{12} \\ \sigma_{23} \\ \sigma_{13} \end{bmatrix} = \begin{bmatrix} C_{11}^0 & C_{12}^0 & C_{13}^0 & & & \\ C_{12}^0 & C_{22}^0 & C_{23}^0 & & & \\ C_{13}^0 & C_{23}^0 & C_{33}^0 & & & \\ & & & C_{44}^0 & & \\ & & & & C_{55}^0 & \\ & & & & & C_{66}^0 \end{bmatrix} \begin{bmatrix} \varepsilon_{11} \\ \varepsilon_{22} \\ \varepsilon_{33} \\ \varepsilon_{12} \\ \varepsilon_{23} \\ \varepsilon_{13} \end{bmatrix} \tag{12}$$

where C_{ij} is a 6 x6 symmetric damaged matrix, whose non-zero terms can be written as:

$$C_{11} = (1 - d_f)E_1(1 - v_{23}v_{32})\Gamma$$
$$C_{22} = (1 - d_f)(1 - d_m)E_2(1 - v_{13}v_{31})\Gamma$$
$$C_{33} = (1 - d_f)(1 - d_m)E_3(1 - v_{12}v_{21})\Gamma$$
$$C_{12} = (1 - d_f)(1 - d_m)E_1(v_{21} - v_{31}v_{23})\Gamma$$
$$C_{23} = (1 - d_f)(1 - d_m)E_2(v_{32} - v_{12}v_{31})\Gamma \tag{13}$$
$$C_{31} = (1 - d_f)(1 - d_m)E_1(v_{31} - v_{21}v_{32})\Gamma$$
$$C_{44} = (1 - d_f)(1 - s_{mt}d_{mt})E_1(1 - s_{mc}d_{mc})G_{12}$$
$$C_{55} = (1 - d_f)(1 - s_{mt}d_{mt})E_1(1 - s_{mc}d_{mc})G_{23}$$
$$C_{66} = (1 - d_f)(1 - s_{mt}d_{mt})E_1(1 - s_{mc}d_{mc})G_{13}$$

where the global fibre and matrix damage variables as well as the constant Γ are also defined as:

$$d_f = 1 - (1 - d_{ft})(1 - d_{fc})$$
$$d_m = 1 - (1 - d_{mt})(1 - d_{mc})$$
$$\Gamma = 1/(1 - v_{12}v_{21} - v_{23}v_{32} - v_{13}v_{31} - 2v_{21}v_{32}v_{13}) \tag{14}$$

where E_i is the Young's modulus in the i direction, G_{ij} is the shear modulus in the i–j plane and v_{ij} is the Poisson's ratio for transverse strain in the j-direction, when the stress is applied in the i-direction. The Young's moduli, shear's moduli, Poisson's ratios and strengths of the CFPP are given in Table 2 and 3. The factors s_{mt} and s_{mc} in the definitions of the shear moduli are introduced to control the reduction in shear stiffness caused by tensile and compressive failure in the matrix respectively. The following values are recommended in [ABAQUS Theory Manual (2011)]: $s_{mt} = 0.9$ and $s_{mc} = 0.5$.

Table 2. Properties data for the CFRP tube

E_1 (MPa)	E_2 (MPa)	E_3 (MPa)	G_{12} (MPa)	G_{13} (MPa)	G_{23} (MPa)	v_{12} (MPa)	v_{13} (MPa)	v_{23} (MPa)	ρ (kg/m³)
70	70	10	8.6	8.6	8.6	0.1	0.3	0.3	1600

Table 3. Damage initiation data for the CFRP tube

X_{1T} (MPa)	X_{1C} (MPa)	X_{2T} (MPa)	X_{2C} (MPa)	S_{12} (MPa)	S_{13} (MPa)	S_{23} (MPa)
600	570	600	570	280	280	280

2.2.2. Strain-rate effects in the mechanical properties

The effects of strain-rate on the mechanical properties of a composite material are typically modelled using strain-rate dependent functions for both the elastic modulus and the strength. Yen developed logarithmic functions to account for strain-rate effects in a composite material as follows [Yen (2012)]:

$$\{S_{RT}\} = \{S_0\}\left(1 + C_1 In \frac{\dot{\bar{\varepsilon}}}{\dot{\bar{\varepsilon}}_0}\right)$$

$$\{E_{RT}\} = \{E_0\}\left(1 + C_2 In \frac{\dot{\bar{\varepsilon}}}{\dot{\bar{\varepsilon}}_0}\right)$$

(15)

Where

$$\{\dot{\bar{\varepsilon}}\} = \left\{|\dot{\varepsilon}_1| \quad |\dot{\varepsilon}_2| \quad |\dot{\varepsilon}_1| \quad |\dot{\varepsilon}_2| \quad |\dot{\varepsilon}_{12}| \quad |\dot{\varepsilon}_{13}| \quad |\dot{\varepsilon}_{23}|\right\}^T$$

$$\{S_{RT}\} = \left\{X_{1t} \quad X_{2t} \quad X_{1c} \quad X_{2c} \quad S_{12} \quad S_{13} \quad S_{23}\right\}^T$$

$$\{E_{RT}\} = \left\{E_1 \quad E_2 \quad E_3 \quad G_{12} \quad G_{13} \quad G_{23}\right\}^T$$

(16)

and the subscript RT refers to the rate-adjusted values, the subscript 0 refers to the static value, $\dot{\bar{\varepsilon}} = 1^{-1}$ is the reference strain-rate, $\dot{\bar{\varepsilon}}$ is the effective strain-rate, C_1 and C_2 are the strain-rate constants, respectively.

2.3 Cohesive elements and material properties

The resin layer at the interface between 0° lateral axis and oriented at 90° across its diameter plies was modelled using cohesive elements available in ABAQUS [ABAQUS Users Manual (2011)]. The elastic response was defined in terms of a traction-separation law with uncoupled behaviour between the normal and shear components. The default choice of the constitutive thickness for modelling the response, in terms of traction versus separation, is 1.0, regardless of the actual thickness of the cohesive layer. Thus, the diagonal terms in the elasticity matrix and density should be calculated using the true thickness of the cohesive layer as follows:

$$K_{nn} = \frac{E_n}{t_c}, \quad K_{ss} = \frac{E_s}{t_c}, \quad K_{tt} = \frac{E_t}{t_c}, \quad \rho = \rho_c t_c$$

(17)

The quadratic nominal stress and energy criterion were used to model damage initiation and damage evolution, respectively. Damage initiated when a quadratic interaction function, involving the nominal stress ratios, reached unity. Damage evolution was defined based on the energy conjunction with a linear softening law. The mechanical properties of the cohesive elements were obtained from [Karagiozova et al. (2010)].

3 Implementation of the material model in ABAQUS/Explicit

The user defined VUMAT subroutine was developed to implement the material model and failure criteria described in the previous sections in ABAQUS/Explicit. During each time step of computation, this subroutine is compiled and enables ABAQUS/ Explicit to obtain the required information regarding the state of the material and the material mechanical response at each integration point of each element. The Hashin's 3D failure criteria outlined in Eq. (7-10) are calculated, and the elastic modulus and strength values are adjusted for strain-rate effects using Eq. (14) base on these stresses computed within the VUMAT subroutine using the given strains and the material stiffness coefficients. The element status, which determined by the failure criteria, is then changed from 1 to 0 when an element fails. Accompanying the change of element status, the stresses at that material point are reduced to zero and it no longer contributes to the model stiffness. The element is removed from the mesh when all of the material status points of an element have been reduced to zero.

The 3D tube reinforced foam panel consisted of the foam, the composite and the cohesive layers as three separate parts. The PVC foam core and composite layers for CFRP tubes were meshed using C3D8R elements, which are eight-noded, linear hexahedral elements with reduced integration and hourglass control. The mesh generation and boundary conditions shown in Figure 2. The interfaces between the composite layers were created using eight-node 3D cohesive elements (COH3D8). The core size is 30×30×20 (in mm) and the diameters of the tube modelled included 8, 10 and 12.5 mm. The loading platens on both the top and bottom of the panel are meshed using rigid surface elements. The compressive load is applied to the top platen, with an only degree of freedom in the vertical direction. The bottom platen is fully fixed. Given that the panels were symmetric in nature, a half of each panel can be modeled with the appropriate boundary conditions applied along the planes of symmetry. A condition of general contact interaction was defined between the two neighboring layers of composites. Surface-based tie constraints were imposed between the composite layer and the cohesive layer to model adhesion between the adjacent layers. The contact interaction property for interaction between the foam and composite layer was also defined.
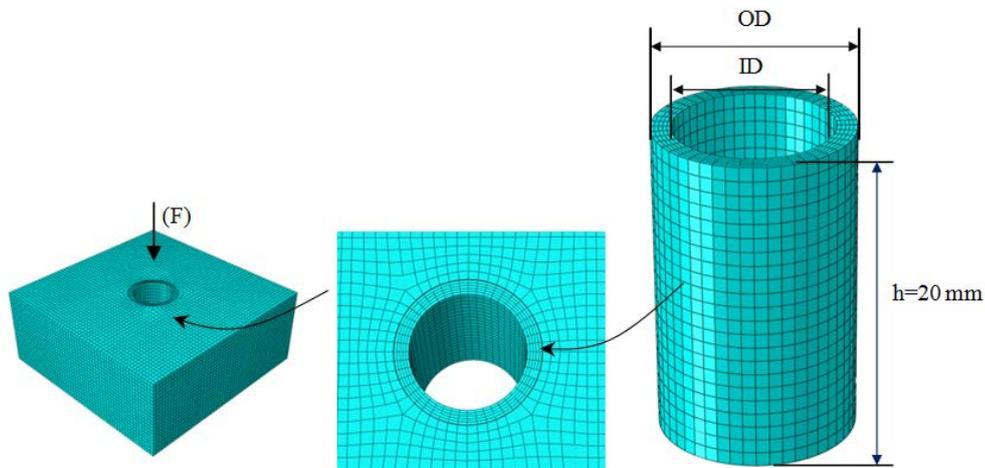
Figure 2 The geometry, mesh, boundary and loading conditions of the tube reinforced foam model.

4 Results and discussion
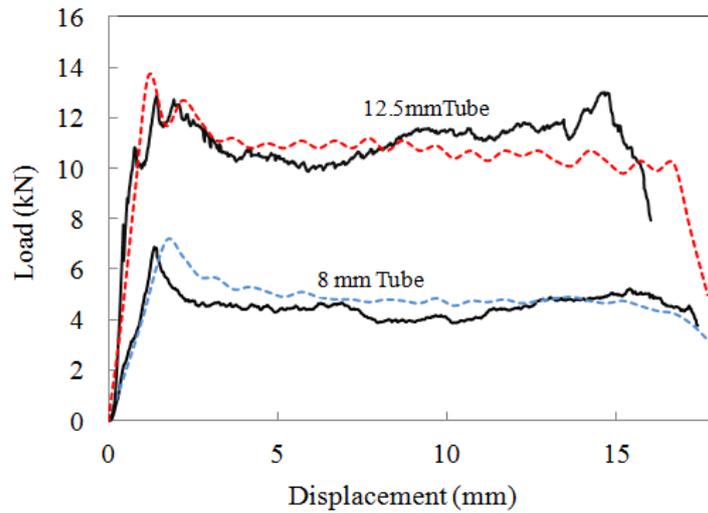
Figure 5 Load-displacement traces of individual carbon tube in diameter of 8, 10 and 12.5 mm under compression load.

a. Progress failre of crushing

b. Top view of the crushing of tube
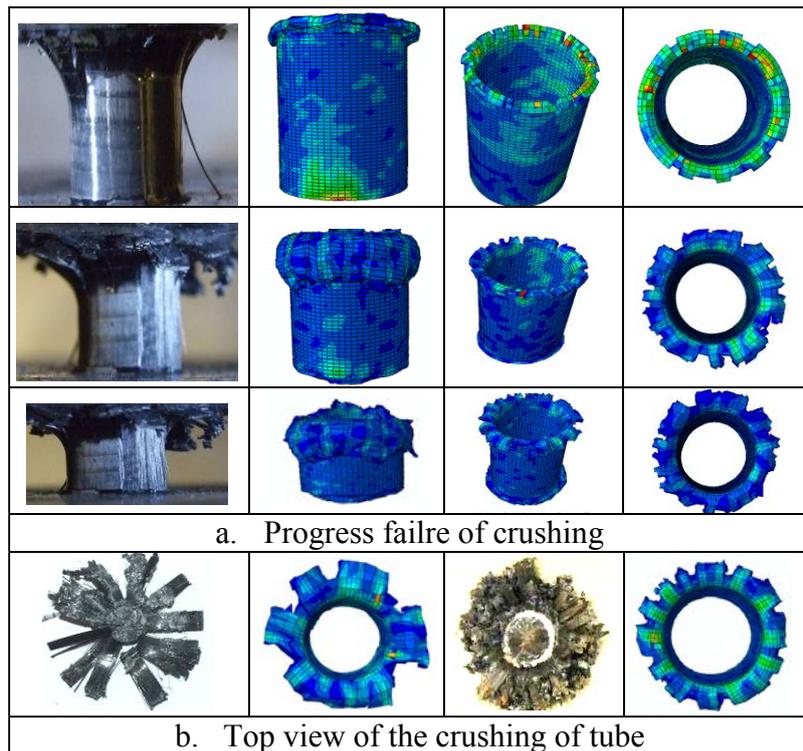
Figure 6 Comparison of progress deformation and failure for 10 mm CFRP tubes between test FE modeling.

The modeling for tube reinforced foam panel was developed based the rods reinforced foam. Prior to model the foam panel with embedded tube, the individual tube with foam support has been simulated and valuated with the experimental test on the tube. Figure 5 shows comparison of load-displacement traces obtained from FE prediction and the corresponding experimental results of the 8 and 12.5 mm tube individually. Reasonably good correlation has been obtained between the measurements and the numerical modeling in terms of the initial stiffness, the first peak load,

plateau load and damage evolution. It clearly shows that the predicted load from FE modelling is in a reducing trend after the first peak load during the compression process.

Figure 6 shows the comparison of progress deformation and failure for 10 mm CFRP tubes between test FE modelling. The basic features of the extensive splaying, fibre fracture and matrix cracking for the crushing tube were captured. The failed tubes is displayed the crushing states in which indicates a progress collapse of tubes. However, the failure modes of FE shows less extensive splaying of fibre may caused by the automatic remove of the failed element. The exact material data for the resin and fiber in longitudinal and circular direction can improve the failure modes.
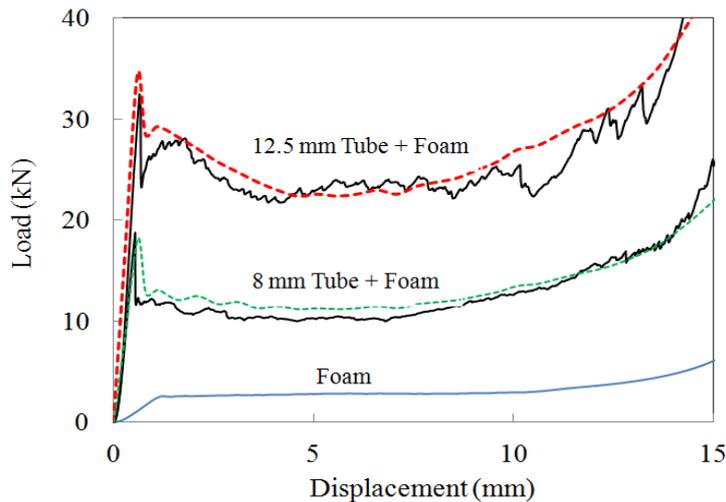
Figure 7 Comparison of load-displacement traces of C130 foam embedded in carbon tube in diameter of 8, 10 and 12.5 mm.

Here, PVC foam panels with densities of 40, 80 and 130 kg/m3 are embedded CFRP tubes in three diameters, i.e. 8, 10 and 12.5 mm. Figure 7 shows load-displacement traces obtained from numerical modeling and the corresponding experimental results of the 8 and 12.5 mm embedded in 130 kg/m^3 PVC panel. The test results of a plain PVC foam panel without any CF tube are also shown in the figure to evident the enhancement of carbon tube. Again, agreements between the experimental results and the finite element simulations are very good, with well captured features in the initial stiffness, the peak load, the damage evolution and the densification. Clearly, the resistance load increased significantly up to a average plateau load of 11.3 and 22.7 kN for the 8, and 12.5 mm tube reinforced foam panel respectively. It noted that the plateau load of 12.5 mm reinforced form panel is 8 times of the plain form panel without embedded tube. It also indicates that the resistance force increased with diameter of embedded tube form 11.3 kN for the 8 mm tube to 22.7 kN for the 12.5 mm tube reinforced foam.

C130T8 Test   C130T8 FE   C130T10 Test   C130T10 FE

a)   Cross-sections

b)   Ctushed foam panels
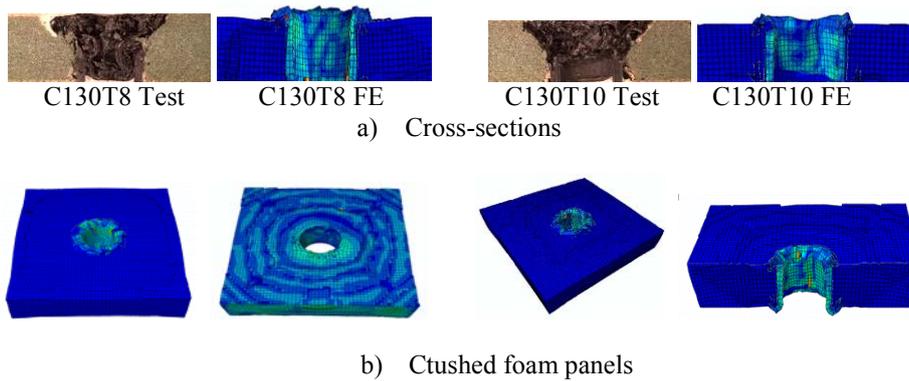
Figure 8 The comparison of cross-secion on tube embeded foam and fialure modes of curshed panel.

Figure 8 shows the cross-sections comparison of deformation and failure modes for C80 foam core panel with embedded CFRP tubes obtained from test and FE modelling. The core structure was deformed by 75% from its original configuration. The basic features of the foam crushing failure and the pin failure were captured. The failure modes of FE shows less crushing debris due to the failed elements removed automatic by the element control. The failed tubes is displayed the crushing states in which indicates a progress collapse of tubes embedded in the PVC foam. The more of the crushed failure modes of the tube reinforced panel are show in figures. The failure also indicates that strong constraint from the foam forces the CFRP tubes failure crushing along their longitudinal axis, which explains the enhancement of foam. However, the crushing failed elements in the modeling may cause element penetration with each other, which underestimates resistance of the tube to the compressive load.
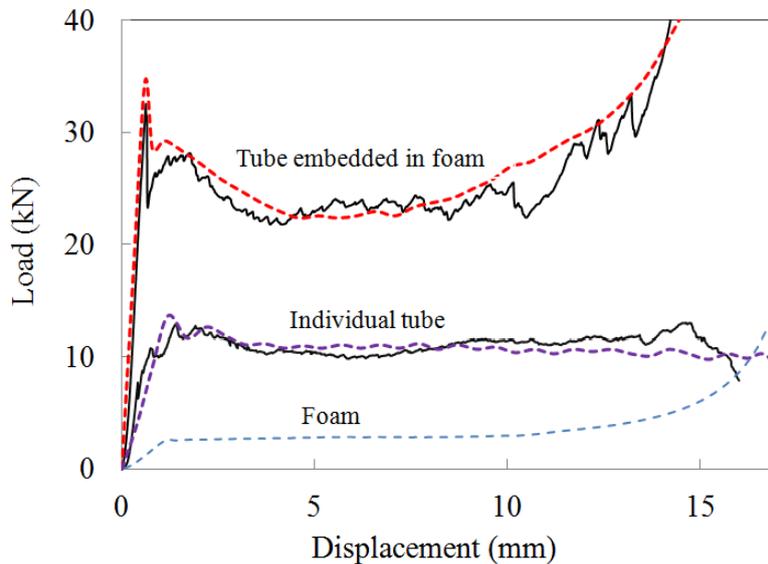
Figure 9 Comparison of load-displacement traces between individual carbon tube in diameter 12.5 mm and tube embedded in C130 foam under compression load.

The comparison of the load-displacement traces between individual tube, tube embedded in foam, and whole panel with a density of 130 kg/m^3 and embedded CF tube in diameters of 12.5 mm are exhibited in Figure 9. The dash line corresponding to the FE prediction whilst the solid line corresponding to the experimental test. An examination of the response of tube shows that the individual tube without foam constraint in a reduce trend during compression crushing, whilst the tube embedded in foam contribute huge resistance and energy absorption during crushing, also shows a increasing trend at the final stage of foam densification. The predicated load-displacement

cure of the embedded curve evident that the tube contribute over 80% load and energy absorption of foam panel and the embedded tube over perform more than twice of individual tube without foam support for the case of 12.5 mm tube embedded in C130 foam. This evidence clearly supports the suggestion that embedded tube in foam panel can modify the failure process and greatly enhance the crush performance of the tubes.
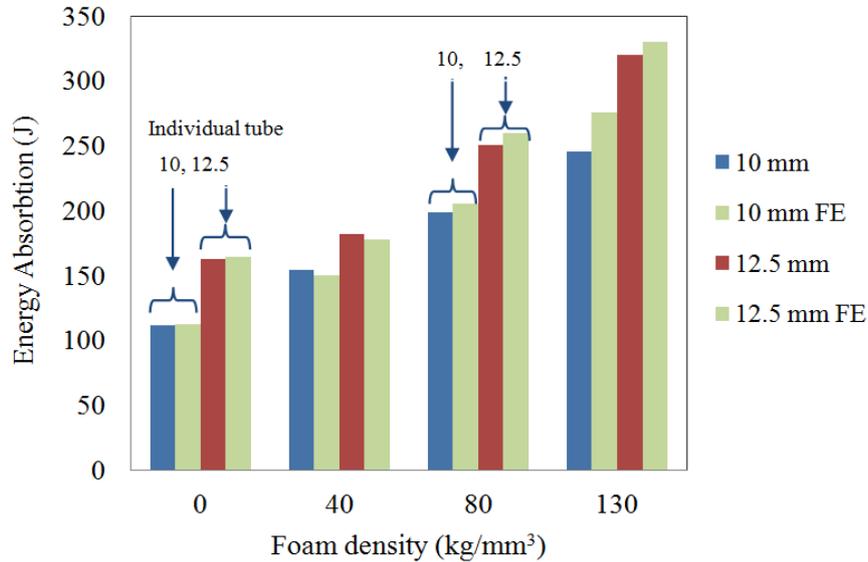
Figure 10 Comparison of energy absorption for individual tube and tube reinforced foam den between FE prediction and experimental test results.

Figure 10 shows the comprehensive comparison of energy absorptions obtained from experimental tests and FE predictions for individual and embedded carbon tube in diameters of 8, 10 and 12.5 mm in foam panel with density form 40 to 130 kg/mm^3. The green bar is FE prediction in the bar chart. In general, correlation is quite good between test data and FE predication for the individual tube on the energy bar chart with a difference less than 5%, whereas the difference on the tube reinforced foam slight higher, which may caused by the estimated parameters for the interaction between tube and foam core. The FE predictions for the 40 kg/m^3 foam panels are slightly lower than those of experimental measurements, whilst such the predictions for the higher density panels are slightly higher. The possible reason is that due to the weak constraint offered by the foam with the lower density foam, in the modeling such crushing causes element penetration with each other, which underestimates resistance of the tube to the compressive load. In the case of tube in higher density foam, the strong constraint from the foam forces the CFRP tubes failure crushing along their longitudinal axis, which offer continue resistance load without buckling failure and crushed totally. The FE simulation is ideal situation compare to the real state of experiment test.

Conclusions

User-defined constitutive models and strain-rate dependent failure criteria have been developed and implemented into finite element models to simulate load-displacement traces of PVC foam panels with embedded carbon fibre tubes, which are compared with the corresponding test results. Reasonably good correlation has been obtained between the experimental results and FE predictions, in terms of the initial stiffness, the peak load and the damage evolution. Here, three densities of the foam and three sizes of the CFRP tubes are investigated. In addition, energy absorption features of the sandwich core structures are captured by the modelling. The results show that the embedment of CFRP tubes inside PVC foam core is a very effective way to enhance energy absorption of this novel sandwich structure.

This evidence clearly supports the suggestion that embedded tube in foam panel can modify the failure process and greatly enhance the crushing performance of the tubes. This study also shows the advantage of FE modeling to predict and present the load of individual components for structural design analysis and optimization. The user-defined subroutines can be further developed to simulate more complex failure mechanisms of fibre reinforced composites.

References

Alia R.A., Cantwell W.J., Langdon G.S., Yuen S.C.K., Nurick G.N., (2014), The energy-absorbing characteristics of composite tube-reinforced foam structures, *Composites: Part B*, 61 , 127–135

ABAQUS, *Theory Manual. Version 6.11* （2011）. Pawtucket: Hibbitt, Karlsson & Sorensen, Inc.

ABAQUS/Explicit, Users Manual, Version 6.11（2011）, Hibbitt, Karlsson & Sorensen, Inc.;.

Beard S.J., Chang F.K. (2002), Design of braided composites for energy absorption, *Journal of Thermoplastic Composite Materials,* 15, 3–12

Cartie D. and Fleck N（2003）， The effect of pin reinforcement upon the through-thickness compressive strength of foam-cored sandwich panels, *Composites Science and Technology*,Vol. 16, 2401-2409.

Deshpande V.S.. and Fleck N.A. （2001） "Multi-axial yield behavior of polymer foams", *Acta mater.*, Vol. 49, 1859–1866

Farley, G. L. and Jones, J. M. (1986), Analogy for the effect of Material and geometrical variables on energy absorption capability of composite tube, *Journal of Composite Materials*, Vol. 20, 390-400

Flesher, N.D. (2006), Crash Energy Absorption of Braided Composite Tubes *Ph. D. Thesis, Department of Mechanical Engineering, Stanford University, Stanford, CA, USA.*

Farley, G. L. and Jones, J. M., Analogy for the effect of Material and geometrical variables on energy absorption capability of composite tube, *Journal of Composite Materials*, Vol. 20, 1986, pp. 390-400

Guan Z.W, Aktas A., Potluri P., Cantwell W.J., Langdon G., Nurick G.N. (2014), The blast resistance of stitched sandwich panels, International Journal of Impact Engineering 65, 137e145

Hashin Z. Failure criteria for unidirectional fiber composites(1980), J Appl Mech. 47,329–34.

Hamada H. et al, (1993) 38th International SAMPE Symposium and Exhibition, 38, pp952-966.

Jacob G. C., Fellers J. F. Simunovic, S., and Starbuck J. M. (2002), Energy absorption in polymer composites for automotive crashworthiness, *Journal of Composite Materials*, vol. 36, no. 7, 813–850,

Karagiozova D, Langdon GS, Nurick GN, Yuen SCK (2010). Simulation of the response of fibre–metal laminates to localised blast loading. Int J Impact Eng;37(6):766–82.

Mamalis, A.G, Manolakos, D.E, Ioannidis M.B, Papapostolou D.P. (2006), The static and dynamic axial collapse of CFRP square tubes: finite element modeling. *Composite Structures*, 74, 213–225

McGregor C.J., Vaziri R., Poursartip A., Xiao X. (2007), Simulation of progressive damage development in braided composite tubes under axial compression. *Composites Part A*: Applied Science and Manufacturing, 38, 2247–2259

Mc Gregor Carla, Vaziri Reza, Xiao Xinran, (2010) Finite element modelling of the progressive crushing of braided composite tubes under axial impact. *International Journal of Impact Engineering*, 37 662–672

Mc Gregora Carla , VazirIA Reza, Xiao Xinran (2010), Finite element modelling of the progressive crushing of braided composite tubes under axial impact. *International Journal of Impact Engineering,* Volume 37, 6, 662–672

Pinho S.T., Camanho P.P., Moura M.F. De, (2004), Numerical simulation of the crushing process of composite materials *International Journal of Crashworthiness,* 9, 263–276

Rouse M., Ambur D.M., Bodine J. and Dopker B (1997). Evaluation of a composite sandwich fuselage side panel with damage and subjected to internal pressure. *NASA Technical Memorandum* 110309.

Shockey A.A., Giovanola J.H., Simons J.W, Erlich D.C., Klopp R.W. and Skaggs S.R (1997). Advanced armor technology: application potential for engine fragment barriers for commercial aircraft, *Federal Aviation Administration Report* No. DOT/FAA/AR-97/53.

Silcock M.D., Hall W. and Fox B.L. (2006), Finite element modeling of composite tubular crash structures with an explicit code, in 6th Annual SPE Automotive Composites Conference, 669–682.

Thornton, P.H. (1979), Energy Absorption in Composite Structures, *Journal of Composite. Materials*, 13:247-262.

Thornton, P. H. and Edwards P. J. (1982), Energy Absorption in Composite Tubes, *Journal of Composite Materials*,16. 521–545.

Vo TP, Guan ZW, Cantwell WJ, Schleyer GK,(2012). Low-impulse blast behaviour of fibre–metal laminates. Composite *Structures,* 94(3),:954–65.

Vo TP, Guan ZW, Cantwell WJ, Schleyer GK(2013). Modelling of the low-impulse blast behaviour of fibre–metal laminates based on different aluminium alloys, *Composites: Part B* 44, 141–151

Xiao X., M. Botkin, Johnson N.L. (2009), Axial crush simulation of braided carbon tubes using Mat58 in LS-DYNA, *Thin-Walled Structures,* 47, 740–749

Yen CF. Ballistic impact modeling of composite materials (2002). In: *Proceedings of the 7th international LS-DYNA users conference*, vol. 6. 15–23.

Zhou J, Hassan M.Z., Guan Z.W. and Cantwell W.J（2012）. The low velocity impact response of foam-based sandwich panels; *Composites Science and Technology*, **72**, 1781-1790.

Numerical modelling of perforation impact damage of fibre metal laminates

* Jin Zhou[1], Zhongwei Guan[1] and Wesley Cantwell[2]

[1]School of Engineering, University of Liverpool, Liverpool L69 3GQ, UK
[2]Aerospace Research and Innovation Center (ARIC), Khalifa University of Science, Technology and Research (KUSTAR), Po.Box127788, Abu Dhabi, UAE.

*Jin Zhou: jinzhou@LIV.AC.UK

Abstract

Perforation damage of fibre metal laminates (FMLs) subjected to projectile impact was modelled using the finite element (FE) analysis. Here, FMLs studied covered stacking sequences of 2/1 and 3/2 FMLs, which were made with different aluminium alloys (6161-O, 6061-T6, 7075-O) and glass fibre reinforced polymer (GFRP) layers. A vectorized user-defined material subroutine (VUMAT) was developed to define Hashin's 3D rate-dependant damage criteria for the GFRP. The subroutine was implemented into the commercial finite element code ABAQUS/Explicit to simulate the deformation and failure of FMLs. The aluminium alloy layers were modelled as an isotropic elasto-plastic material by Johnson-Cook plasticity and the related damage criterion. The resin layer was simulated using cohesive elements, defined in terms of traction-separation. Good agreement was obtained between the simulations and the experimental results, in terms of the load–displacement traces, the deformation and failure modes.

Keywords: Impact, Fibre metal laminates, Hashin 3-D Criteria, Finite Element, Progressive Failure

1. Introduction

Fibre metal laminates (FMLs) are advanced composite structural materials that have been attracting interest from a number of researchers to investigate the impact resistance [Reyes and Cantwell (2000); Vogelesang and Vlot (2000)]. In recent years, a number of studies had been conducted to investigate the low and high velocity impact behaviour of fibre metal laminates. Caprino et al. [Caprino et al. (2004)] performed low-velocity impact tests on fibre metal laminates made of 2024-T3 sheets and S2-glass/epoxy prepreg layers. Various impact masses, velocities, and energies were applied in the tests to investigate the influence of these factors on the impact response. For comparison purposes, similar tests were also performed on monolithic 2024-T3 sheets with the equivalent thickness. Abdullah and Cantwell [Abdullah and Cantwell (2006)] studied the impact behaviour of a glass fibre reinforced polypropylene FMLs and the results showed that the FML offered an impressive resistance subject to low and high velocity impact. They found that FMLs absorb more energy during plastic deformation in the aluminium and composite layers. A low velocity impact tests on glass fibre-based FMLs has been conducted by Vlot and Fredell. The FMLs offer a superior impact to both an aluminium alloy and a carbon fibre reinforced composite. [Vlot and Fredell (1993)]. Vlot also conducted impact tests on an aluminium alloy and different types of FML and composites [Vlot (1996)]. There was a crack at the carbon and aramid fibres based FMLs and the energy absorption lower than that of a glass fibre reinforced FML. A inspection on the tested specimens showed that the FMLs exhibited a similar indentation in size to those plain aluminium alloy. Vlot et al. conducted impact tests on GLARE and plain aluminium and showed that the FML exhibited an initial cracking energy. They alos found that the impact damage resistance of these FMLs increased with increasing glass/epoxy content [Vlot et al (1999)].

A numbers of FE modeling have been developed to simulate the impact response of FMLs using numerical echniques. Guan et al investigated the impact response of fibre metal laminates based on a woven polypropylene (PP) fibre reinforced composite and an aluminium alloy at velocities up to 150 m/s. Both the predicted failure modes and displacement of the FMLs was good agree with the test data [Guan et al,(2009)]. Payeganeh et al. developed a number of FE models to investigate the resistance force traces, deflection, in-plane strains and stresses in of FMLs subjected to low velocity impact loading [Payeganeh et al. (2010)]. The results shown that the stacking sequence, the masses and velocities of the impactor were important parameters in determining the impact response of the FMLs. Lannucci et al studied the failure mode the impact load on FMLs. Modelling of composite damage subjected to impact within the intermediate strain rate regime may be generally categorized into four approaches [Lannucci (2006)], i.e. (1) failure criteria, (2) fracture mechanics, (3) plasticity or yield surface, and (4) damage mechanics. The Tsai-Wu failure criterion describes the failure surface in stress or strain space [Tsai and Wu (1971)]. However, it is a significant disadvantage to use stress-based failure criteria to model brittle materials as the scale effect in relation to the crack length in the same stress field cannot be modelled properly. Lee et al. [Lee et al. (2001)] investigated the penetration and perforation behaviour of a 6061-T6 aluminium plate and a C12K33 carbon fibre reinforced 6061-T6 aluminium metal-matrix composite plate subjected to projectile impact using an explicit finite element code, LS-DYNA3D. Perforation of the plate was found to occur under all of the studied impact conditions. The deformation behaviour of the plate and projectile as well as the projectile post-perforation velocity and the deceleration of the projectile were strongly dependent on the plate properties and impact velocity. Payeganeh et al. developed explicit FE models to investigate the contact force history, deflection, in-plane strains and stresses of 2024-O 2/1, 5/4 and 2024-T3 2/1, 5/4 FMLs subjected to low-velocity impact [Payeganeh et al. (2010)]. Failure shear strain and tension cut-off stresses were specified as failure criteria for aluminium layers. The failure of fibre laminate was simulated using Tsai–Wu failure criterion by specifying tensile cut-off stress based on the ultimate tensile stress of the fibre.

Although a few numerical modeling developed to simulate the response of composite using commercial software LS-DYNA and Abaqus. However, those such as ABAQUS only has a number of failure criteria for composite materials modeled using 2D elements, such as plane stress and continuum shell elements [Carla McGregor (2010)]. Further, none of these criteria consider strain-rate effects in composite materials, which is clearly important in dynamic studies. The 2D elements, with the existing failure criteria, are not capable of taking large through-the-thickness rate-dependent deformations into account. Therefore, it is necessary to develop a constitutive model with associated failure criteria suitable for simulating a composite material using 3D solid elements.

A limited numerical modeling were developed to investigate the structural response of composite using three-dimensional 3D solid elements. Recently, Thuc et al. developed a FE models which were validated using experimental data from tests on FMLs based on a 2024-O aluminium alloy and a woven glass–fibre/polypropylene composite. The rate-dependent failure criteria for a unidirectional composite were used, which were based on the modified Hashin's 3D failure criteria [Thuc et al. (2013]. The constitutive model and failure criteria were then implemented in ABAQUS/Explicit using the VUMAT subroutine. Based on the previous research [Thuc et al. (2013], A further parametric studies were carried out to investigate the influence of the properties of the aluminium alloy on the blast resistance of FMLs for aerospace applications. A vectorized user material subroutine (VUMAT) was employed to define Hashin's 3D rate-dependant damage constitutive model of the GFPP. [Thuc et al. (2014)] Sandwich panels based on three-dimensional woven S-glass/epoxy skins and a crosslinked PVC core were modelled using finite element techniques to investigate the effect of through-the-thickness stitching on the blast resistance of the

panels by [Guan et al. (2014)]. The finite element model accurately predicted the failure modes and deformed shapes of the sandwich panels over the range of impulsive loading conditions.

This paper presents numerical modeling of structural behavior of fibre metal laminates subjected to impact loading for aerospace applications. Here, Johnson–Cook strain hardening and damage criterion were employed to simulate the fialure of aluminiym layers. A vectorized user material subroutine (VUMAT) was employed to define Hashin's 3D damage criteria for the composite layer to model the corresponding deformation and failure mechanisms. Energy absorption of the fibre metal laminates plates made with different configurations of the liminates layers was also investigated. Modeling results were compared with the experimental results, in terms of load-displacement relationships, deformation and failure modes.

3 Finite element modeling

The two material layers in the FMLs, i.e. the composite and the aluminum alloy, exhibit very different mechanical behaviors. Therefore, different constitutive models were used to simulate the behavior of the metal and composite plies. The aluminium alloy layers were modelled as an isotropic elasto-plastic material by Johnson-Cook plasticity and the related damage criterion. ABAQUS/Explicit [Hibbitt et al. (2011)] was used to develop numerical simulations of the FMLs subjected to projectile impact. Numerical modeling was undertaken on the 6061-O, 6061-T6 and 7075-O FMLs outlined in Table 1.

Table 1 Johnson–Cook constants and static tensile strength for aluminium alloys

Aluminum type	A (MPa)	B (MPa)	n	C	D_1	D_2	D_3	D_4	Strength (MP)
Al 6061-T6	324	114	0.42	0.002	0.13	0.13	-1.5	0.011	332
Al 6061-O	360	105	0.73	0.083	0.013	0.025	-1.7	-0.4	310
Al 7075-O	535	658	0.71	0.024	-0.068	0.451	-0.95	0.036	551

3.1. Aluminium layers

The aluminium alloy was modelled as an elasto-plastic material included a rate-dependent behaviour. Temperature effects in the aluminium alloy were not taken into account. The Johnson–Cook material model was used in the form as below:

$$\sigma = [A + B(\bar{\varepsilon}_{pl})^n]\left[1 + C\ln\left(\frac{\dot{\bar{\varepsilon}}_{pl}}{\dot{\varepsilon}_0}\right)\right]$$

(1)

where $\bar{\varepsilon}_{pl}$ is the equivalent plastic strain; $\dot{\bar{\varepsilon}}_{pl}$ and $\dot{\varepsilon}_0$ are the equivalent plastic and reference strain rate and A, B, C and n are material parameters. Damage in the Johnson–Cook material model is predicted using the following cumulative damage law:

$$D = \sum\left(\frac{\Delta\bar{\varepsilon}_{pl}}{\bar{\varepsilon}_f^{pl}}\right)$$

(2)

in which

$$\bar{\varepsilon}_f^{pl} = [D_1 + D_2\exp(D_3\sigma^*)]\left[1 + D_4\ln\left(\frac{\dot{\bar{\varepsilon}}_{pl}}{\dot{\varepsilon}_0}\right)\right]$$

(3)

where σ^* is the mean stress normalised by the equivalent stress and $\Delta \bar{\varepsilon}_{pl}$ is the increment of equivalent plastic strain during an increment in loading. D, is a function of the mean stress and the strain rate. The parameters D_1, D_2, D_3, and D_4 are constants. Failure is assumed to occur when D = 1. Hence the current failure strain, $\bar{\varepsilon}_f^{pl}$, and thus the accumulation of damage,. The constants in the Johnson−Cook model for the three alluminium alloys used in this study are given in Table 2. The Young's modulus, Poisson's ratio and density of the various aluminium alloys were taken as E = 73.5 GPa, v = 0.3 and ρ = 2700 kg/m3, respectively.

3.2. Glass fibre reinforced composite layers

3.2.1. The 3D damage model for the composite material
A constitutive model and failure criteria suitable for simulating the solid geometry composite using 3D solid elements was employed to simulate the failure mechanism of glass fire layers. Failure criteria for laminated composites are available in ABAQUS, which can be applied for panel coordinate and continuum shell elements only. However, none of these existing criteria consider the third direction through-the-thickness and strain-rate effects in the composite material in a coordinate using 3D solid elements. In order to develop a constitutive model and failure criteria suitable for simulating the composite tube using 3D solid elements, a 3D rate-dependent failure criteria for a anisotropic composite is developed by modifying Hashin's 3D failure criteria [Hashin (1980), Thuc et.al (2012)], to include rate-dependent elastic moduli and strength properties. The failure criteria, with the related constitutive model, are implemented into ABAQUS/Explicit using a VUMAT subroutine provided by ABAQUS [ABAQUS Theory Manual. 6.11(2011)].

Given that a woven glass fibre composite layer is produced by placing fibres in a [0/90] pattern, the material behaviour within the plane of the laminate is similar in those two directions according to the material test data provided by the manufacturer. Therefore, the developed Hashin's 3D failure criteria [Thuc et.al (2013] be able to simulate overall response of a roll wrapped composite layer in a cylindrical coordinate. The failure functions may be expressed as follows:

Fibre tension: $\quad (\sigma_{11} \geq 0): F_f^t = \left(\dfrac{\sigma_{11}}{X_{1t}}\right)^2 + \left(\dfrac{\sigma_{12}}{S_{12}}\right)^2 + \left(\dfrac{\sigma_{13}}{S_{13}}\right)^2, d_{ft} = 1$ (4)

Fibre compression: $\quad (\sigma_{11} < 0): F_f^c = \dfrac{|\sigma_{11}|}{X_{1t}}, d_{fc} = 1$ (5)

Matrix tension: $\quad (\sigma_{22} + \sigma_{33} \geq 0): F_m^t = \dfrac{(\sigma_{22}+\sigma_{33})^2}{X_{2t}^2} + \dfrac{\sigma_{23}^2 - \sigma_{22}\sigma_{33}}{X_{23}^2} + \dfrac{\sigma_{12}^2+\sigma_{13}^2}{X_{12}^2}, d_{mt} = 1$ (6)

Matrix compression: $(\sigma_{22} + \sigma_{33} \geq 0)$:

$$F_m^c = \left[\left(\frac{X_{2c}}{2S_{23}}\right)^2 - 1\right]\frac{(\sigma_{22}+\sigma_{33})}{X_{2c}^2} + \frac{(\sigma_{22}+\sigma_{33})^2}{4S_{23}^2} + \frac{\sigma_{23}^2 - \sigma_{22}\sigma_{33}}{X_{23}^2} + \frac{\sigma_{12}^2+\sigma_{13}^2}{X_{12}^2}, d_{mc} = 1$$ (7)

where X_{1t}, X_{1c}, X_{2t}, X_{2c}, S_{12}, S_{13} and S_{23} are the various strength components and d_{ft}, d_{fc}, d_{mt} and d_{mc} are the damage variables associated with the four failure modes.

The response of the material after damage initiation (which describes the rate of degradation of the material stiffness once the initiation criterion is satisfied) is defined by the following equation:

$$\sigma = C(d) \cdot \varepsilon \quad , \quad \sigma_{ij} = C_{ij} \cdot \varepsilon_{ij} \tag{8}$$

$$\begin{bmatrix} \sigma_{11} \\ \sigma_{22} \\ \sigma_{33} \\ \sigma_{12} \\ \sigma_{23} \\ \sigma_{13} \end{bmatrix} = \begin{bmatrix} C_{11}^0 & C_{12}^0 & C_{13}^0 & & & \\ C_{12}^0 & C_{22}^0 & C_{23}^0 & & & \\ C_{13}^0 & C_{23}^0 & C_{33}^0 & & & \\ & & & C_{44}^0 & & \\ & & & & C_{55}^0 & \\ & & & & & C_{66}^0 \end{bmatrix} \begin{bmatrix} \varepsilon_{11} \\ \varepsilon_{22} \\ \varepsilon_{33} \\ \varepsilon_{12} \\ \varepsilon_{23} \\ \varepsilon_{13} \end{bmatrix} \tag{9}$$

where C_{ij} is a 6 x6 symmetric damaged matrix, whose non-zero terms can be written as:

$$C_{11} = (1 - d_f)E_1(1 - v_{23}v_{32})\Gamma$$

$$C_{22} = (1 - d_f)(1 - d_m)E_2(1 - v_{13}v_{31})\Gamma$$

$$C_{33} = (1 - d_f)(1 - d_m)E_3(1 - v_{12}v_{21})\Gamma$$

$$C_{12} = (1 - d_f)(1 - d_m)E_1(v_{21} - v_{31}v_{23})\Gamma$$

$$C_{23} = (1 - d_f)(1 - d_m)E_2(v_{32} - v_{12}v_{31})\Gamma \tag{10}$$

$$C_{31} = (1 - d_f)(1 - d_m)E_1(v_{31} - v_{21}v_{32})\Gamma$$

$$C_{44} = (1 - d_f)(1 - s_{mt}d_{mt})E_1(1 - s_{mc}d_{mc})G_{12}$$

$$C_{55} = (1 - d_f)(1 - s_{mt}d_{mt})E_1(1 - s_{mc}d_{mc})G_{23}$$

$$C_{66} = (1 - d_f)(1 - s_{mt}d_{mt})E_1(1 - s_{mc}d_{mc})G_{13}$$

where the global fibre and matrix damage variables as well as the constant Γ are also defined as:

$$d_f = 1 - (1 - d_{ft})(1 - d_{fc})$$

$$d_m = 1 - (1 - d_{mt})(1 - d_{mc}) \tag{11}$$

$$\Gamma = 1/(1 - v_{12}v_{21} - v_{23}v_{32} - v_{13}v_{31} - 2v_{21}v_{32}v_{13})$$

where E_i is the Young's modulus in the i direction, G_{ij} is the shear modulus in the i–j plane and v_{ij} is the Poisson's ratio for transverse strain in the j-direction, when the stress is applied in the i-direction. The Young's moduli, shear's moduli, Poisson's ratios and strengths of the CFPP are given in Table 2 and 3. The factors s_{mt} and s_{mc} in the definitions of the shear moduli are introduced to control the reduction in shear stiffness caused by tensile and compressive failure in the matrix respectively. The following values are recommended in [ABAQUS Theory Manual (2011)]: s_{mt} = 0.9 and s_{mc} = 0.5.

Table 2. Properties data for the GFRP composite

E_1 (MPa)	E_2 (MPa)	E_3 (MPa)	G_{12} (MPa)	G_{13} (MPa)	G_{23} (MPa)	v_{12} (MPa)	v_{13} (MPa)	v_{23} (MPa)	ρ (kg/m^3)
13	13	2.4	1.72	1.72	1.72	0.1	0.3	0.3	1800

Table 3. Damage initiation data for the GFRP composite

X_{1T} (MPa)	X_{1C} (MPa)	X_{2T} (MPa)	X_{2C} (MPa)	S_{12} (MPa)	S_{13} (MPa)	S_{23} (MPa)
320	240	320	240	140	140	140

2.2.2. Strain-rate effects in the mechanical properties

The effects of strain-rate on the mechanical properties of a composite material are typically modelled using strain-rate dependent functions for both the elastic modulus and the strength. Yen [Yen (2012)] developed logarithmic functions to account for strain-rate effects in a composite material as follows:

$$\{S_{RT}\} = \{S_0\}\left(1 + C_1 In \frac{\dot{\varepsilon}}{\dot{\varepsilon}_0}\right)$$

$$\{E_{RT}\} = \{E_0\}\left(1 + C_2 In \frac{\dot{\varepsilon}}{\dot{\varepsilon}_0}\right)$$

(12)

Where

$$\{\dot{\varepsilon}\} = \left\{|\dot{\varepsilon}_1| \quad |\dot{\varepsilon}_2| \quad |\dot{\varepsilon}_1| \quad |\dot{\varepsilon}_2| \quad |\dot{\varepsilon}_{12}| \quad |\dot{\varepsilon}_{13}| \quad |\dot{\varepsilon}_{23}|\right\}^T$$

$$\{S_{RT}\} = \{X_{1t} \quad X_{2t} \quad X_{1c} \quad X_{2c} \quad S_{12} \quad S_{13} \quad S_{23}\}^T$$

$$\{E_{RT}\} = \{E_1 \quad E_2 \quad E_3 \quad G_{12} \quad G_{13} \quad G_{23}\}^T$$

(13)

and the subscript RT refers to the rate-adjusted values, the subscript 0 refers to the static value, $\dot{\varepsilon} = 1^{-1}$ is the reference strain-rate, $\dot{\varepsilon}$ is the effective strain-rate, C_1 and C_2 are the strain-rate constants, respectively.

2.3 Cohesive elements and material properties

The resin layer at the interface between 0° lateral axis and oriented at 90° across its diameter plies was modelled using cohesive elements available in ABAQUS [ABAQUS Users Manual (2011)]. The elastic response was defined in terms of a traction-separation law with uncoupled behaviour between the normal and shear components. The default choice of the constitutive thickness for modelling the response, in terms of traction versus separation, is 1.0, regardless of the actual thickness of the cohesive layer. Thus, the diagonal terms in the elasticity matrix and density should be calculated using the true thickness of the cohesive layer as follows:

$$K_{nn} = \frac{E_n}{t_c}, \quad K_{ss} = \frac{E_s}{t_c}, \quad K_{tt} = \frac{E_t}{t_c}, \quad \rho = \rho_c t_c$$

(14)

The quadratic nominal stress and energy criterion were used to model damage initiation and damage evolution, respectively. Damage initiated when a quadratic interaction function, involving the nominal stress ratios, reached unity. Damage evolution was defined based on the energy conjunction with a linear softening law. The mechanical properties of the cohesive elements were obtained from [Karagiozova et al. (2010)].

3 Implementation of the material model in ABAQUS/Explicit

The user defined VUMAT subroutine was developed to implement the material model and failure criteria described in the previous sections in ABAQUS/Explicit. During each time step of computation, this subroutine is compiled and enables ABAQUS/ Explicit to obtain the required information regarding the state of the material and the material mechanical response at each integration point of each element. The Hashin's 3D failure criteria outlined in equations (4-7) are calculated, and the elastic modulus and strength values are adjusted for strain-rate effects using equations (11) base on these stresses computed within the VUMAT subroutine using the given strains and the material stiffness coefficients. The element status, which determined by the failure criteria, is then changed from 1 to 0 when an element fails. Accompanying the change of element status, the stresses at that material point are reduced to zero and it no longer contributes to the model stiffness. The element is removed from the mesh when all of the material status points of an element have been reduced to zero.

The fibre metal laminates consisted of the aluminum, the composite and the cohesive layers as three separate parts. The aluminum and composite layers for CFRP tubes were meshed using C3D8R elements, which are eight-noded, linear hexahedral elements with reduced integration and hourglass control. The mesh generation and boundary conditions shown in Figure 1. The interfaces between the composite layers were created using eight-node 3D cohesive elements (COH3D8). The plate size is 75×75 (in mm). The initial velocity applied to the projectile, with an only degree of freedom in the vertical direction. The plate edges are fully fixed. Given that the panels were symmetric in nature, a quart of each panel was modeled with the appropriate boundary conditions applied along the planes of symmetry. A condition of general contact interaction was defined between the two neighboring layers of composites. Surface-based tie constraints were imposed between the composite layer and the cohesive layer to model adhesion between the adjacent layers. The contact interaction property for interaction between the aluminum and composite layer was also defined.
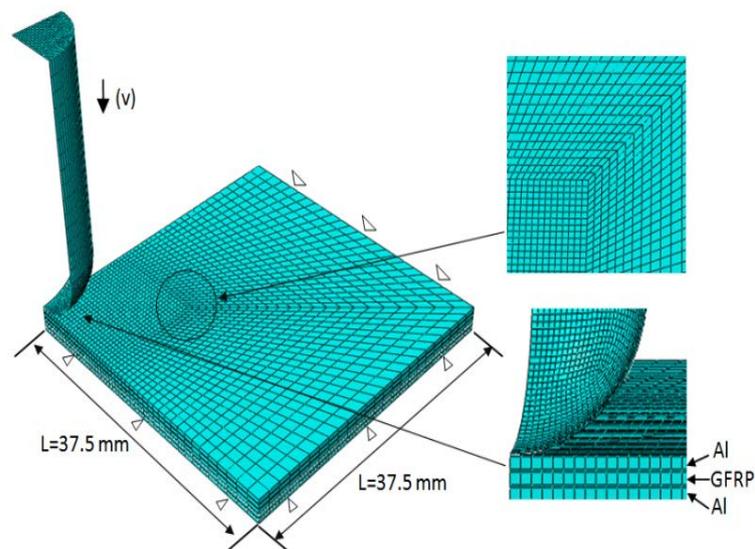
Figure 1 The geometry, mesh, boundary and loading conditions of the model for FMLs.

4 Results and Discussion

The developed finite element models have been simulated to predict the structural behaviour of fibre metal laminates subject to perforation loading. Modeling results were compared with the experimental results, in terms of load-displacement relationships, energy absorption, deformation and failure modes. The perforation loading on individual layers has been modeling firstly to validation the FE modeling. The perforation prediction on FMLs has been compared with the experimental data.

a. Aluminum layers   b. GFRP composite layers
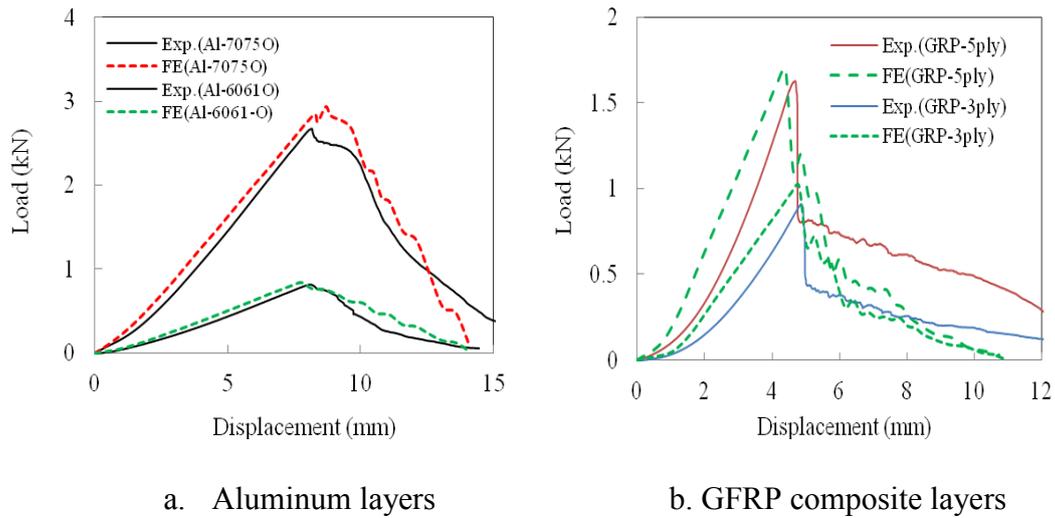
Figure. 2 Comparison of load-displacement traces of perforation tests for individual layers on aluminum and GFRP

Figures 2a show the comparison between the experimental and the numerical load-displacement traces for the individual layers of the 6061-o and 7075-O aluminums. The peak loads from the numerical predictions and the experimental tests for these layers were 750 and 3105 N, respectively. The former are only 3.1% and 9.2% higher than the latter, respectively. Also the predicted initial stiffness and the displacement at the peak load for the two aluminum were shown a good agreement with the corresponding experimental results. The predicted perforation energies were 5.8 and 20.8 J respectively, which are only 5.8% higher and 6.4% higher than the corresponding experimental results. Figures 2b present the comparison between the experimental and the numerical load-displacement traces for the 0.5 mm 3-ply and 1 mm 5-ply composite layers respectively subjected to a low velocity impact. The load-displacement traces shows that the stiffness is lower at beginning and the traced start liner up after the displacement of 1 mm. The peak loads from the numerical predictions and the experimental tests for these layers were 1015 and 1692 N, respectively. The former are only 4.1% and 7.2% higher than the latter, respectively. Also the predicted initial stiffness and the displacement at the peak load for the two GRP skins were in reasonably good agreement with the corresponding experimental results. The predicted perforation energies were 3.35 and 5.51 J respectively, which are only 5.8% higher and 6.4% higher than the corresponding experimental results. The fibre metal laminate was simulated using the validated individual layers.

a. 6061-O based FMLs   b. 7075-O based FMLs
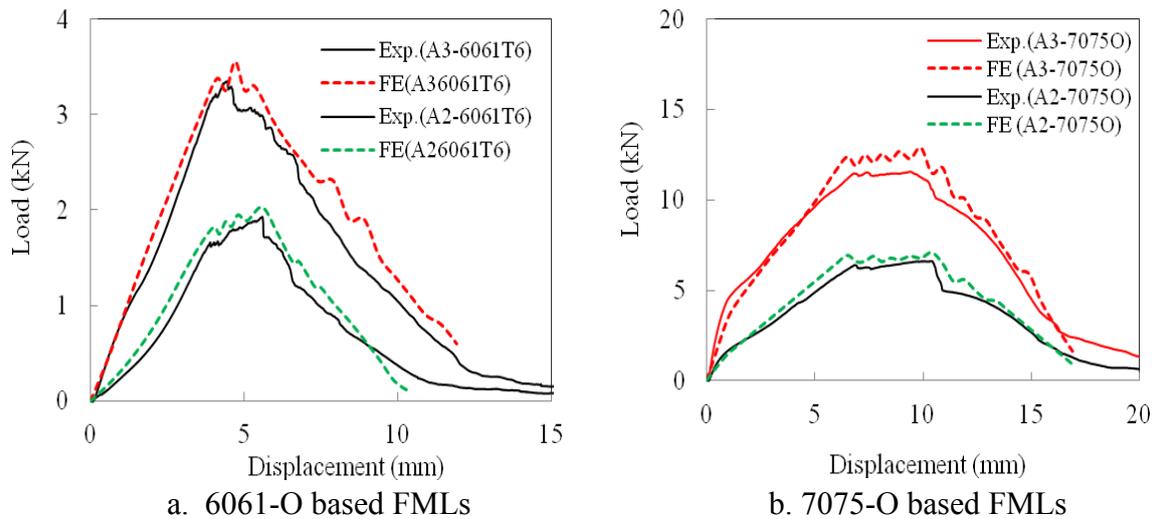
Figure. 3 Comparison of load-displacement traces of 2/1 and 3/2 fibre metal laminates between FE and experimental tests.

The finite element models using the constitutive models and failure criteria presented in the figure 2 were developed to simulate the critical perforation impact tests of various fibre metal laminates. Figures 3a and 3b show the simulated and the related experimental load-displacement traces of 2/1 and 3/2 FMLs plates made with 3-ply and 5-ply composite layers respectively subjected to low velocity impact. Figure 3a shows the 0.5 mm thick 6061-O aluminum and 3-ply GFRP based FMLs. The load-displacement traces show a linear up stiffness until the fist peak load. The predicted peak loads for the 6061-O based 2/1 and 3/2 FMLs plates were 203 and 355 Newtons, respectively, which are 2.4% and 6.0% higher than the experimental results respectively. The predicted initial stiffness and the displacement at the peak load for the targets were in good agreement with the corresponding experimental results. The predicted perforation energies were 11.09 and 23.65 Joules, respectively. In comparison with the experimental results they were slightly higher, respectively.

Figure 3b shows the 1 mm thick 7075-O aluminum and 5-ply GFRP based FMLs. The load-displacement traces show a linear up stiffness up to the displacement of 1 mm. The stiffness reduced and shows a lower liner stiffness before up to first peak load. The reduced stiffness maybe caused by the delamination between aluminum and composite layer which cased the stiffness reduced. The predicted peak loads for the 7075-O based 2/1 and 3/2 FMLs plates were 7100 and 12900 Newtons, respectively, which are 5.4% and 8.1% higher than the experimental results respectively. The predicted initial stiffness and the displacement at the peak load for the targets were in reasonably agreement with the corresponding experimental results. The predicted perforation energies were 74.4 and 132.5 Joules, respectively. In comparison with the experimental results the results still slightly higher.

0.5mm 6061-O + 0.5mm 3-ply GRP

1mm 7075-O + 1mm 5-ply GRP
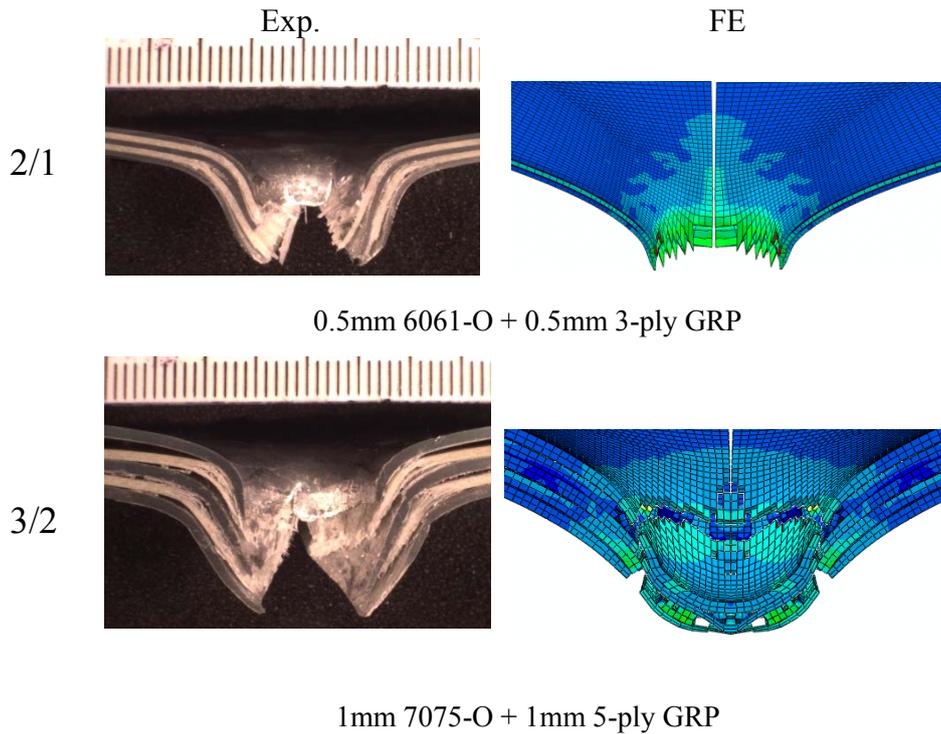
Figure 5 Comparison of the simulated and experimental failure modes of 3/2 FMLs plates made with 3-ply and 5-ply composite

Figures 5 show the comparison of the simulated and experimental failure modes of 3/2 FMLs plates made with 3-ply and 5-ply composite subjected to an on-set perforation impact. The basic features of the experimental failure modes for all the FMLs plates were well simulated, in terms of the cross cracks at the rear face and the local deformation mode at the target centre. Since the difference between the FMLs plates was thickness of aluminum and the number of composite plies in the composite layer, the experimental failure modes for these two FMLs plates were quite similar. The FE simulate the delamination of resin between composite and aluminum.
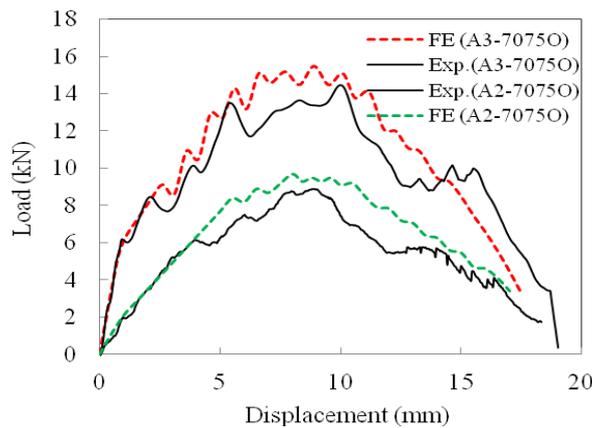
Figure. 6 Comparison of load-displacement traces of impact tests on 2/1 and 3/2 7075-O based fibre metal laminates between FE and experimental tests.

Finite element models of other types of FMLs plates subjected to a low velocity impact were also developed include the strain rate effects to broaden the validation. Figures 6 show the numerical simulations of the experimental load-displacement traces for the 3/2 FMLs plates made with 5-ply composite cores respectively subjected to an on-set perforation impact. Very good correlation was obtained between the experimental results and the numerical simulations, in terms of the overall

initial stiffness, the peak load and the perforation process. The predicted peak loads for these two FMLs plates were 970 and 1551 N, respectively, which are only 5.4 % and 5.8 % higher in comparison with the experimental results, respectively. The predicted initial stiffness and the predicted displacement at the peak load were also shown reasonably agreement with the corresponding experimental results. The predicted perforation energies for these two plates were 103.4 and 175.5 J, respectively. Compared to the experimental results, the FE were reasonable higher.
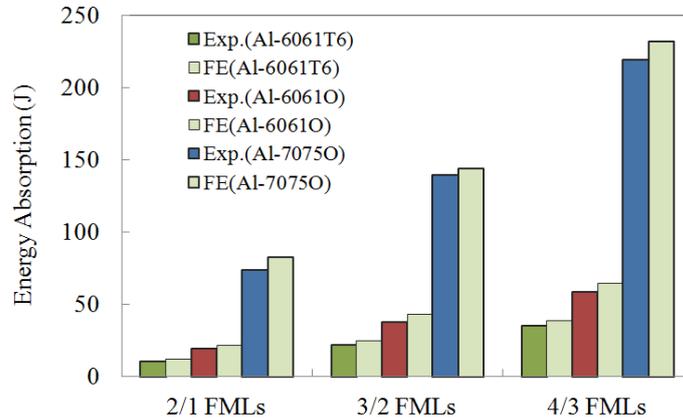
Figure. 7 Comparison of energy absorption of 2/1, 3/2 and 4/3 fibre metal laminates between FE and experimental tests.

Figures 7 show the comparison between the the perforation energy and the corresponding test results in a chart form. Clearly, very good correlation was obtained. The green bar show the FE prediction of the modeling and all the FE results are slight higher than the experimental data as the load-displacement traces shown in early figure. The possible reason may caused by the contact parameters used for the contact between projectile and FMLs, the strain rate of the modeling, elements control of the modeling. In further studies, more points can to be predicted by using validated numerical models in order to draw out the reliable relationship, In fact, the finite element models developed are well validated based on the reasonably good prediction compare to the test results.

Conclusions

Finite element models have been developed to simulate the structural behaviour of fibre metal laminates with various stacking sequences and three different aluminium alloys subjected to impact loading. Hashin's 3D failure criteria, incorporating strain-rate effects in the GFPP is implemented into ABAQUS/Explicit using a vectorized user-defined material subroutine (VUMAT). Very good correlation has been obtained between the numerical simulations and the experimental results, in terms of load-displacement traces, peak load and perforation energy. A reasonable agreemetn bas been shown in deformation mode and failure mode.

The validated finite element models, which cover the configurations of 2/1, 3/2 and 4/3 laminates made with different layers included 3-ply and 5-ply composite and various thinckness of aluminiums are ready to be used for further parametric studies of FMLs subjected to different loading conditions. The evidence suggests that the impact resistance and energy absorption increased with the increasing of laminates thickness and area density. Both the peak load and erergy absorption of 6061-T6 overperform than the 6061-O based fibre metal laminates, however the specifii erergy absorption of later slight higher than former. It also a suggests that the 7075-O alloy offers the best impact resistance and energy absorptions.

References

ABAQUS/Explicit, Users Manual, Version 6.11（2011）, Hibbitt, Karlsson & Sorensen, Inc.

ABAQUS, Theory Manual. Version 6.11（2011）. Pawtucket: Hibbitt, Karlsson & Sorensen, Inc.

Abdullah MR, Cantwell WJ (2006) The impact resistance of polypropylene-based fibre metal laminates. *Compos Sci Techno*l;66:1682‒93.

Caprino G, Spataro G, Del LS (2004). Low-velocity impact behaviour of fibreglassaluminium laminates. *Compos Part A,* 35:605‒16.

Deshpande V.S.. and Fleck N.A.（2001）"Multi-axial yield behavior of polymer foams", *Acta mater*., Vol. 49, 1859–1866.

Fan J., Guan Z.W. and Cantwell W.J.（2011）Numerical modelling of perforation failure in fibre metal laminates subjected to low velocity impact loading. *Composite Structures*, Vol. 93, pp2430–2436.

ABAQUS/Explicit, *Users Manual, Version 6.11（2009）*, Hibbitt, Karlsson & Sorensen, Inc.

ABAQUS, *Theory Manual. Version 6.11（2012）*. Pawtucket: Hibbitt, Karlsson & Sorensen, Inc.

Guan ZW, Cantwell WJ, Abdullah R. Numerical modeling of the impact response of fibre-metal laminates. *Polym Compos* 2009;30:603‒11.

Guan Z.W, Aktas A., Potluri P., Cantwell W.J., Langdon G., Nurick G.N. (2014), The blast resistance of stitched sandwich panels, *International Journal of Impact Engineering* 65, 137e145

Hashin Z. Failure criteria for unidirectional fiber composites(1980), *J Appl Mech.* 47,329–34.

Karagiozova D, Langdon GS, Nurick GN, Yuen SCK (2010). Simulation of the response of fibre–metal laminates to localised blast loading. *Int J Impact Eng;37*(6):766–82.

Lee W S, Lai C H, Chiou S T (2001). Numerical study on perforation behaviour of 6061-T6 aluminium matrix composite. *J Mater Process Technol*, , 117: 125–131

Payeganeh GH, Ghasemi FA, Malekzadeh K. (2010) Dynamic response of fibre-metal laminates (FMLs) subjected to low-velocity impact. *Thin-Walled Struc*;48:62‒70

Reyes G, Cantwell WJ (2000). The mechanical properties of fibre-metal laminates based on glass fibre reinforced polypropylene. *Compos Sci Technol*;60:1085–94.

Lannucci L (2006). Progressive failure modelling of woven carbon composite under impact. *Int J Impact Eng,* , 32: 1013–1043

Tsai S W, Wu E (1971). A general theory of strength for anisotropic materials. J Compos Mater, , 5: 58–80

Vlot A, Fredell RS (1993). Impact damage resistance and damage tolerance of fibre metal laminates. *In: Proceedings of the 9th international conference on composite materials Spain*, vol.6;. p. 51‒8.

Vlot A (1996). Impact loading on fibre metal laminates. Int J Impact Eng;18:291‒307.

Vlot A, Vogelesang LB, Vries T (1999). Towards application of fibre metal laminates in large aircraft. *Aircraft Eng Aerospace Technol;*71:558‒70.

Vogelesang LB (2000), Vlot A. Development of fibre metal laminates for advanced aerospace structures. *J Mater Process Technol*;103:1–5.

Vo TP, Guan ZW, Cantwell WJ, Schleyer GK,(2012). Low-impulse blast behaviour of fibre–metal laminates. *Compos Structures,* 94(3),:954–65.

Vo TP, Guan ZW, Cantwell WJ, Schleyer GK(2013). Modelling of the low-impulse blast behaviour of fibre–metal laminates based on different aluminium alloys, *Composites: Part B* 44, 141–151

Effects of inlet pressure and inlet flow rate on the flow field in a pressure-swirl atomizer

L. X. Zhang[1],* Z. M. Liu[1], H.L. Zheng[2], C. T. Pang[2], Z.Y. Kang[2],T. Zhang[2]

[1]College of Mechanical Engineering and Applied Electronics Technology, Beijing University of Technology, Beijing 100124, China

[2]Aviation Key Laboratory of Science and Technology on Precision Manufacturing, Precision Engineering Institute for Aircraft Industry, Beijing 100076, China

*Corresponding author: lzm@bjut.edu.cn

The dual-fuel circuit pressure-swirl atomizers are numerically studied by the finite volume method with the inside kerosene flowing under the boundary condition of mass-flow-inlet (300L/h~580L/h) and pressure-inlet (1.44MPa~3.04MPa) respectively. It is suggested that the velocity magnitude of the area near the wall of outlets, as significant parts of the internal flow filed of the atomizer, increases nearly linearly with the flow rate and the pressure of inlets. Furthermore, outlet flow rate slowly increases with the increasing inlet pressure under the pressure-inlet boundary condition and the fuel supply pressure of vice orifice increases fast with the increasing inlet flow rate under the mass-flow-rate boundary condition. Therefore, the inlet pressure and flow rate significantly affect the performances of atomizers.

Keywords: Swirl atomizers, inlet flow rate, inlet pressure, atomization performance

0   Introduction

Due to recent trends toward direction of higher thrust to weight ratio, high power to weight ratio, high reliability and low fuel consumption of aero-engine technology, how to improve the performance of the engine is still the focus of research in the field of the aviation. The atomizer of aviation engine is an inevitable necessity component due to fuel into the combustion chamber of the aviation engine in the form of droplets or spray is burned, the characteristics of atomization directly determine the combustion efficiency and stability of aero-engine. This process involves the fuel atomization performed by centrifugal atomizer. There are many types of centrifugal atomizers, which highlight the advantages of large fuel adjustment range, and ensure that better fuel atomization quality can still be obtained at the condition of low volume, meeting the requirements of the stability and complete combustion of the aviation turbine in different altitude. It is believed that in-depth study of the dual centrifugal atomizer to improve engine performance has a crucial role.

During the last decade, a number of numerical and experimental researches by many scholars were conducted on the atomizer considering different perspective and methods. [Jain M et al.(2014)] conducted a detailed experimental study to understand the role of Reynolds number and geometry on the flow coefficient, spray angle and

droplet size in a spray atomizer. [Chen et al.(2010)] experimentally compared the effects of different operating conditions under which primary and secondary fuel line working at the same time on spray angle, indicating that the gap between the two situations is obvious. [Tratnig, A et al. (2010)] carried out an experimental study to evaluate mean diameter of liquid droplet caused by physical properties of working fluid. Although researches on the atomizer has achieved some profound understandings [C. J. Wang Et al. (2009); Zhang Et al. (2003); Chatterjee S et al. (2014); M. Yue Et al. (2003); Y. D. Kong Et al. (2007); Han Z et al. (1997); Datta AFan Y, et al. (2000); Fan Y et al.(2014)], due to small size of atomizer and the complexity of its internal flow field, significant details of the flow field cannot be captured only relying on experiments. And researches regarding the effects of inlet pressure, inlet flow on the flow field inside the atomizer are not deep enough and comprehensive. Henceforth, this study is devoted to investigate the law of dependence of inlet pressure, inlet flow on the flow field inside the atomizer using FLUENT software of computational fluid dynamics.

1 Physical model and calculation methods

1.1 computational model

A schematic of structure of the double line pressure swirl centrifugal atomizer and fuel flow inside the atomizer is shown in Fig. 1. Primary and secondary fuel line is consist of swirl chamber, swirl groove and fuel orifice respectively. When it is working, fuel is accelerated in swirl groove of primary and secondary fuel line and then rotated in swirl chamber and finally sprayed to orifice in the form of rotating film which is spread into a cone by centrifugal force. The fuel becomes relatively small particles under the effect of the air.
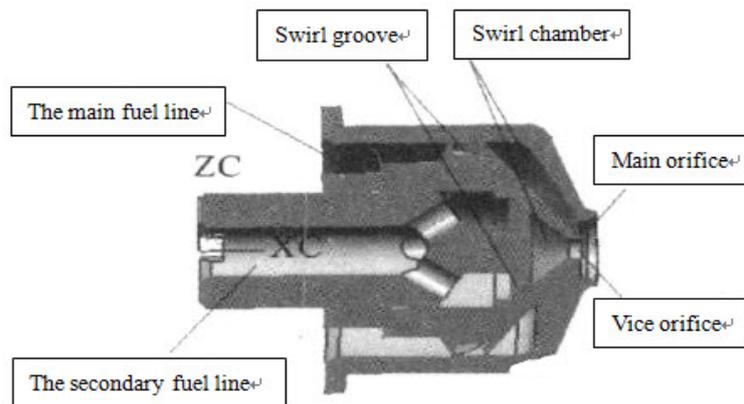
Figure 1. A schematic of cross-section of the atomizer and fuel flow inside it[C. J. Wang Et al. (2009))]

1.2 numerical calculation method

Due to fuel was injected into the atomizer at a relatively high speed at the inlet and then generated high-speed rotation at swirl chamber, the flow belonged to turbulent flow. Therefore, numerical simulation of the flow field of fuel in the atomizer also involves turbulence model, and standard $k-\varepsilon$ turbulence model was selected in this study. To analyze the flow characteristics of this three-dimensional model and simplify the problem, the following assumptions of the flow are adopted: (1)

incompressible; (2) gravity cannot be ignored. The governing equations are expressed as follows:

Continuity equation:

$$\nabla \cdot (\rho \vec{\upsilon}) = 0 \tag{1}$$

Momentum equation:

$$\rho \left[\frac{\partial \vec{\upsilon}}{\partial_t} + (\vec{\upsilon} \cdot \nabla)\vec{\upsilon} \right] = \rho \vec{f} - \nabla p + \mu \nabla^2 \vec{\upsilon} \tag{2}$$

Turbulence equations:

$$\frac{\partial(\rho k)}{\partial_t} + \frac{\partial(\rho k u_i)}{\partial_{x_i}} = \frac{\partial}{\partial_{x_j}} \left[\left(\mu + \frac{\mu_t}{\sigma_k} \right) \frac{\partial k}{\partial_{x_j}} \right] + G_k + G_b - \rho\varepsilon - Y_M + S_k$$

$$\frac{\partial(\rho \varepsilon)}{\partial_t} + \frac{\partial(\rho \varepsilon u_i)}{\partial_{x_i}} = \frac{\partial}{\partial_{x_j}} \left[\left(\mu + \frac{\mu_t}{\sigma_\varepsilon} \right) \frac{\partial \varepsilon}{\partial_{x_j}} \right] + C_{1\varepsilon}\frac{\varepsilon}{k}(G_k + C_{3\varepsilon}G_b) - C_{2\varepsilon}\rho\frac{\varepsilon^2}{k} + S_\varepsilon \tag{3}$$

$C_{1\varepsilon}$、$C_{2\varepsilon}$、$C_{3\varepsilon}$ is empirical constant and $C_{1\varepsilon} = 1.44$、$C_{2\varepsilon} = 1.92$ is the default value in FLUENT. Where ρ is density; μ is kinematic viscosity; $\vec{\upsilon}$ is velocity vector; p is pressure; $\vec{f}$ is combined external force.

1.3 mesh-independent verification

In the current work, CFD (Computational fluid dynamics, CFD) FLUENT software was used for numerical simulation. In order to improve the efficiency and accuracy, structural hexahedral mapped mesh generated by the pre-processing software ICEM was used, as shown in Figure 2. By comparing the amount of each of the grid 1.247 million and 1.181 million, the velocity magnitude of the line y = o through diameter at the outlet is shown in Figure 3, indicating that two lines almost overlap by the mesh-independent verification. Therefore, using 1.181 million mesh could meet the requirements of numerical simulation and analysis.
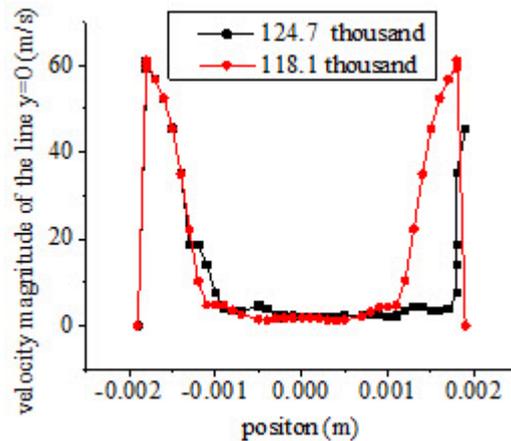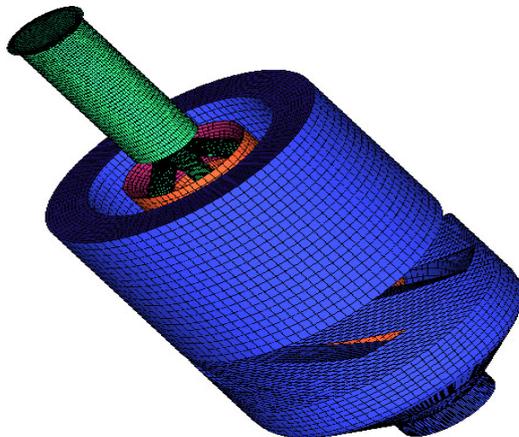
Figure 2. Mesh of calculated domain   Figure 3. Mesh independence verification

1.4 boundary conditions

During the calculation, the 3rd aviation kerosene is used as the working medium. The properties of the 3rd aviation kerosene used in the computation are a density $=800kg/m^3$, the dynamic viscosity coefficient=0.00144Pa.s, the surface tension =0.0268N/s, the boundary conditions are as follows:

（1） The inlet pressure is 1.44, 1.84, 2.04, 2.64, 3.04Mpa respectively;

（2） The inlet flow rate is 300, 370, 440, 510, 580 L/h respectively;

（3） Total Guage pressure is 0MPa at the outlet;

2   Results and discussion

2.1 local flow field (pressure, velocity distribution) inside atomizer

Internal filed of the atomizer under the condition of primary and secondary fuel supply pressure as the same 370MPa was numerically calculated. The pressure and velocity distribution is shown in Figure 4 wherein Figure (a), Figure(b) displays the contours of the pressure and velocity in plane of Z=0 respectively. It can clearly be seen from Fig. 4 that the pressure has an apparent gradient in swirl chamber of primary and secondary fuel line, furthermore, due to the large pressure gradient in the secondary fuel line, there is an obvious area of low pressure in the center of swirl chamber. Correspondingly, after the fuel inflowing the inlet, velocity of the fuel was gradually increased, and increased rapidly after flowing into swirl groove and then fuel generated the acceleration and rotation in swirl chamber, finally was sprayed in a cone under the centrifugal force. The velocity maldistribution of the fuel in the swirling groove is dedicated to a sudden contraction of flow channel and its irregular structure. The geometric structure of the swirl chamber is gradually tapered, thus velocity magnitude of the fuel increases rapidly, reaching a maximum at the outlet, which provides the necessary conditions for the fuel in form of droplets into the combustion chamber.
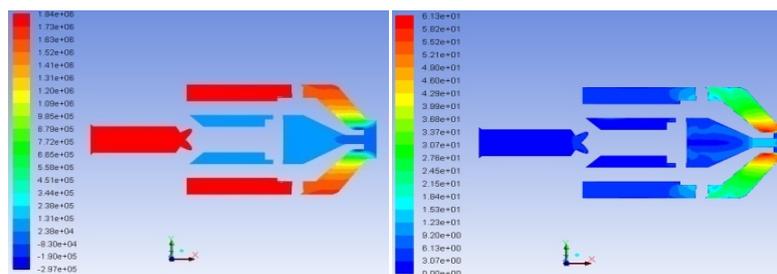
(a) Contours of pressure   (b) Contours of velocity

Figure 4. Schematic of flow field in the atomizer

2.2 Effect of inlet pressure on average velocity of annular area at outlet and outflow rate .

Jet speed at the outlet where near the central area of the outlet is air and near the edge of the annular area of the outlet is the fuel directly determines the fuel atomization quality. The average velocity of the area was extracted to explore the dependence of fuel inlet pressure and inlet flow rate on it. Figure 5 (a) is the curve of the effect of

inlet pressure on average velocity of annular area at outlet. It shows that, when the inlet pressure is 1.44, 1.84, 2.24, 2.64, 3.04Mpa, the average velocity in the annular area is 35.1, 39.5, 43.6, 47.6, 51.4 m/s correspondingly. When the inlet pressure equally spaced (0.4Mpa) increases, the velocity increases 4.4, 4.1, 4.0, 3.8 m/s separately. Namely as the inlet pressure (1.44-3.04Mpa) increases, the velocity magnitude of annular area at outlet is almost linearly increased, however the growth rate of velocity slows gradually. Therefore, in the situation of sufficient fuel supply and complete combustion, increasing inlet pressure of atomizer, the more fuel into the combustion chamber, greater combustion power can be generated. However, increasing the inlet pressure is not completely converted to an increase of the average velocity. Accordingly, when design the atomizer, a fact that with increased pressure (a range), the increasing trend of combustion power became slowly should be fully considered.

Outlet flow rate is one of the most critical indicators of the performance of the atomizer. Exploring the influence of geometric parameters or different operating conditions (inlet pressure) on the outflow rate has an important significance to guide the design and development of the atomizer. This paper focused on the effects of inlet pressure on the outlet flow rate, as shown in Figure 5 (b). It apparently demonstrates that when the inlet pressure is1.44, 1.84, 2.24, 2.64, 3.04Mpa, outlet flow rate is 0.136, 0.153, 0.169, 0.183, 0.196 kg/s respectively. When the inlet pressure equivalently spaced (0.4Mpa) increases, the outlet flow rate increases 0.017, 0.016, 0.014, 0.013 kg/s accordingly. That is, as the inlet pressure (1.44-3.04Mpa) increases, outflow rate is almost linearly increased. but a trend of the increase rate gradually becoming slower has appeared. It is proved once again that a fact that with inlet pressure (a range)increased, the trend of increase of burning power becoming slower should be took full account when design atomizer.

(a) Velocity of annular area of the orifice changes with inlet pressure

(b) The outflow rate changes with inlet pressure
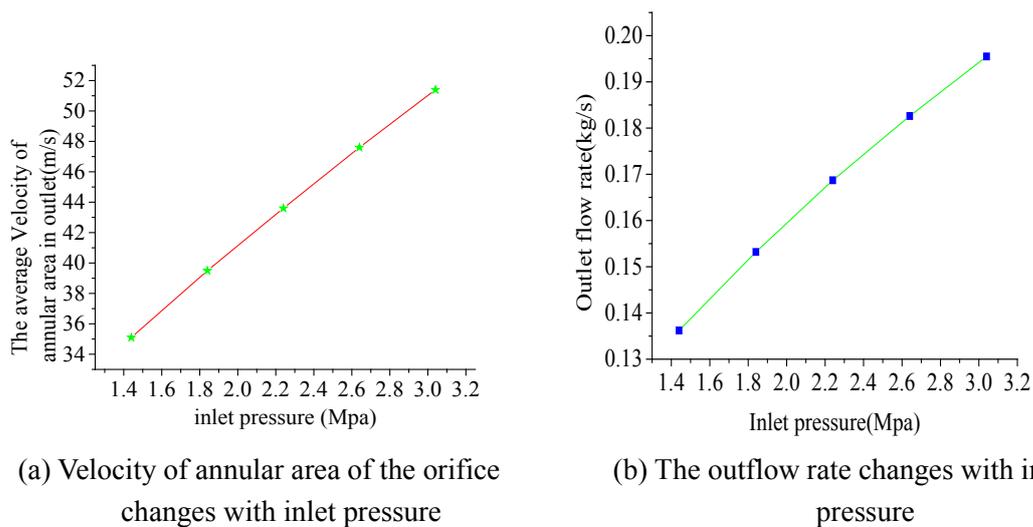
Figure 5. The effect of inlet pressure on the local flow field

2.3 Effect of inlet flow rate on the average velocity of annular area of the outlet and fuel supply pressure of the secondary line

As described in section 2.2, studying the average velocity of annular area of the outlet has great significance. The law of the impact of the inlet flow rate on the velocity is shown in Fig 6(a). Seen from the figure, when the flow rate of the inlet is 300, 370, 440, 510, 580L/h, the average velocity of annular area of the outlet is 14.8, 18.7, 22.8, 26.6, 32.5 m/s separately. While the inlet flow rate equally spaced (70L/h) increases, the velocity increases 3.9, 4.1, 4.8, 5.9 m/s respectively. That is with the increase of the inlet flow rate (300-510L/h), the velocity of annular area of the outlet is gradually increased, and the increase rate becomes gradually faster. Consequently, in condition of combustion chamber is large enough and fuel combustion is compete, compared to adjust the inlet pressure, adjusting inlet flow rate (in a range) can remarkably enhance the combustion efficiency .

Fuel supply pressure is one of the indicators of the performance evaluating the atomizer, maximization of fuel supply efficiency in the atomizer is one of the fundamental purposes of numerous researches. The relationship between flow rate and supply pressure of the fuel in the secondary line was considered in this paper as shown in Figure 6(b).It is shown that when the inlet flow rate is 300, 370, 440, 510, 580L/h, supply pressure of the fuel of the secondary line is 5.5, 8.4, 11.9, 16.2, 20.9Mpa respectively. While the inlet flow equally spaced (70 L/h) increases, supply pressure of the fuel of the secondary line increases 2.9, 3.5, 4.3, 4.7Mpa respectively. That is, with inlet flow rate (300-510L/h) increased, the supply pressure of the fuel of the secondary line increases, and the increase rate gradually becomes faster. It can be explained that the fuel in the secondary domain, the more fuel supplied to the inlet passage (300-510L/h) ,leading to area of air contact with fuel is larger, the flow resistance is increased accordingly. It is seen that, when the primary and secondary fuel line operate simultaneously, a consideration that the increasing the flow rate (a range) makes the resistance of secondary fuel line increase and the effect that increase rate becomes faster should be premeditated in design of the atomizer.
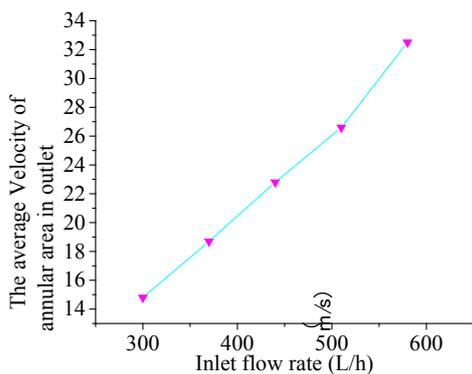
(a) Average velocity of annular area in the outlet with inlet flow rate changes
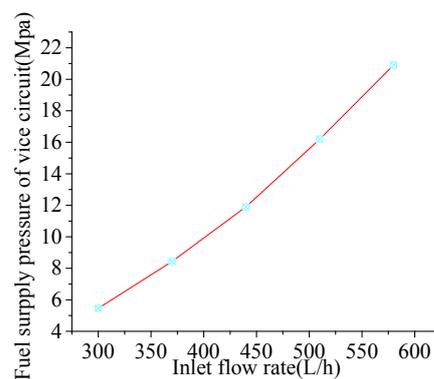
(b) Supply pressure of the fuel in the secondary line changes with inlet flow rate

Figure 6. The effect of inlet flow rate on the local flow field

3 Conclusions

（1） With inlet pressure (1.44-3.04Mpa) increased, the average velocity of annular area of the outlet almost linearly increases, slightly emerging a trend that as the pressure increases, the increase rate will be slower. But the velocity sharply increases with the increasing inlet flow rate (300-510L/h), while the increase rate gradually goes up. Compared to adjust the inlet pressure (a range) , changing the inlet flow rate(a range) can more validly enhance combustion efficiency in the time of atomizer design.

（2） With the inlet pressure (1.44-3.04Mpa) increased, the outlet flow rate almost linearly increases, effect of air resistance on the fuel also reveals a relationship of linear increase, but increase of the flow rate tends to be slower. When improve combustion efficiency of the atomizer, this effect that increasing pressure (a range), the outlet flow increases correspondingly, but the increase rate slows down should be given full consideration.

（3） With the increasing inlet flow rate, the fuel supply pressure of secondary line increases, but the increase rate gradually becomes faster. The more fuel supplied to secondary passage, the greater chance of fuel contacting with air is, and the air resistance on the fuel becomes larger. Designing a atomizer especially primary and secondary passage operates at the same time, it is concentrated on the fact that with inlet flow rate increased, resistance of secondary line is increased and the growth rate becomes faster.

References

Jain M, John B, Iyer K N, et al. Characterization of the full cone pressure swirl spray atomizers for the nuclear reactor containment spray system[J]. Nuclear Engineering and Design, 2014, 273: 131-142.

J. Chen Et al. Experimental investigation of spray characteristics of a double line pressure-swirl atomizer [J]. Journal of Aerospace Power, 2010 (4): 774-779.

Tratnig A, Brenn G. Drop size spectra in sprays from pressure-swirl atomizers[J]. International Journal of Multiphase Flow, 2010, 36(5): 349-363.

C. J. Wang Et al. Numerical study of the cone angle of dual centrifugal atomization atomizer [J]. Energy Conservation, 2009, 28(2): 11-13.

Z. Zhang Et al. Experimental study on spray characteristics of dual-orifice pressure-swirl atomizer[J]. Journal of Engineering Thermophysics, 2003, 24(1): 153-156.

Chatterjee S, Das M, Mukhopadhyay A, et al. Effect of Geometric Variations on the Spray Dynamics of an Annular Fuel Sheet in a Hybrid Atomizer[J]. Atomization and Sprays, 2014.

M. Yue Et al. Numerical study of two-phase flow in atomizer [J]. Journal of Engineering Thermophysics, 2003, 24(5): 888-890.

Y. D. Kong Et al. Numerical Simulation of Spray Angle in Aero engine Combustor Swirl Atomizer [J]. Computer Simulation, 2007, 24(10): 45-47.

Han Z, Parrish S, Farrell P V, et al. Modeling atomization processes of pressure-swirl hollow-cone fuel sprays[J]. Atomization and Sprays, 1997, 7(6).

Datta A, Som S K. Numerical prediction of air core diameter, coefficient of discharge and spray cone angle of a swirl spray pressure atomizer[J]. International Journal of Heat and Fluid Flow, 2000, 21(4): 412-419.

Fan Y, Hashimoto N, Nishida H, et al. Spray characterization of an air-assist pressure-swirl atomizer injecting high-viscosity Jatropha fuels[J]. Fuel, 2014, 121: 271-283.

Parameter estimation approach for particle flow model of rockfill materials using response surface method

*Shouju Li[1], De Li[1], Lijuan Cao[2], Zichang Shangguan[3]

[1] State Key Laboratory of Structural Analysis for Industrial Equipment, Dalian University of Technology, Dalian, China
[2] School of Mechanical and Power Engineering, Dalian Ocean University, Dalian, 116024, China
[3] Institute of Marine and Civil Engineering, Dalian Ocean University, Dalian, 116024, China

*Corresponding author: lishouju@dlut.edu.cn

Abstract

Particle flow code (PFC) is widely used to model deformation and stress states of rockfill materials. The accuracy of numerical modelling with PFC is dependent upon the model parameter values. How to accurately determine model parameters remains one of the main challenges. In order to determine model parameters of particle flow model of rockfill materials, some triaxial compression experiments are performed, and the inversion procedure of model parameters based on response surface method is proposed. Parameters of particle flow model of rockfill materials are determined according to observed data in triaxial compression tests for rockfill materials. The investigation shows that the normal stiffness, tangent stiffness and friction coefficient of rockfill materials will slightly increase with increase of confining pressure in triaxial compression tests. The experiments in laboratory show that the proposed inversion procedure behaves higher computing efficiency and the forecasted stress-strain relations agree well with observed values.

Keywords: micromechanical model, rockfill materials, parameter inversion, triaxial compression tests, response surface method

1. Introduction

Rockfill materials are widely used to construct dams. The deformation characteristics of rockfill materials commonly are numerically simulated by distinct element method and PFC software. The accuracy of numerical modelling with PFC is dependent upon the model parameter values. How to accurately determine model parameters remains one of the main challenges. Some researchers have tried to determine the micromechanical model parameters of granular materials experimentally. Masson performed a set of distinct element simulations of the filling and the discharge of a plane rectangular silo with variable values of particle mechanical parameters. The analysis of the influence of friction and stiffness of contacts showed that these parameters played a major role in the flow kinematics and in the stress field during filling and discharge processes [Masson and Martinez (2000)]. Bagherzadeh developed a novel approach for the two-dimensional numerical simulation of the phenomenon in rockfill using combined DEM and FEM. All particles were simulated by the discrete element method as an assembly and after each step of DEM analysis, each particle was separately modeled by FEM to determine its possible breakage [Bagherzadeh et al. (2011)]. Hosseininia presented a model to simulate the breakage of two-dimensional polygon-shaped particles. In the model, each uniform (uncracked) particle was replaced with smaller inter-connected sub-particles which are bonded with each other [Hosseininia and Mirghasemi (2006)].

Renzo performed a mathematical modification of Mindlin's tangential solution and demonstrated formally its advantages with respect to the commonly used model [Renzo and Maio (2005)]. Coetzee presented a method for determining the parameters of cohesionless granular material. The particle size and density were directly measured and modeled. The particle shapes were modeled using two to four spheres clumped together. The remaining unknown parameter values were determined using confined compression tests and angle of repose tests [Coetzee et al. (2010)]. Koyama proposed a numerical procedure to determine the equivalent micro-mechanical properties of intact rocks using a stochastic representative elementary volume (REV) concept and a particle mechanics approach. More than 200 models were generated in square regions with side lengths varying from 1 to 10 cm, using the Monte Carlo simulation technique [Koyama and Jing (2007)]. Kulatilake performed laboratory experiments and numerical simulations to study the behavior of jointed blocks of model material under uniaxial loading. The effect of joint geometry parameters on the uniaxial compressive strength of jointed blocks was investigated [Kulatilake et al. (2001)]. Each particle has material parameters (micro-parameters) that influence the particle macro-behaviors. The accuracy of PFC model depends on the micro-parameters of model. How to accurately determine PFC model parameters remains one of the main challenges.

PFC^{2D} models the movement and interaction of circular particles by the distinct element method (DEM), as described by Cundall and Strack (1979). The overall constitutive behavior of a material is simulated in PFC^{2D} by associating a simple constitutive model with each contact. The constitutive model acting at a particular contact consists of three parts: a stiffness model, a slip model, and a bonding model. The stiffness model provides an elastic relation between the contact force and relative displacement. The slip model enforces a relation between shear and normal contact forces such that the two contacting balls may slip relative to one another. The bonding model serves to limit the total normal and shear forces that the contact can carry by enforcing bond-strength limits. González-Montellano performed the experimental to determine values for several of the microscopic properties-the particle density, modulus of elasticity, particle-wall coefficient of restitution, particle-particle coefficient of restitution, and the particle-wall coefficient of friction-of maize grains and olives, required for use in DEM simulations [González-Montellano et al. (2012)]. Yoon developed a new approach for calibrating contact-bonded particle models using 'experimental design' and 'optimization' in uniaxial compression simulation. These were applied to calculate an optimum set of microparameters used in generation of models to be tested in uniaxial compression simulations [Yoon (2007)]. Belheine calibrated the micro-mechanical properties of the numerical material using numerical triaxial tests in order to match the macroscopic response of the real material. Numerical simulations were carried out under the same conditions as the physical experiments. The pre-peak, peak and post-peak behaviors of the numerical material were studied [Belheine et al. (2009)]. Chen investigated the failure mechanism and the limit support pressure of a tunnel face in dry sandy ground by using discrete element method. The contact parameters of the dry sand particles were obtained by calibrating the results of laboratory direct shear tests. A series of three-dimensional DEM models for different ratios of the cover depth to the diameter of the tunnel were then built to simulate the process of tunnel face failure [Chen et al. (2011)]. Deluzarche proposed a methodology to define the resistance of the 2D particles so that the same probability of breaking blocks may be reproduced as in a 3D material. The model used the discrete element code PFC^{2D} and considered breakable clusters of 2D balls. The different parameters were determined from experimental data obtained from

laboratory tests performed on rock blocks [Deluzarche and Cambou (2006)]. Alaei simulated single crushing tests and triaxial tests on the Purulia dam's material to validate the presented model for rockfill material. The obtained results demonstrated the accuracy of the adopted model and the model's capability for considering a rockfill material's strength, deformation and crushing behaviour [Alaei and Mahboubi (2012)]. Even if some procedures has been proposed to determine micromechanical parameters of rockfill materials, the common drawback of these estimating procedures lies in lower fitting and predicting precision. Response surface methodology is a collection of statistical and mathematical techniques useful for developing, improving, and optimizing processes in which a response of interest is influenced by several variables and the objective is to optimize this response. Response surface methodology has been widely applied in inverse solution of soil-water transport model parameters [Saha et al. (2010)], parameter optimization [Muthuvelayudham and Viruthagiri (2010)], nutritional parameter optimization [Kunamneni et al. (2005)]. The aim of the paper is to propose a new procedure for determining PFC model parameters of rockfill materials from triaxial compression tests and to validate effectiveness of proposed inversion approach through experiments in laboratory.

2. Numerical simulations for triaxial compression tests using PFC software

PFC model is based on the simulations of the motion of granular material as separate particles. Using the soft particle approach, each particle contact is modeled with a linear spring both in the contact normal direction and contact tangential direction, as shown in Figure 1. The particles are allowed to overlap and the amount of overlap is used in combination with the spring stiffness to compute the contact force components.
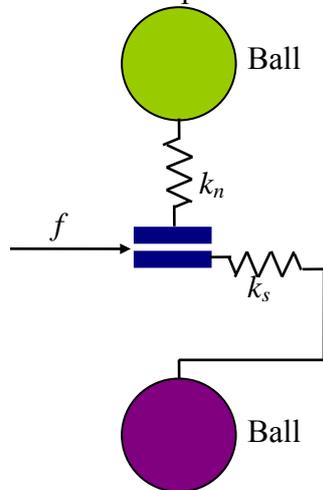
Figure 1. DEM contact model

The normal stiffness of a particle is secant stiffness. The relation between normal force and normal displacement is expressed as follows

$$F_n = k_n U_n \tag{1}$$

Where F_n denotes total normal force, k_n denotes normal stiffness, U_n denotes total normal displacement. The shear stiffness of a particle is a tangent stiffness. The relation between increment of tangent force and increment of tangent displacement is expressed as follows

$$\Delta F_s = -k_s \Delta U_s \tag{2}$$

Where ΔF_s denotes the increment of shear force, k_s denotes tangent stiffness, ΔU_s denotes the increment of shear displacement. The slip model is defined by the friction coefficient at the contact f [dimensionless], where f is taken to be the minimum friction coefficient of the two contacting entities.

In order to determine the model parameters of rockfill materials, some triaxial compression tests of rockfill materials are performed in laboratory. The largest size of rockfill particles is 100mm, as shown in Figure 2. The smallest size is 0.1mm. The diameter of test model is 300mm. The height is 700mm, as shown in Figure 3. Variation of deviatoric stress (principal stress difference: σ_1-σ_3) versus axial strain in triaxial compression test of rockfill materials is depicted in Figure 4. σ_1 is major stress (axial stress), and σ_3 is minor stress (confining pressure). These test data are available for parameter estimation of PFC model of rockfill materials.
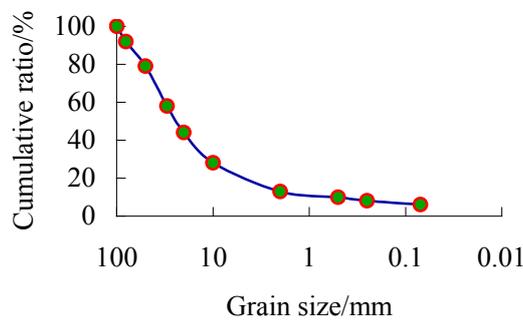
Figure 2. Particle size distribution for rockfill materials
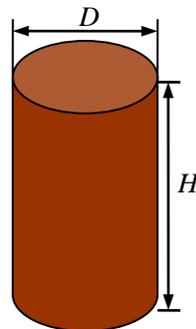
Figure 3. Triaxial compression test of rockfill materials
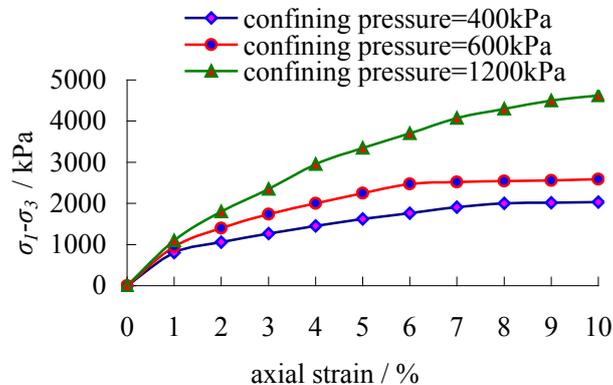
Figure 4. Variation of deviatoric stress versus axial strain in triaxial compression test of rockfill materials

238
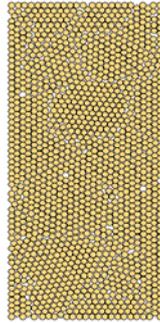
Figure 5. Simplified PFC^{2D} model of triaxial compression test of rockfill materials

After taking into account of symmetrical characteristic of triaxial compression test model, PFC model is simplified into two dimensions for simulating triaxial compression test of rockfill materials, as shown in Figure 5. The radius of rockfill particle in PFC^{2D} model is approached as 20mm according to the average radius of rockfill particle. Influences of normal stiffness, tangent stiffness and friction coefficient of rockfill materials on stress-strain relation are simulated with PFC model, as shown in Figure 6, 7 and 8.
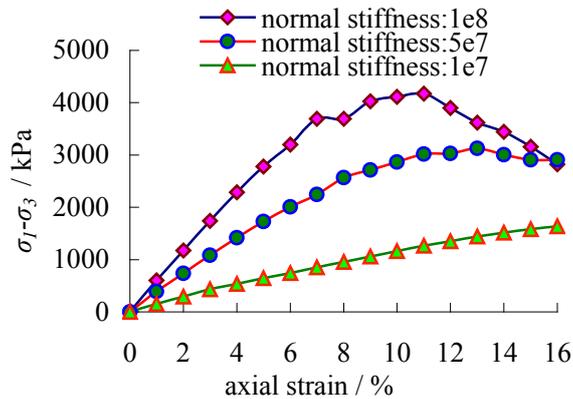
Figure 6. Influence of normal stiffness of rockfill materials on stress-strain relations (Confining pressure 1200kPa)
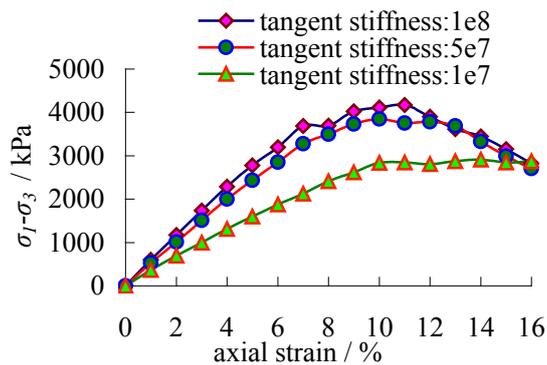
Figure 7. Influence of tangent stiffness of rockfill materials on stress-strain relations (Confining pressure 1200kPa)
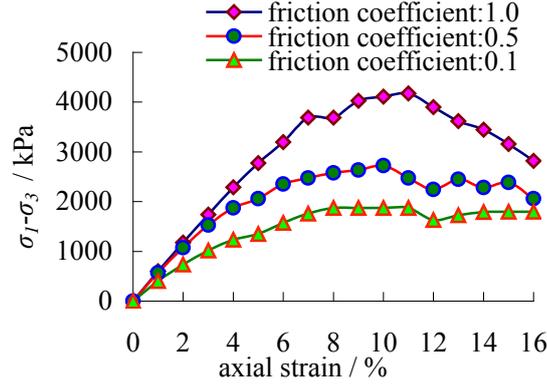
Figure 8. Influence of friction coefficient of rockfill materials on stress-strain relations (Confining pressure 1200kPa)

3. Parameter inversion procedures for PFC model using response surface method

Based on the response surface method, the relation between unknown PFC model parameters of rockfill materials and deviatoric stress in triaxial compression test is approached as [Rosa et al. (2009); Bas and Boyaci (2007)].

$$s_k(\mathbf{\bar{x}}) = a + \sum_{i=1}^{3} b_i \bar{x}_i + \sum_{i=1}^{3} c_i \bar{x}_i^2 \tag{3}$$

Where $s_k(\mathbf{\bar{x}})$ is principal stress difference (σ_1-σ_3) at loading step k, a, b_i and c_i are unknown coefficients, $\mathbf{\bar{x}}$ is unknown model parameter vector after dimensionless procedure.

$$\mathbf{\bar{x}} = \{\bar{x}_1, \bar{x}_2, \bar{x}_3\}^T = \{\bar{k}_n, \bar{k}_s, \bar{f}\}^T \tag{4}$$

$$\bar{k}_n = \frac{k_n}{\tilde{k}_n}, \bar{k}_s = \frac{k_s}{\tilde{k}_s}, \bar{f} = \frac{f}{\tilde{f}} \tag{5}$$

Where $\tilde{k}_n$, $\tilde{k}_s$ and $\tilde{f}$ denote initial evaluating values of model parameters according to prior to information.

Taking the first loading step as an example, the left items of following equations can be calculated by simulations using PFC^{2D} software under the given model parameter combinations

$$s_1^1(\mathbf{\bar{x}}) = s(\bar{k}_n, \bar{k}_s, \bar{f}) \tag{6}$$

$$s_1^2(\mathbf{\bar{x}}) = s(\bar{k}_n + \Delta\bar{k}_n, \bar{k}_s, \bar{f}) \tag{7}$$

$$s_1^3(\mathbf{\bar{x}}) = s(\bar{k}_n - \Delta\bar{k}_n, \bar{k}_s, \bar{f}) \tag{8}$$

$$s_1^4(\mathbf{\bar{x}}) = s(\bar{k}_n, \bar{k}_s + \Delta\bar{k}_s, \bar{f}) \tag{9}$$

$$s_1^5(\mathbf{\bar{x}}) = s(\bar{k}_n, \bar{k}_s - \Delta\bar{k}_s, \bar{f}) \tag{10}$$

$$s_1^6(\mathbf{\bar{x}}) = s(\bar{k}_n, \bar{k}_s, \bar{f} + \Delta\bar{f}) \tag{11}$$

$$s_1^7(\overline{\mathbf{x}}) = s(\overline{k}_n, \overline{k}_s, \overline{f} - \Delta\overline{f}) \qquad (12)$$

Where $\Delta\overline{k}_n = 0.1, \Delta\overline{k}_s = 0.1, \Delta\overline{f} = 0.1, s_1^i$ denotes principal stress difference computed in the first loading case under i-th parameter combination, which is computed by using PFC^{2D} software. There exist 7 unknown coefficients and 7 equations. So, the 7 unknown coefficients in response surface functions in the first loading case can be determined by solving linear equation set with MATLAB software. The unknown coefficients in response surface functions for other loading steps may be deduced by analogy.

Initial evaluating parameter values of PFC model of rockfill materials are listed in Table 1, where ρ denotes particle density, which is a known constant, σ_c denotes confining pressure in triaxial compression test.

Table 1. Initial evaluating parameter values of PFC model of rockfill materials

$\tilde{k}_n$ / N/m	$\tilde{k}_s$ / N/m	$\tilde{f}$	ρ / kg/m^3	σ_c / kPa
8.0e7	8.0e7	0.9	2800	400
1.2e8	1.2e8	0.9	2800	600
1.4e8	1.4e8	1.0	2800	1200

After performing a lot of numerical simulations for triaxial compression test with PFC software, the coefficients of response surface functions for every load step under different confining pressure are computed and listed in Table 2, 3 and 4.

Table 2. Coefficients of response surface functions for every load step (Confining pressure 400kPa)

Load step	a	b_1	b_2	b_3	c_1	c_2	c_3
1	-26.0	-1505.0	870.0	1230.0	850.0	-400.0	-600.0
2	-1255.0	-2800.0	2625.0	3515.0	1600.0	-1250.0	-1650.0
3	-120.0	-3325.0	685.0	4040.0	1850.0	-250.0	-1800.0
4	1793.0	2870.0	-5005.0	-1260.0	-600.0	2550.0	900.0
5	6447.0	-12430.0	1370.0	-155.0	6500.0	-700.0	450.0
6	9443.0	-12030.0	-4455.0	-720.0	6400.0	2250.0	800.0
7	-4291.0	3205.0	5305.0	1240.0	-950.0	-2750.0	100.0
8	-20630.0	25270.0	5960.0	11940.0	-12000.0	-3100.0	-5400.0
9	-24579.0	31465.0	-2645.0	20555.0	-15150.0	1250.0	-8850.0
10	-54751.0	52535.0	21190.0	36745.0	-26050.0	-10600.0	-17050.0

Table 3. Coefficients of response surface functions for every load step (Confining pressure 600kPa)

Load step	a	b_1	b_2	b_3	c_1	c_2	c_3
1	-3733.0	5510.0	1670.0	1090.0	-2600.0	-800.0	-500.0
2	-2702.0	4855.0	1130.0	875.0	-2250.0	-500.0	-250.0
3	-3957.0	7280.0	1170.0	1445.0	-3500.0	-500.0	-350.0
4	-13691.0	19685.0	945.0	9075.0	-9750.0	-250.0	-4050.0
5	-6514.0	19105.0	-2110.0	-1610.0	-9150.0	1100.0	1400.0
6	3683.0	26660.0	-13130.0	-17845.0	-13000.0	6500.0	9550.0
7	-9866.0	42955.0	-9960.0	-10210.0	-21150.0	5000.0	5900.0
8	-12339.0	59555.0	-15075.0	-15865.0	-29650.0	7250.0	8850.0
9	1064.0	47920.0	-26850.0	-19710.0	-23700.0	13000.0	10900.0
10	-61362.0	101415.0	-10325.0	34545.0	-50450.0	4850.0	-15950.0

Table 4. Coefficients of response surface functions for every load step (Confining pressure 1200kPa)

Load step	a	b_1	b_2	b_3	c_1	c_2	c_3
1	457.0	-625.0	835.0	-265.0	550.0	-350.0	150.0
2	-234.0	160.0	1090.0	850.0	300.0	-400.0	-300.0
3	1150.0	-2275.0	620.0	1380.0	1850.0	-100.0	-500.0
4	48694.0	-98960.0	-3070.0	2385.0	52700.0	1800.0	-850.0
5	-3356.0	4500.0	2495.0	1905.0	-1000.0	-1050.0	-250.0
6	-10869	15845.0	-2075.0	11170.0	-6750.0	1150.0	-4800.0
7	-3229.0	9035.0	4210.0	-4680.0	-3050.0	-1900.0	3600.0
8	29202.0	-5755.0	-17810.0	-30510.0	4050.0	8700.0	16200.0
9	-89748	75030.0	52945.0	53025.0	-36500.0	-25750.0	-24550
10	-94191	126100.0	16225.0	48875.0	-62400.0	-8150.0	-22050
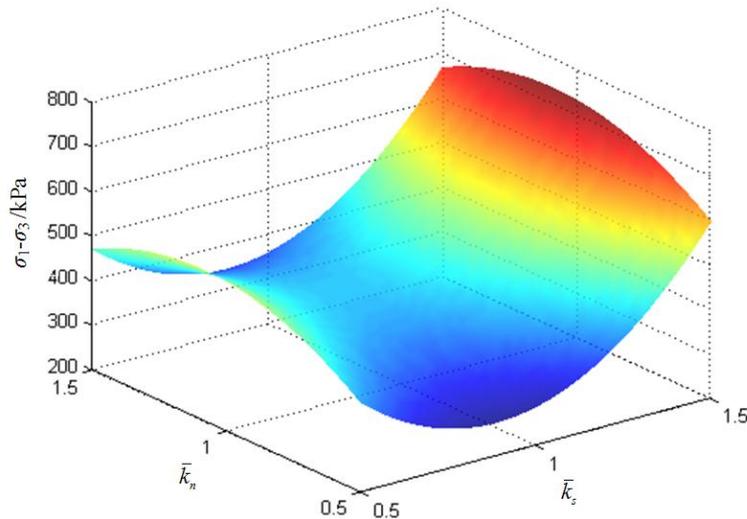
Figure 9. Response surface of deviatoric stress (*f*=1.0, Confining pressure=400kPa)

The objective function of estimating PFC model parameters for rockfill materials is defined as Root Mean Square (RMS)

$$\min J = \sqrt{\frac{1}{N}\sum_{k=1}^{N}[s_k(\overline{\mathbf{x}}) - s_k^m]^2} \qquad (13)$$

Where *J* is objective function of parameter inversion, s_k^m is the observed principal stress differences for the-*k* loading step in triaxial compression tests of rockfill materials, *N* is the number of loading step. Equation (13) is an optimization problem with non-constrained conditions and can be solved with some optimization algorithms. So, the inverse problem for parameter estimation is transformed into optimization problem and can be solved with BFGS optimization algorithm [Broyden (1970); Andonegi et al. (2011)]. According to observed data in triaxial compression tests of rockfill materials, as shown in Figure 4, and response surface functions, as shown in Table 2, 3, and 4, as well as BFGS optimization algorithm, unknown PFC model parameters of rockfill materials are identified and listed in Table 5.

Table 5. Identified PFC model parameters of rockfill materials

k_n /MN/m	k_s /MN/m	f	σ_c / kPa
86.504	83.224	0.9175	400
128.88	124.45	0.9192	600
144.31	137.12	1.0389	1200

From Table 5, it will be found that the normal stiffness is slightly larger than tangent stiffness and nearly equal to tangent stiffness. Based on identified PFC model parameters of rockfill materials, variations of deviatoric stress versus axial strain in triaxial compression test of rockfill materials under different confining pressure are simulated again. The differences between observed deviatoric stresses and predicted ones are depicted in Figure 10, 11 and 12. From these figures, we can find that predicted values by PFC model agree well with the experimental ones.
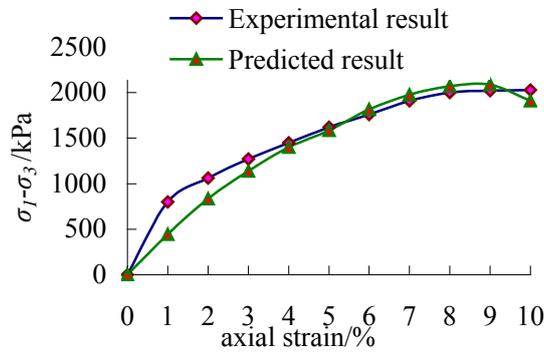
Figure 10. Comparison between experimental values and predicted ones in triaxial compression test (Confining pressure 400 kPa)
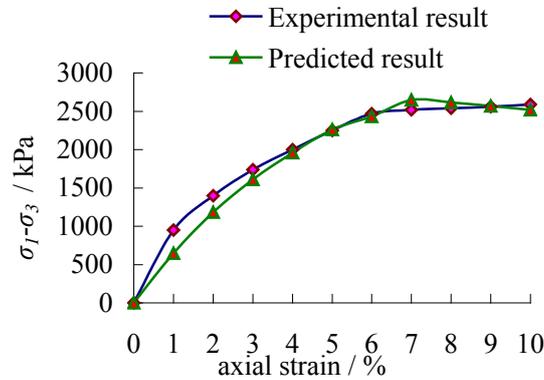
Figure 11. Comparison between experimental values and predicted ones in triaxial compression test (Confining pressure 600 kPa)
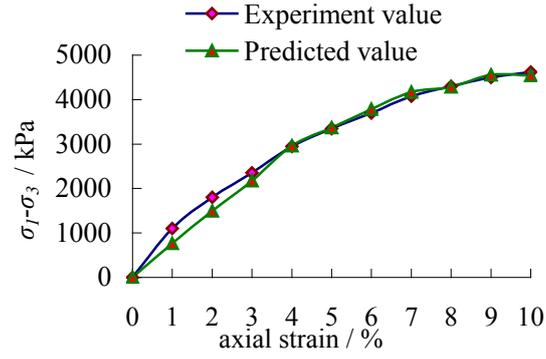
Figure 12. Comparison between experimental values and predicted ones in triaxial compression test (Confining pressure 1200 kPa)

The further investigation facts that the normal stiffness, tangent stiffness and friction coefficient will increase with the increase of confining pressure σ_3, as shown in Table 5 and in Figure 13 and 14. The relations between constitutive model parameters of particles and confining pressures can be expressed as follows

$$k_n = \alpha(\frac{\sigma_3}{P_a})^\beta \tag{14}$$

$$k_s = \psi(\frac{\sigma_3}{P_a})^\zeta \tag{15}$$

$$f = m(\frac{\sigma_3}{P_a})^n \tag{16}$$

Where α, ψ and m are coefficients of empirical equations, β, ζ and n are exponents of empirical equations, P_a is atmosphere pressure, P_a =100kPa. After regression analysis, the coefficients and exponents of empirical equations are determined as follows: α=51.1, β =0.437, ψ =50.5, ζ=0.423, m=0.763, n=0.120.
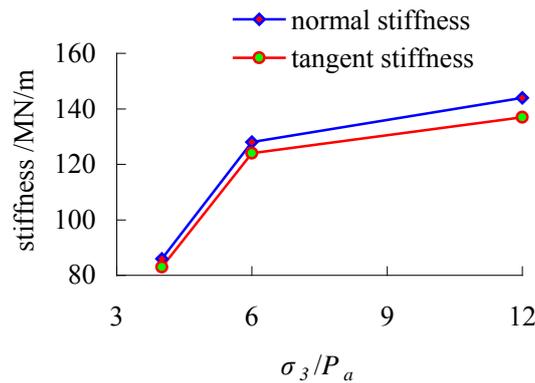
Figure 13. Variation of normal stiffness and tangent stiffness versus confining pressure
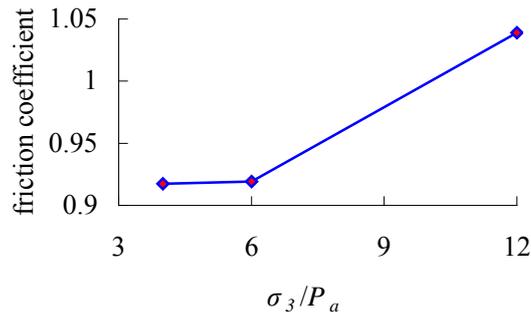
Figure 14. Variation of friction coefficient versus confining pressure

4. Conclusions

1) A new inversion procedure is proposed to determine PFC model parameters of rockfill materials. Based on the response surface method, the relation between unknown PFC model parameters of rockfill materials and deviatoric stress in triaxial compression test is approached. By comparing forecasted stress-strain curves with observed ones, the effectiveness of proposed model parameter inversion procedure is validated by experiments in laboratory.

2) The investigation facts that the normal stiffness is slightly larger than tangent stiffness and nearly equal to tangent stiffness. The normal stiffness, tangent stiffness and friction coefficient will increase with the increase of confining pressure.

3) The nonlinear relations between constitutive model parameters of particles and confining pressures are presented. But the expressions and its coefficients only supply references because the number of samples is not large enough. How to determine pro-fractured mechanical characteristics of rockfill materials should be further investigated in the future.

5. Acknowledgement

References

Masson, S. and Martinez, J. (2000) Effect of particle mechanical properties on silo flow and stresses from distinct element simulations, Powder Technology **109**, 164–178.

Bagherzadeh Kh, A., Mirghasemi, A. A. and Mohammadi, S. (2011) Numerical simulation of particle breakage of angular particles using combined DEM and FEM, Powder Technology **205**, 15–29.

Seyedi Hosseininia, E. and Mirghasemi, A. A. (2006) Numerical simulation of breakage of two-dimensional polygon-shaped particles using discrete element method, Powder Technology **166**, 100–112.

Di Renzo, A. and Di Maio, F. P. (2005) An improved integral non-linear model for the contact of particles in distinct element simulations, Chemical Engineering Science **60**, 1303–1312.

Coetzee, C. J., Els, D. N. J. and Dymond, G. F. (2010) Discrete element parameter calibration and the modelling of dragline bucket filling, Journal of Terramechanics **47**, 33–44.

Koyama, T. and Jing, L. (2007) Effects of model scale and particle size on micro-mechanical properties and failure processes of rocks-A particle mechanics approach, Engineering Analysis with Boundary Elements **31**, 458–472.

Kulatilake, P. H. S. W., Malama, B. and Wang, J. (2001) Physical and particle flow modeling of jointed rock block behavior under uniaxial loading, International Journal of Rock Mechanics and Mining Sciences **38**, 641–657.

González-Montellano, C., Fuentes, J. M., Ayuga-Téllez, E. and Ayuga, F. (2012) Determination of the mechanical properties of maize grains and olives required for use in DEM simulations, Journal of Food Engineering **111**, 553–562.

Yoon, J. (2007) Application of experimental design and optimization to PFC model calibration in uniaxial compression simulation, International Journal of Rock Mechanics and Mining Sciences **44**, 871–889.

Belheine, N., Plassiard, J. P., Donzé, F. V., Darve, F. and Seridi, A. (2009) Numerical simulation of drained triaxial test using 3D discrete element modeling, Computers and Geotechnics **36**, 320–331.

Chen, R. P., Tang, L. J., Ling, D. S. and Chen, Y. M. (2011) Face stability analysis of shallow shield tunnels in dry sandy ground using the discrete element method, Computers and Geotechnics **38**, 187–195.

Deluzarche, R. and Cambou, B. (2006) Discrete numerical modelling of rockfill dams, International Journal for Numerical and Analytical Methods in Geomechanics **30**, 1075-1096.

Alaei, E. and Mahboubi, A. (2012) A discrete model for simulating shear strength and deformation behaviour of rockfill material, considering the particle breakage phenomenon, Granular Matter **14**, 707–717.

Saha, R., Upadhyaya, S. K., Wallender, W. W. and Slaughter, D. C. (2010) Inverse solution of soil-water transport model parameters using response surface methodology, Transactions of the ASABE **53**, 1137–1145.

Muthuvelayudham, R. and Viruthagiri, T. (2010) Application of central composite design based response surface methodology in parameter optimization and on cellulase production using agricultural waste, International Journal of Chemical and Biological Engineering **3**, 97–104.

Kunamneni, A., Kumar, K. S. and Singh, S. (2005) Response surface methodological approach to optimize the nutritional parameters for enhanced production of-amylase in solid state fermentation by Thermomyces lanuginosus, African Journal of Biotechnology **4**, 708–716.

Rosa, J. L., Robin, A., Silva, M. B., Baldan, C. A. and Peres, M. P. (2009) Electrodeposition of copper on titanium wires: Taguchi experimental design approach, Journal of Materials Processing Technology **209**, 1181–1188.

Baş, D. and Boyacı, İ. H. (2007) Modeling and optimization I: Usability of response surface methodology, Journal of Food Engineering **78**, 836–845.

Broyden, C. G. (1970) The convergence of a class of double-rank minimization algorithms 1. general considerations, IMA Journal of Applied Mathematics **6**, 76–90.

Andonegi, E., Fernandes, J. A., Quincoces, I., Irigoien, X., Uriarte, A., Pérez, A., ... and Stefánsson, G. (2011) The potential use of a Gadget model to predict stock responses to climate change in combination with Bayesian networks: the case of Bay of Biscay anchovy, ICES Journal of Marine Science **68**, 1257–1269.

Alternately Moving Road Method for the FEM/DEM Simulation of Tire-Sand Interactions

C.L. Zhao, *M.Y. Zang

School of Mechanical & Automotive Engineering, SCUT, Guangzhou 510640, China

*Corresponding author: myzang@scut.edu.cn

Abstract

The Finite-Discrete Element Method (FEM/DEM) is a promising tool to analyze the tire-sand interactions. However, it usually requires a long driving distance to investigate the tire running behavior on the sand road which will lead to a large-scale simulation model. The Alternately Moving Road Method is proposed in this study to reduce the size of the simulation model: the sand road which has been passed over by the tire is removed and same size of road specimen is laid in front of the tire simultaneously. This method possesses the ability to keep the road scale constant and acceptable in the simulation of arbitrary length sand roads. Numerical model of tire driving on sand road is established to verify the feasibility of the method. And the simulation results are compared with the current experimental results to validate the feasibility and effectiveness of the method.

Keywords: Alternately Moving Road Method, Tire-sand interactions, FEM/DEM, Running behavior

1 Introduction

The in-depth study of the tire-sand interactions is significant to the design and parameter match of off-road vehicles. Recently, as the rapidly developed of computer technology, numerical method becomes an efficient and economic approach for the research of this field. The Finite Element Method (FEM) and the Discrete Element Method (DEM) are two frequently used methods. The FEM, which is based on the continuous theory, possesses the advantage of describing the tire characteristics [BIRIS et al., (2011); Cuong et al., (2013); González Cueto et al., (2013); Li and Schindler, (2013); Moslem and Hossein, (2014)], and the DEM is appropriate to model the granular futures such as large displacement of the sand [Khot et al., (2007); Knuth et al., (2012); Nakashima et al., (2007); Smith and Peng, (2013); Zhang et al., (2012)]. Thus, it is quite nature to use the DEM and the FEM together (FEM/DEM) to taken into account the advantages of the two method in the investigation of tire-sand interactions, where the sand can be modeled by using the DEM and the tire model can be discretized into finite elements.

The FEM/DEM method has been used by David [David et al., (2001)], Nakashima [Nakashima and Oida, (2004); Nakashima et al., (2008), (2009)] and Zhao [Zhao and Zang, (2014a), (2014b)] to investigate the tire-soil interactions and proved to be an effective tool. In these literatures, the discrete elements were contact with each other and with the finite elements, and the contact detection was the most time consuming part. Although various kinds of contact detection algorithms were applied by researchers to improve the computing efficiency, it is still the bottleneck problem for the application of this method due to large amount of discrete elements, especially for a longish test road. In this study, the 3D FEM/DEM is applied to investigate the tire-sand interactions and the Alternately Moving Road Method (AMRM) is proposed to keep a constant number of discrete elements for the simulation of arbitrary length test

roads.

The structure of the paper is as follows: Sect. 2 briefly introduces the basis for the application of the FEM/DEM to the investigation of the tire running behavior on sand terrain. Sect. 3 illustrates the principle of the AMRM according to two-dimensional schematics. Sect. 4 presents the detailed FEM/DEM numerical example of a rigid tire running on sand road, where the feasibility of the AMRM is also displayed. The conclusions of this study are listed in Sect. 5.

2 The basis for analyzing tire running behavior by using FEM/DEM

The motions of the discrete elements and the finite element nodes are governed by the Newton's Second Law. For arbitrary element i, the equations are expressed by Eq. (1) (used for both discrete elements and finite element nodes) and Eq. (2) (only used for discrete elements).

$$m_i(d^2\boldsymbol{u}_i / dt^2) = \boldsymbol{F}_i \tag{1}$$

$$I_i(d^2\boldsymbol{\theta}_i / dt^2) = \boldsymbol{M}_i \tag{2}$$

where m_i and I_i are the mass and inertia moment of element i, respectively; $\boldsymbol{u}_i$ and $\boldsymbol{\theta}_i$ are the displacement and the rotation angle of element i, respectively; $\boldsymbol{F}_i$ and $\boldsymbol{M}_i$ are the total external force and centroidal moment of element i, respectively. Eqs. (1) and (2) are solved by the explicit finite difference method.

The contact models for elements are shown in Fig. 1, where $\boldsymbol{h}_{ij}$ is the overlap of two contact elements; $\boldsymbol{v}_i$, $\boldsymbol{v}_j$, $\boldsymbol{\omega}_i$ and $\boldsymbol{\omega}_j$ are the velocity and angular velocity of element i and j, respectively; O_i, O_j are the mass center of the discrete element i and j, respectively; C is the contact point of the elements; $\boldsymbol{F}_n$ is the normal force, and $\boldsymbol{F}_s$, taken Coulomb friction law into account, is the tangential force among elements. $\boldsymbol{F}_{n,e}$ and $\boldsymbol{F}_{n,v}$ are the normal spring and the normal damping forces, respectively; $\boldsymbol{F}_{s,e}$ and $\boldsymbol{F}_{s,v}$ are the tangential spring and the tangential damping forces, respectively. The spring and damping forces are calculated by the Hertz-Mindlin theory [Balevičius et al., (2004)] for both the two types of contact, where the finite elements are regarded as spheres with infinite radius [Han et al., (2000)]; μ is the friction coefficient.
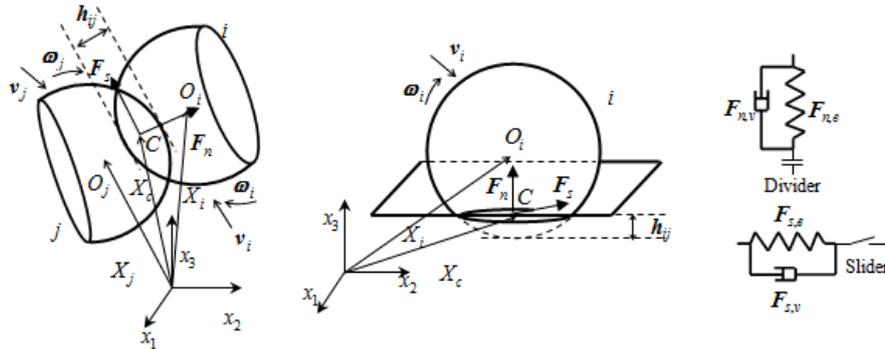
(a) Discrete elements model (b) Discrete and finite element model (c) Interaction forces
Figure 1. Contact models among elements

The concept of analyzing tire-sand interactions by using the FEM/DEM is illustrated in Fig. 2. The discrete elements contact with each other and with the finite element tire. Consequently, the tire drawbar pull N, vertical reaction force P and slip ratio s can be derived from Eqs. (3) - (5).

$$N = G - |\boldsymbol{R}| \tag{3}$$

$$P = \sum f_y \tag{4}$$

$$s = (1 - v / (r\omega)) \tag{5}$$

where f is the contact force between the finite elements and the discrete elements; $G = \sum f_{x+}$ and $R = \sum f_{x-}$ are the gross traction force and the resistance, respectively; v and ω are the translational speed and the angular velocity of the tire; r is the tire radius.
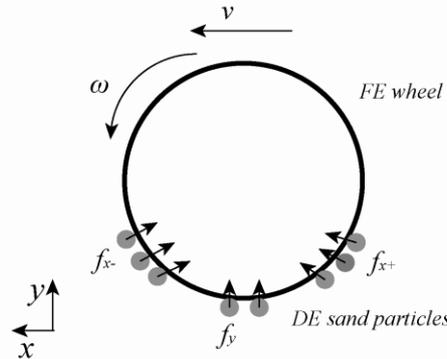
Figure 2. The tire-sand analysis system by using the FEM/DEM

3 The Alternately Moving Road Method

It is obvious that the sand outside a certain distance of the tire center have less influence on the tire running behavior. Thus, during the tire running process, the sand which is run over by the tire could be removed and new sand could be laid in front of tire to form new road. Accordingly, the Alternately Moving Road Method is proposed and the specific steps are as follows: first, the sand road sample, which is a section of the whole road, is established. Then, the initial sand road is assembled by combining two road samples in sequence. After that, the tire is placed on the sand road and starts to run. The alternation is performed when the tire travels a proper distance. The execution flowchart of the method is shown in Fig. 3, where T is the current calculation time; T_{ter} is the termination time, Δt denotes the time step of the explicit time integration.
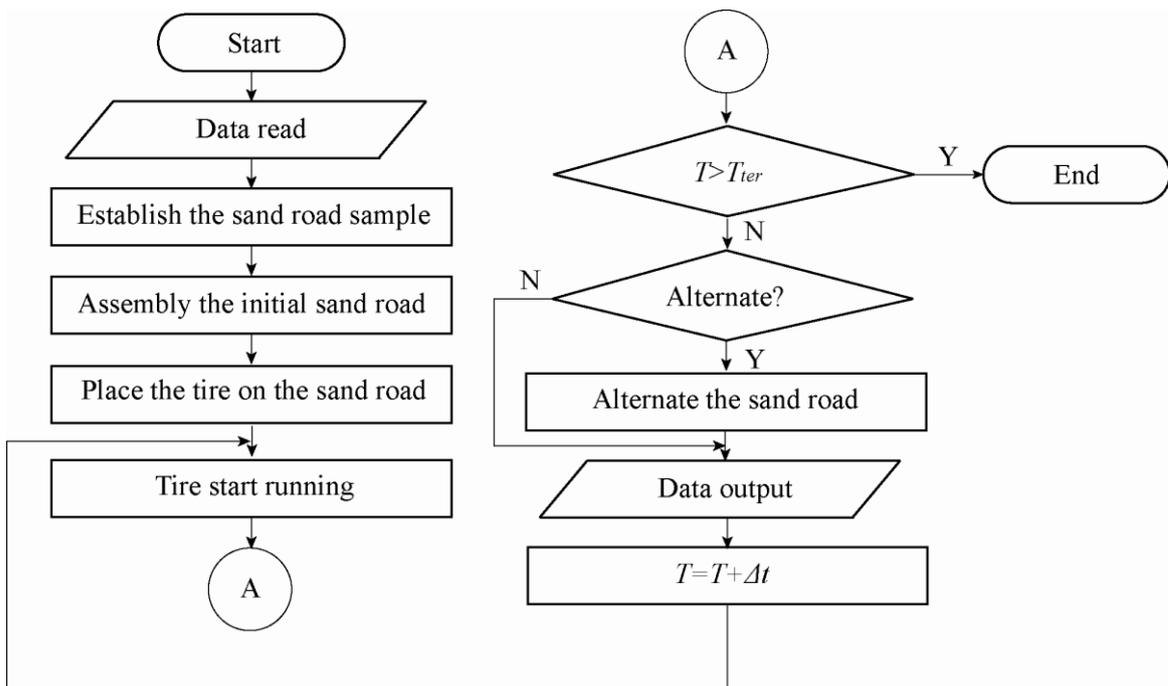
Figure 3. The execution flowchart

3.1 The establishment of sand road sample

Firstly, the discrete elements which are randomly distributed in a given domain are generated. There might be contact among the elements but no overlaps, contact forces or confining stress at this stage. Thus the elements should be rearranged to a steady state under self-weight to simulate the real sand. The boundary of the domain is constrained by rigid walls during the rearrangement process and the C-grid algorithm [Williams et al., (2004)] is applied for the contact detection among the discrete elements. The Discrete Element Set (DES) at the stable stage is the so-called sand road sample (recorded as DES *S*), as illustrated by two-dimensional schematic in Fig. 4. To facilitate the descriptions later, the element sequence numbers in the schematic are recorded as $1 \sim N$; the coordinates of arbitrary element *i* are recorded as X: $S_{x,i}$, Y: $S_{y,i}$; the length of the sample is equal to *a*; The constraint boundaries of the rigid wall are X: $[0,a]$, Y: $[0,b]$. The contact detection regions for the C-grid algorithm are also set to be X: $[0,a]$, Y: $[0,b]$.

Figure 4. The sand road sample

3.2 The assembly of the initial sand road

The initial sand road consists of two sand road samples mentioned in Sect. 3.1. The sand road sample DES *S* is duplicated into two DESs and recorded as DES A_0 and DES A_1. The DES A_0 and DES A_1 are arranged in sequence along the X direction to assemble the initial sand road, as illustrated in Fig. 5. The total element number of the initial sand road is *2N*. This is realized by adding the constant value *a* to the X coordinate of each element in DES A_1. Simultaneously, the element sequence numbers of DES A_1 are changed into $N+1 \sim 2N$. The coordinates of arbitrary element *i* in DES A_1 are X: $S_{x,i-N}+a$, Y: $S_{y,i-N}$. Other parameters of the elements are unchanged. The constraint boundaries of the rigid wall are altered into X: $[0,2a]$, Y: $[0,b]$. The contact detection regions for the discrete elements are also altered as X: $[0,2a]$, Y: $[0,b]$.
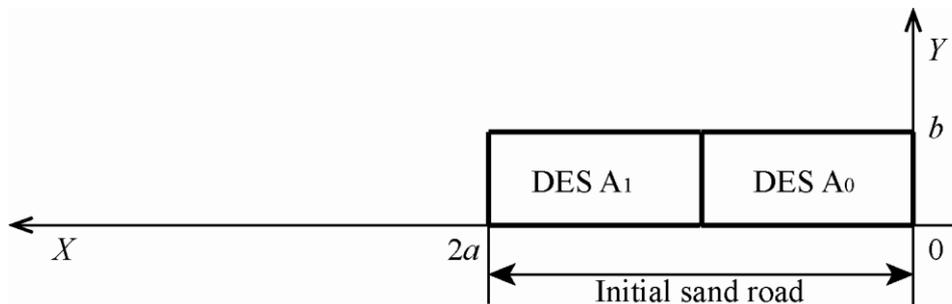
Figure 5. The initial sand road

3.3 The Alternately Moving Road process

The tire is placed at the center position of the DES A_0, as illustrated in Fig. 6, and

vertical load including self-weight of the tire and external load is loaded to the tire mass center. A constant angular velocity ω and corresponding translational speed v are enforced to the mass center of the tire after the vertical reaction force between the tire and the road is equal to the given vertical load. Then the tire starts to run along the X direction under specific slip ratio conditions.
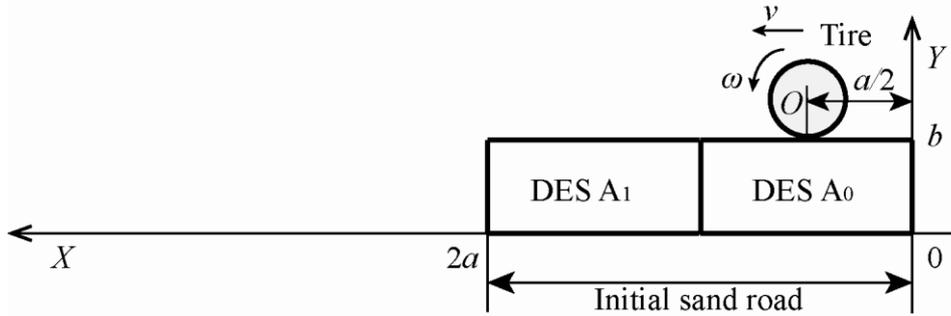
Figure 6. Initial position of the tire

The Alternately Moving Road process is performed when the tire arrives at the alternate point, where DES A_0 has almost no influence on the tire running behavior and the distal end of the DES A_1 is not serious damaged. In this study, the alternate point is at the center position of the DES A_1 as illustrated in Fig. 7.
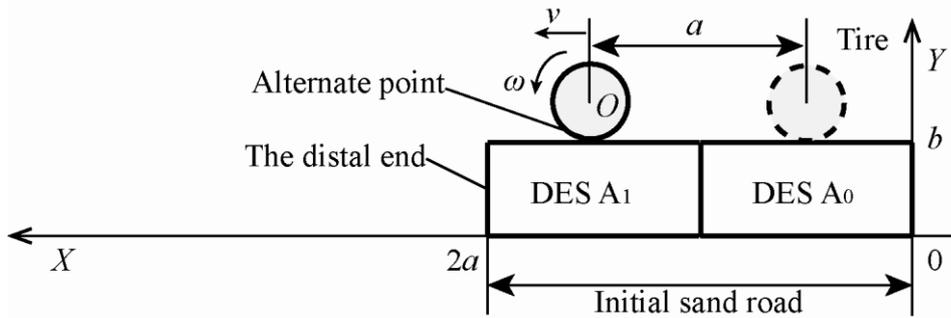
Figure 7. The alternate point for the AMRM

Then the elements of DES A_0 are removed and the road sample DES S (established in Sect. 3.1) is duplicated again and recorded as DES A_2. The element coordinates of DES A_2 are altered by adding the constant value $2a$ to the X coordinate of each element and the element sequence numbers are recorded as 1 to N. The coordinates of arbitrary element i are X: $S_{xi}+2a$, Y: S_{yi} after the alternation. Then, the DES A_1 and DES A_2 form a new sand road, as illustrated in Fig. 8. The constraint boundaries of the rigid wall are altered into X: $[0,3a]$, Y: $[0,b]$. The contact detection regions for the discrete elements are also changed into X: $[0,3a]$, Y: $[0,b]$.
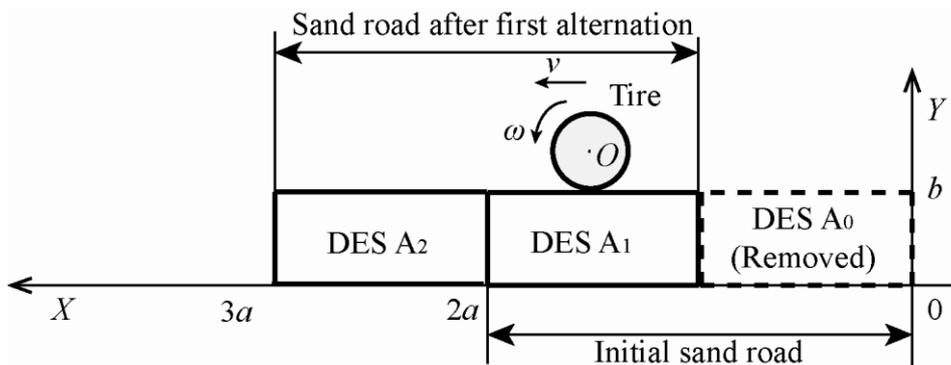
Figure 8. The first time of sand road alternation

251

Repeat the alternate process along the tire running direction and the tire running behavior on arbitrary length of sand road can be investigated with constant number of discrete elements. The alternate principles are as follows: for the kth alternation, if k is odd, the element sequence number is set to be 1~N; else if k is even, the element sequence number is set to be N+1~$2N$; The coordinates for arbitrary element i should be changed into X: S_{xi}+(k+2)a, Y: S_{yi} for k is odd, and to be X: S_{xi-N}+(k+2)a, Y: S_{yi-N} for k is even; The constraint boundaries of the rigid wall and the contact detection regions for the discrete elements should be changed into X: $[ka,(k+2)a]$, Y: $[0,b]$.

4 Numerical examples

Three-dimensional numerical model of tire running on sand road is established based on the soil-bin experiment in [Shinone et al., (2010)] to validate the feasibility and effectiveness of the AMRM in the investigation of the tire-sand interactions, where the sand road is modeled by discrete elements and the tire is discretized into finite elements. And the tire running behavior under different slip ratio is also investigated.

4.1 The sand road sample

Firstly, the discrete elements, which are randomly distributed in a given domain of X: $[0,735]$, Y: $[0,480]$, Z: $[0,280]$, are generated, as illustrated in Fig. 9. The friction coefficient between the discrete element and the rigid wall is set to be 0.3. The displacement contour of the discrete elements in Z direction during the rearrangement process under self-weight is shown in Fig. 10. The time history of the total gravitational potential energy (TGPE) is shown in Fig. 11. The value shows a decreasing trend and tends to be stable after 1.1 s of rearrangement. Then the rearrangement process is completed. The porosity value for the final state is about 0.32. The discrete element set is stored as the road sample. It should be noticed that this paper is focusing on the validation of the effectiveness of the AMRM, thus the radius range of the discrete element is 6~7 mm which is larger than the real sand.
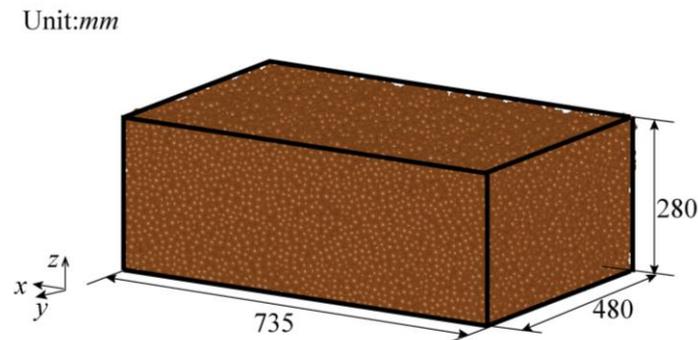
Figure 9. The configuration of the discrete element after the initial generation
Discrete element parameter: Young's Modulus: 75000 MPa, Poisson's Ratio: 0.3, Density: 2400 kg/m^3, Element number: 45551, Friction coefficient: 0.3.
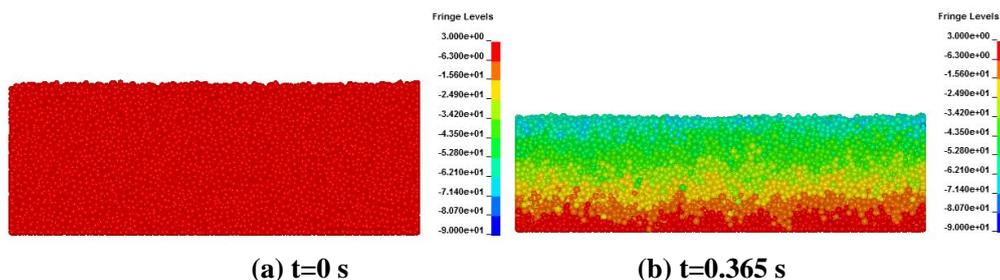
(a) t=0 s   (b) t=0.365 s

(c) t=0.730 s   **(d) t=1.095 s**
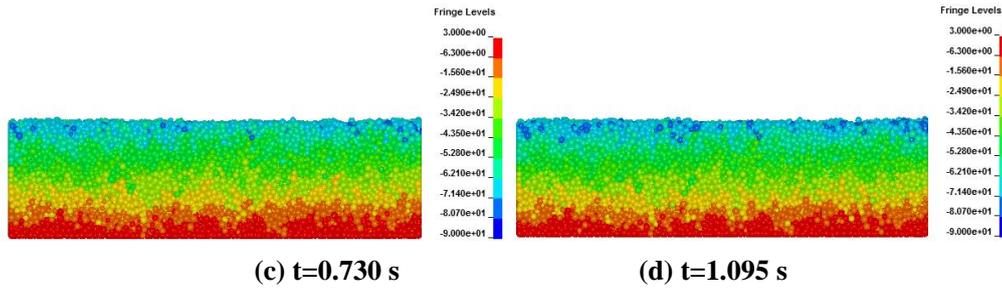
Figure 10. The displacement counter of the discrete elements in *Z* direction during the rearrangement process under self-weight (front view)
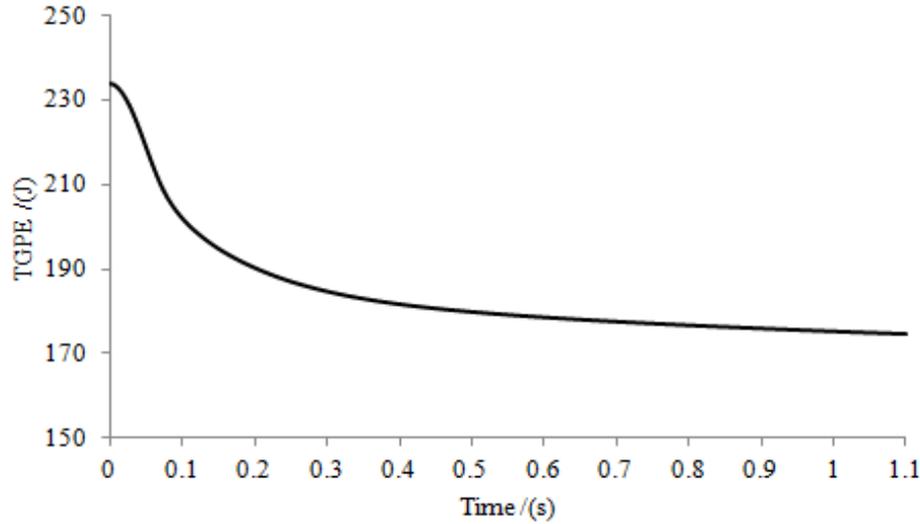
Figure 11. Time history of the TGPE of the discrete elements

4.2 Models of the initial sand road and the tire

The initial sand road consists of two road samples as illustrated in Fig. 12. The model size parameters are refer to the soil-bin experiment in [Shinone et al., (2010)]. The discrete element number is 91102, the length of the initial sand road is 1470 *mm* which is two times of the sand road sample illustrated in Sect. 4.1. The parameters of the discrete elements are the same as the ones in Sect. 4.1.

(a) Front view   **(b) Left view**
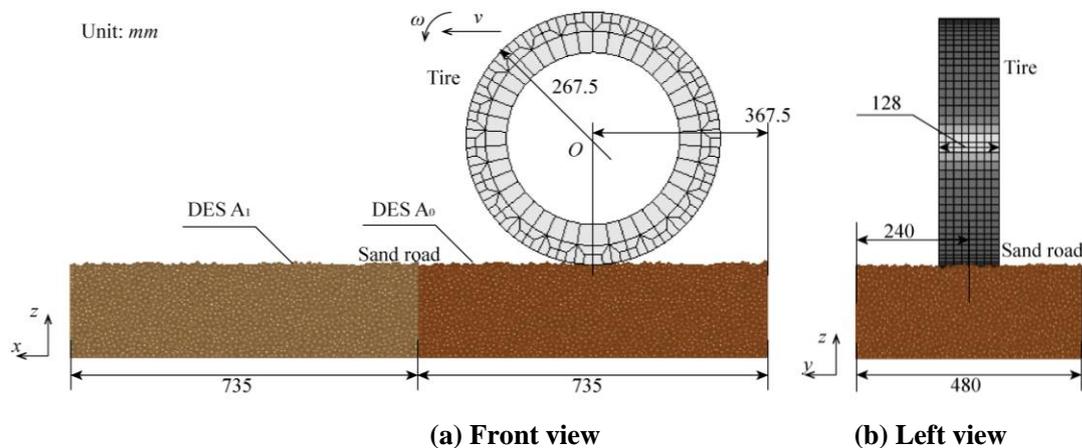
Figure 12. The models of the initial sand road and the finite element tire

Parameters of the tire: Young's Modulus: 2 *MPa*, Poisson's Ratio: 0.49, Density: 1800 *kg/m*³,
Element number: 1344.

The tire is placed on the sand road. Vertical load of 1295 *N* including the self-weight of the tire and the external vertical load is loaded to the center of the tire. The tire

sinks onto the sand road until the vertical reaction force between the tire and road reached 1295 *N*. Afterwards, constant angular velocity of 0.5 *rad/s* and corresponding translational speed are enforced to the tire center, and the tire travels towards the *X* direction under 30% slip ratio. The tire deformation is neglected for the given experimental inflation pressure of the tire in [Shinone et al., (2010)]. The friction coefficient between the tire and the sand road is set to be 0.4.

4.3 The alternately moving road process

The tire running along the *X* direction and the alternation of the sand road is performed when the tire running a distance of *k**735 *mm*, where *k* is alternate times. The displacement counter of the discrete elements in *Z* direction during the traveling process is shown in Fig. 13: Fig. 13(a) is the initial configuration; Fig. 13(b) shows the rut of the tire at time 0.777 *s*; when the tire travels a distance of 735 *mm*, the first alternation is performed, as shown in Fig. 13(c); then, the tire continues to move, and the rut of the tire at 1.58 *s* is shown in Fig. 13(d); After 1470 *mm* of travel, the second alternation is performed, as illustrated in Fig. 13(e); The total tire traveling distance is 1560 *mm*, and the final configuration is shown Fig. 13(f). During the running process, the length of the sand road keeps a constant value of 1470 *mm*, the total number of the discrete elements keeps a constant value of 91102.

The simulation is carried out on a PC. The principal characteristics of the PC are Intel Core i3-2100 1.58GHz (CPU), 2.00GB (RAM) and Windows XP Home Basic SP3 32bit. The elapsed time for the above numerical test is approximately 72 hours.

(a)

(b)

(c)

(d)

(e)

(f)
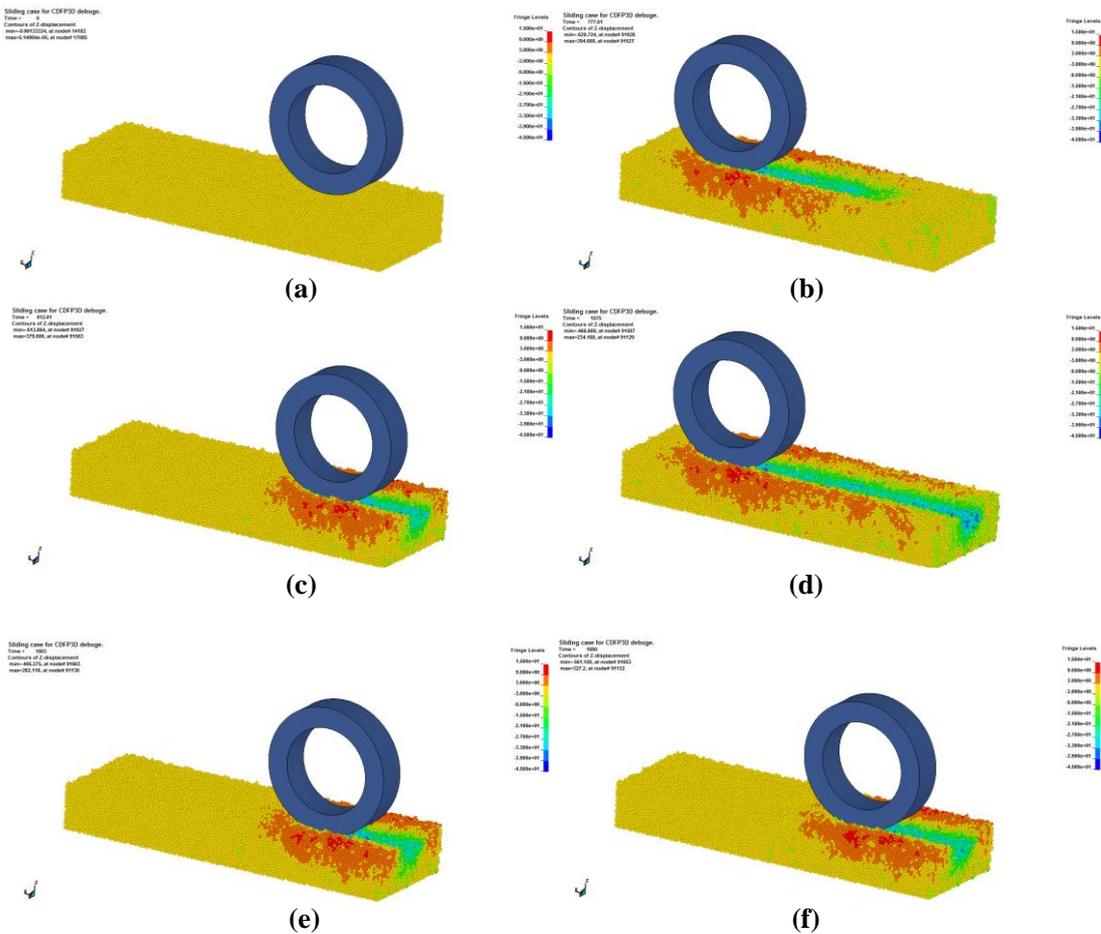
Figure 13. Displacement counter of the discrete elements under the rolling tire in 30% slip ratio

4.4 Tire running behavior

Fig. 14 shows the vertical reaction force as a function of the traveling distance under the 30% slip ratio condition. It can be seen that the vertical reaction force shows an abrupt fluctuation at the initial stage. The possible reason is that there is a vertical downward velocity when the tire was placed on the sand road (see Sect. 4.2 for detail), this lead to an impact between the tire and the road. Afterward, the tire vertical reaction force tend to be stable and its value fluctuates around 1295 *N* which is the given load value.
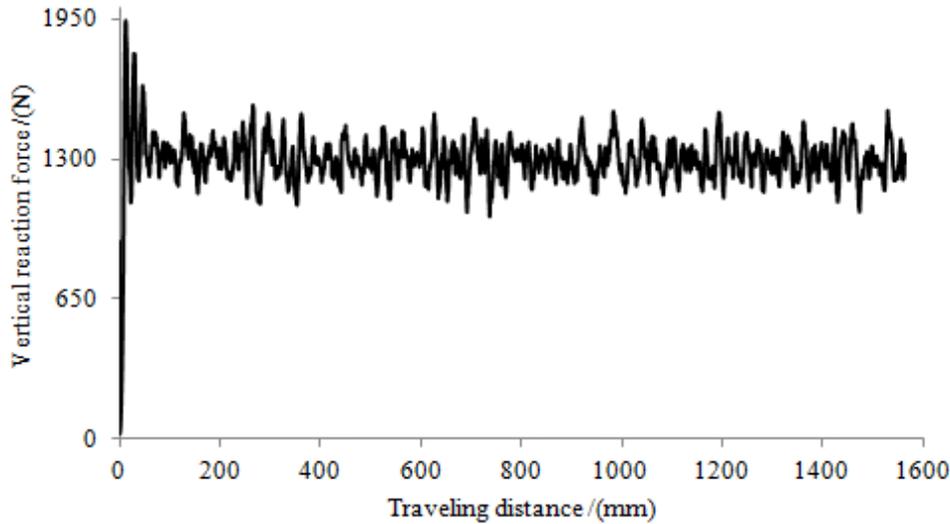
Figure 14. The vertical reaction force during the traveling process

Fig. 15 shows the drawbar pull as a function of the traveling distance under the 30% slip ratio condition. Analogous to the abrupt fluctuations of the vertical reaction force, the drawbar pull also shows a dramatic fluctuation at the initial stage because the tire traction force *G* is proportional to the vertical reaction force. After that the drawbar pull value is relatively stable without any abrupt fluctuations at the alternate point and its value fluctuates around 75 *N*. The possible reason for the fluctuations of the drawbar pull is the large radius values of the discrete elements.
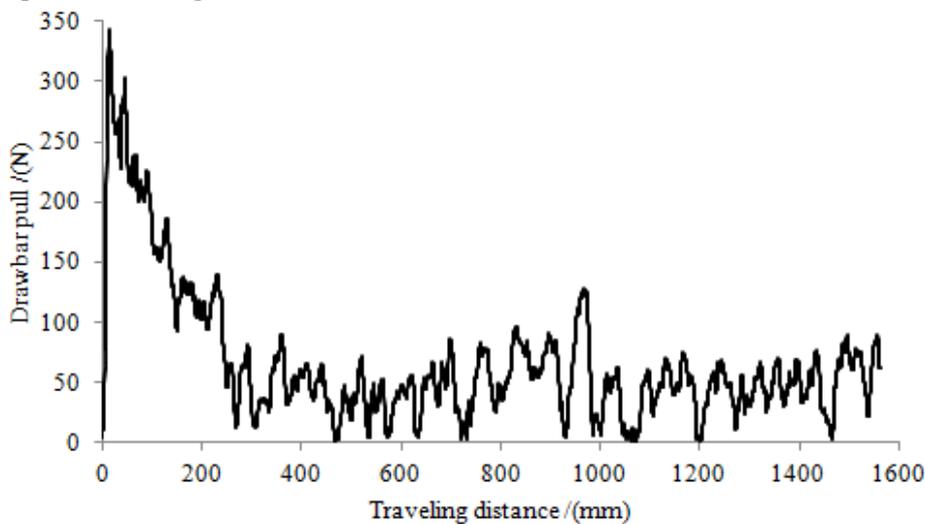
Figure 15. The tire drawbar pull during the traveling process

Fig. 16 shows the tire sinkage value as a function of the traveling distance under the 30% slip ratio condition. It can be seen that the tire sinkage value increases

dramatically at the initial stage due to the unbalance vertical force acting on the tire. Then its value fluctuates around the value of 45 *mm* and there are no abrupt fluctuations at the alternate point.
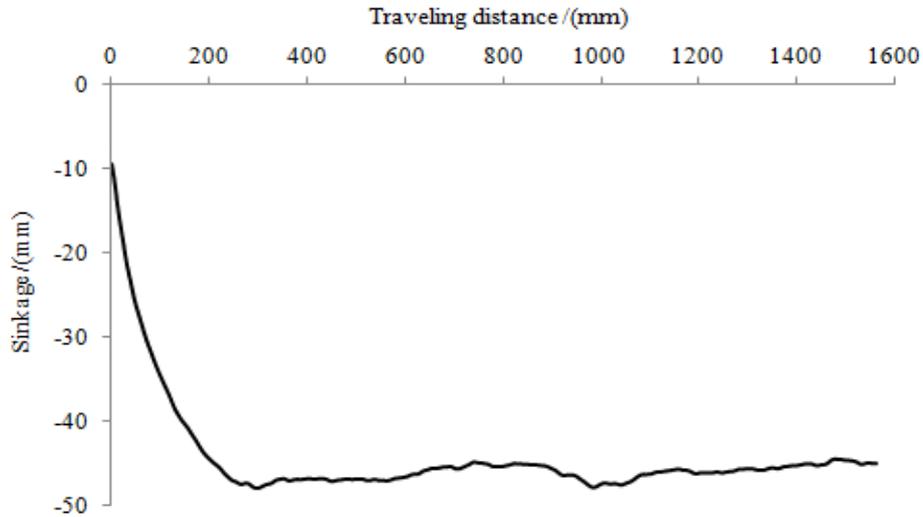
Figure 16. The tire sinkage during the traveling process

Fig. 17(a) presents the flow trend of the discrete elements in the *X-Z* plane under the rolling tire with 30% slip ratio. Herein, the velocity vectors of the elements are used to display their flow trend. It can be seen that the flow trend can be divided into two areas: the forward area flow in clockwise direction due to the bulldozing force of the tire and the rear area in anticlockwise direction because of the traction force of the rolling tire (digging effect). This result agrees qualitatively with the experimental result [Zhuang, (2002)] as illustrated in Fig. 17(b).
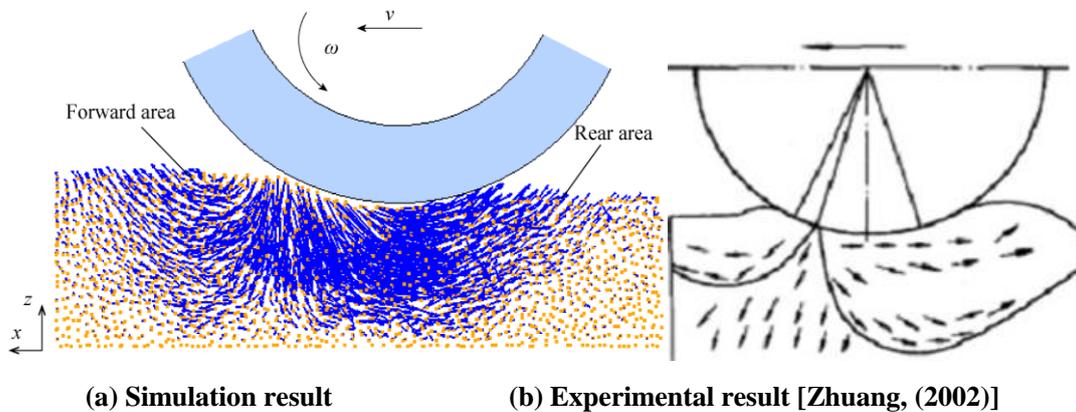
     (a) Simulation result   **(b) Experimental result [Zhuang, (2002)]**
Figure 17. Flow trend of the sand particles under a rolling tire

A constant angular velocity of 0.5 *rad/s* and corresponding translational velocity for different slip ratios according to Eq.(5) are loaded to the mass center of the tire to further analyze the influence of the slip ratio on the tire running behavior. Fig. 18 illustrates the tire equivalent sinkage values (the average sinkage value under each slip ratio) as a function of the slip ratios. It can be seen that the equivalent tire sinkage values are rise with the increase of the slip ratio. And the trend becomes steeper when the slip ratio is larger. This agrees qualitatively with the experimental results in [Shinone et al., (2010)]. However, the simulation results are larger than the experiment results. The possible reason is that the parameters in this study are decided by a trial and error preliminary computation to ensure the numerical stability at this stage. And the selection of the microscopic parameters among the discrete

elements has strong effects on its macroscopic mechanics.
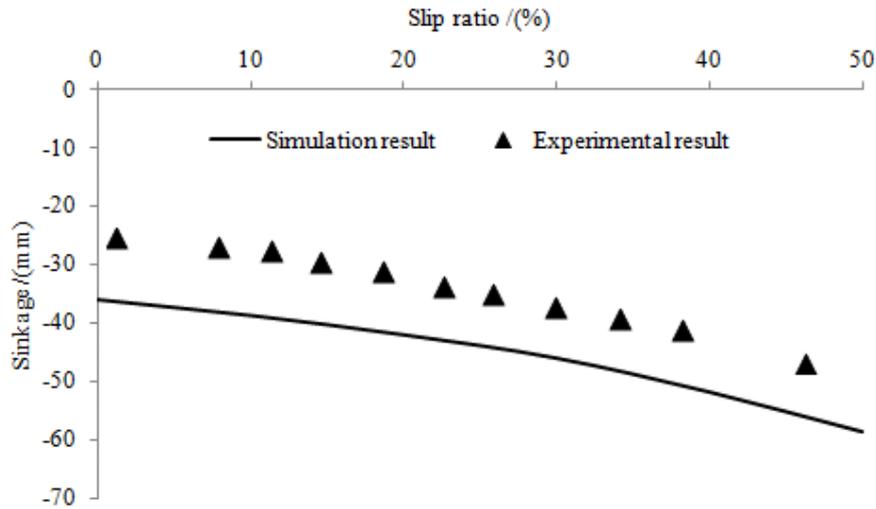
Figure 18. The relation between tire sinkage and slip ratio

Fig. 19 illustrates the equivalent values (the average drawbar pull value under each slip ratio) of the drawbar pull as a function of the slip ratios. It can be seen that the drawbar pull shows an increasing trend when the slip ratio is less than 25%, and its value tend to be stable when the slip ratio is larger than 25%. Such whole developing trend agrees qualitatively with the experimental result in [Shinone et al., (2010)]. It should be noticed that there is a large difference between the simulation results and the experimental results when the slip ratio is less than 15%. The possible reason for this phenomenon is the larger translational speed of the tire for the smaller slip ratio condition according to Eq. (8), and this leads to larger tire bulldozing resistance. The drawbar pull values have a little decrease after the slip ratio value is larger than 35%. This is because the bulldozing force is even larger due to the larger tire sinkage values under these slip ratio conditions. It should also be noticed that all the drawbar pull values are smaller than the experimental results because of the larger sinkage values, as shown in Fig. 19, which leads to larger bulldozing resistance.
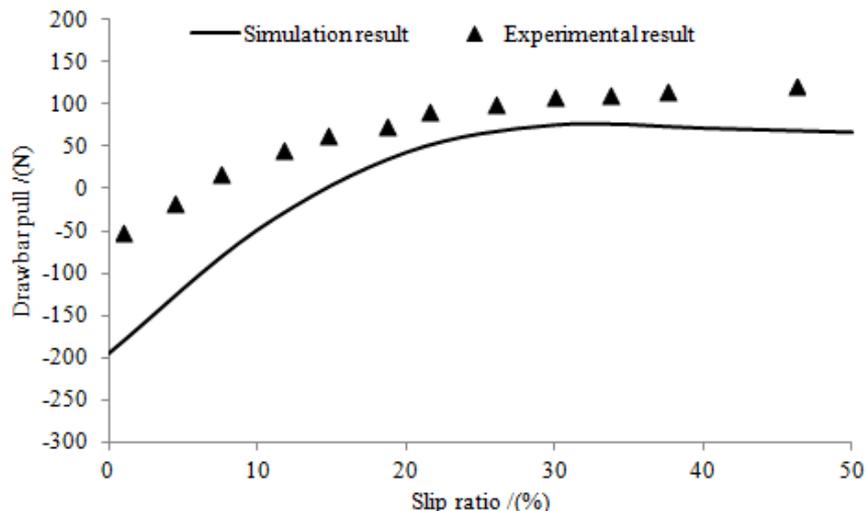
Figure 19. The relation between tire drawbar pull and slip ratio

4.5 Discussions

As can be seen from the Fig. 14, Fig. 15 and Fig. 16, the alternate moving road process is stable and effective for the simulation of tire running behavior on arbitrary

length sand roads.

As can be seen from Fig. 17, Fig. 18 and Fig. 19, the tire slip ratio has strong effect on its running behavior. The sinkage value of the tire rises with the increases of the slip ratio due to the tire traction effect (digging effect). This leads to the increase of the bulldozing resistance which is the main resistance for tire running on sand road. The drawbar pull shows a steeper increasing trend when the slip ratio is less than 25% and the values tend to be stable when the slip ratio is larger due to the dramatically increase of the resistance under these slip ratio conditions.

The comparisons between the simulation results and current experimental results show that the FEM/DEM is a straightforward and effective tool to investigate the tire running behavior on sand road, where the flow trend of the sand particles under a rolling tire, the drawbar pull and the sinkage of the tire and the dynamic parameters such as vertical acceleration value of the tire can be obtained expediently and reasonably.

5. Conclusions

From the above investigation, following results can be obtained:

(1) The Alternately Moving Road Method is proposed and applied to the FEM/DEM simulation of tire running behavior on the sand road. This method possesses the ability of simulating arbitrary length of sand road with constant discrete element numbers. Numerical simulation results show that the method is stable and effective.

(2) The tire running behavior such as the normal reaction force, tire drawbar pull, tire sinkage and flow trend of the sand particles can be obtained conveniently by the FEM/DEM. The comparisons between the simulation and current experimental results show that FEM/DEM is an effective and promising approach to simulate the tire running behavior on the sand terrain. The current research work is not only appropriate for the tire-sand interactions, but also suitable for the investigation of other terramechanics problems such as soil cultivation process.

Plans for the future work are to improve the accuracy of the method. The size effect of the discrete elements and the new discrete element interaction models considering the rolling resistance [Ai et al., (2011); Jiang et al., (2005); Kuhn and Bagi, (2004)] should be investigated.

Acknowledgements

References

Ai, J., Chen, J.-F., Rotter, J.M., Ooi, J.Y. (2011) Assessment of rolling resistance models in discrete element simulations, *Powder Technology* **206**, 269-282.
Balevičius, R., Džiugys, A., Kačianauskas, R. (2004) Discrete element method and its application to the analysis of penetration into granular media, *Journal of Civil Engineering and Management* **1**, 3-14.
BIRIS, S., Ungureanu, N., Maican, E., Murad, E., Vladut, V. (2011) FEM model to study the influence of tire pressure on agricultural tractor wheel deformations, *Engineering for rural development*, 27-27.
Cuong, D.M., Zhu, S., Ngoc, N.T. (2013) Study on the variation characteristics of vertical equivalent

damping ratio of tire–soil system using semi-empirical model, *Journal of Terramechanics*.

David, A.H., John, F.P., Alex, C. (2001) Large scale discrete element modeling of vehicle-soil interaction, *International Journal for Numerical Methods in Engineering* **127**, 1027-1032.

González Cueto, O., Iglesias Coronel, C.E., Recarey Morfa, C.A., Urriolagoitia Sosa, G., Hernández Gómez, L.H., Urriolagoitia Calderón, G., Herrera Suárez, M. (2013) Three dimensional finite element model of soil compaction caused by agricultural tire traffic, *Computers and Electronics in Agriculture* **99**, 146-152.

Han, K., Peric, D., Crook, A.J.L., Owen, D.R.J. (2000) A combined finite/discrete element simulation of shot peening processes–Part I: studies on 2D interaction laws, *Eng Computation* **17**, 593-620.

Jiang, M.J., Yu, H.S., Harris, D. (2005) A novel discrete model for granular material incorporating rolling resistance, *Computers and Geotechnics* **32**, 340-357.

Khot, L.R., Salokhe, V.M., Jayasuriya, H.P.W., Nakashima, H. (2007) Experiment validation of distinct element simulation for dynamic wheel-soil interaction, *Journal of Terramechanics* **44**, 9.

Knuth, M.A., Johnson, J.B., Hopkins, M.A., Sullivan, R.J., Moore, J.M. (2012) Discrete element modeling of a Mars Exploration Rover wheel in granular material, *Journal of Terramechanics* **49**, 10.

Kuhn, M.R., Bagi, K. (2004) Contact rolling and deformation in granular media, *International journal of solids and structures* **41**, 5793-5820.

Li, H., Schindler, C. (2013) Investigation of Tire-Soil Interaction with Analytical and Finite Element Method *Mechanics Based Design of Structures and Machines* **41**, 293-315.

Moslem, N., Hossein, G. (2014) Numerical simulation of tire/soil interaction using a verified 3D finite element model, *Journal of Central South University* **21**, 817-821.

Nakashima, H., Fujii, H., Oida, A., Momozu, M., Kawase, Y., Kanamori, H., Aoki, S., Yokoyama, T. (2007) Parametric analysis of lugged wheel performance for a lunar microrover by means of DEM, *Journal of Terramechanics* **44**, 153-162.

Nakashima, H., Oida, A. (2004) Algorithm and implementation of soil-tire contact analysis code based on dynamic FE-DE method, *Journal of Terramechanics* **41**, 127-137.

Nakashima, H., Takatsu, Y., Shinone, H. (2008) Analysis of tire tractive performance on deformable terrain by finite element-discrete element method, *Journal of computational science and technology* **4**, 423-434.

Nakashima, H., Takatsu, Y., Shinone, H. (2009) FE-DEM analysis of the effect of tread pattern on the tractive performance of tires operating on sand, *Journal of Mechanical Systems For Transportation and Logistics* **2**, 55-65.

Shinone, H., Nakashima, H., Takatsu, Y. (2010) Experimental analysis of tread pattern effects on tire tractive performance on sand using an indoor traction measurement system with forced-slip mechanism, *Engineering in Agriculture, Environment and Food* **3**, 61-66.

Smith, W., Peng, H. (2013) Modeling of wheel-soil interaction over rough terrain using the discrete element method, *Journal of Terramechanics* **50**, 277-287.

Williams, J.R., Perkins, E., Cook, B. (2004) A contact algorithm for partitioning *N* arbitrary sized objects, *Engineering Computations* **21**, 235-248.

Zang, M.Y., Gao, W., Lei, Z. (2011) A contact algorithm for 3D discrete and finite element contact problems based on penalty function method, *Computational Mechanics* **48**, 541-550.

Zhang, R., Liu, F., Zeng, G.Y., Li, J.Q. (2012) Research on dynamic simulation system of interactions between irregular rigid wheel and lunar soil simulant, CMGM-2012, China:Zhangjiajie, pp. 438-444.

Zhao, C.L., Zang, M.Y. (2014a) Analysis of rigid tire traction performance on a sandy soil by 3D finite element-discrete element method, *Journal of Terramechanics*, (Submitted).

Zhao, C.L., Zang, M.Y. (2014b) An improved discrete and finite element contact algorithm for the analysis of wheel-sand interaction, *Computers and Structures*, (submitted).

Zhuang, J.D. (2002) Computational vehicle terramechanics, 1st ed. *CMP*, China.

Coherent and Compatible Statistical Models in Structural Analysis

*M. Nogal[1], E. Castillo[2], A. Calviño[2] and A. J. O'Connor[1]

[1]Department of Civil Engineering, Trinity College Dublin, Dublin 2, Ireland.
[2] Department of Applied Mathematics and Computational Sciences, University of Cantabria, 39005 Santander, Spain.

*Corresponding author: nogalm@tcd.ie

Abstract

Modelling problems in structural analysis requires of a statistical approach that allows us to take into account the random nature of the variables as well as the uncertainties involved in the problem being analysed. However neither all statistical models are valid nor all assumptions are mathematically or physically reasonable. The aim of this paper is twofold: (a) to explain how to build statistical models with mathematical and physical coherence, and (b) to describe the most common mistakes made when building or selecting mathematical and statistical models. We provide some interesting tools to carry out this important task and present some examples that show the inconveniences and consequences derived from an incorrectly established model.

Keywords: Location-scale stable families, Structural analysis statistical models, Specification of multivariate joint distributions, Extreme Values, Probability papers.

Introduction

Before selecting a model to solve a given engineering problem, a very important step consists of dedicating sufficient time to study the problem under consideration in some depth. This means that the engineer must understand the problem, the variables and the implied physical relations, which should be present in the model. For example, an engineer dealing with a breakwater needs to understand that the large waves and winds are the most important agents implied in design. This means that maxima events and then maxima extreme value distributions must be considered. The limited or unlimited range of the random variables involved is also relevant, because this permits excluding either the Weibull or Frechet type of distributions. Ignoring these aspects leads to unconservative or very expensive solutions which are engineeringly regrettable.

It is also convenient to use simple models, that is, as parsimonious as possible and dimensionally consistent. In this line, the Buckingham theorem plays a fundamental role and should be the first step in equation modelling. Apart from reducing the number of variables involved and avoiding us to be concerned about dimensions, it permits us to check if the selected variables are sufficient or need to be completed with additional variables to reproduce a physical problem or phenomena.

Another important decision to be made when building models is the selection of the families of random variables used to reproduce the real ones. In this context, the designer must take into account the variable ranges and be aware that not all distribution families are valid for reproducing all types of variables. In this context, one should know that some distributions are valid only for dimensionless variables (Poisson, beta, binomial, etc.) and that some distributions are not scale (geometric, chi-squared, etc.) or location (gamma, log-normal, etc.) stable. For example, selecting non-scale families means that the resulting models will not be valid for variables when written in terms of different measure units, and then they are inadequate.

Since a statistical analysis requires the joint distribution of all variables involved, the selection of a multivariate model is crucial too. In this line it is important to use feasible models. We point out that in some cases a lack or an excess of simplifying assumptions can lead to undefined or inexistent models, respectively.

As a final example, the designers should be aware of the existence of different probabilistic papers (maximum, minimum, etc.) and that not all data points but only those in the tail of interest must be used when dealing with extremes. Ignoring these facts can be catastrophic.

Since we consider that all these issues are very relevant for engineers, they are discussed with some detail in this paper.

The aim of this paper is twofold. On one hand we introduce some considerations to be taken into account when building statistical models and, on the other hand, we point out some problems we can find when these aspects are not considered. Besides, we provide some tools to facilitate this task together with several examples for a better comprehension of the concepts discussed.

The paper is organized as follows. In Section 2 we present a brief review of some of the statistical models proposed in different Civil Engineering fields. In Section 3 we make some considerations about the units of the random variables and their moments. In Section 4 we emphasize the importance of the Buckingham theorem in order to build parsimonious and dimensionless models. In Section 5 we deal with extreme values and probability papers. In Section 6 we explain and discuss different possibilities to define multivariate models and finally, in Section 7 we give some conclusions.

Some statistical models proposed in the literature

In the Civil engineering literature it is becoming more frequent to find statistical approaches. For example, reliability analysis has reached all engineering fields. Due to the abundant bibliography dealing with this issue, as a sample and for illustration purposes, Table 1 shows a list of some examples of distributions used in the Civil Engineering literature.

Table 1: Some probability distribution families used in the literature together with the corresponding engineering variables.

VARIABLE	DISTRIBUTION	VARIABLE	DISTRIBUTION
Geometric and mechanical properties	log-normal and normal	Maximum wave height	reverse Weibull
Material properties	normal, two- and three-parameter Weibull	Two successive wave periods	bivariate Weibull and bivariate Rayleigh
Excedences of wave height or significant wave height	Generalized Pareto	Significant wave heights	Weibull, generalized gamma, generalized beta kind I and beta kind II
Stress range	Raleigh, wide-band, Weibull, beta, log-normal, Rice's and normal distributions	Significant wave height and wave period	Box-Cox + bivariate normal, bivariate log-normal and bivariate Plackett
Loads	Poisson, Gumbel and normal	Small wave heights in large depths	Rayleigh
Wind speed	Frechet, Gumbel, reverse Weibull and log-normal	Joint density of significant wave height, wave period and current and wind speeds	Marginals transformed to normals by Box-Cox transformation plus multivariate normal
Wave period	log-normal	Wave height	Rayleigh distribution and reverse three-parameter Weibull
Fatigue life	Weibull		

It is relevant to say that some of the used models above are theoretically justified and some are used just for convenience or to facilitate calculations or mathematical derivations. For example, the normal model is justified when the random variable being modelled is the sum of a large enough number of other variables. This occurs frequently in strength of materials where in a cross section of volume all the subelements add efforts or collaborate to resistance. Poisson and gamma distributions have been proved to correspond to rare events and the time of occurrence of the r-th event, respectively. The Weibull, Gumbel and Frechet extreme value distributions and their reverse versions are justified because they are the limit distributions of maxima or minima, which are very important in Civil Engineering design because in general only maxima (waves, winds, earthquakes, temperatures, etc.) and minima (draughts, fatigue strength, temperatures, etc.) values lead to failure. However, it is not uncommon to see minima models erroneously used for maxima or vice versa. The generalized Pareto distribution is justified because it arises as a limit distribution for exceedances (large waves, winds, etc.) over or shortfalls (rain, temperature, etc.) under a threshold. Rice and Rayleigh distributions are also derived from theoretical models of waves.

Contrary, other distributions, such as the log-normal that arises in order to reproduce asymmetric data, the generalized beta and the models based on the Box-Cox transformation that are used to fit different data histograms, etc. have convenience as motivation.

In the Structures field, for example, [O'Connor and Kenshel (2013)] use the normal distribution to describe concrete material properties, [O'Connor and Enevoldsen (2009)] propose Log-normal distributions for modelling structural parameters and uncertainties associated with modelling, [Simiu et al. (1980)] assume the Fretchet distribution for the wind speed and [Pourzeynali and Datta (2005)] suggests the Raleigh distribution to model the stress range.

In the Material Science field, [Castillo and Fernández-Canteli (2009)] develop a fatigue model using a three-parameter Weibull distribution for a normalizing variable representing the whole S-N field based on a unique distribution function, [Koller et al. (2009)] validate the use of a log-Gumbel fatigue regression model and [Przybilla et al. (2011)] propose a method to obtain the distribution of fracture stress as a three-parameter Weibull cumulative distribution function (cdf) referred to a uniaxially and uniformly tensioned surface element. We can also mention the case of Coast and Ocean Engineering where [Ferreira and Guedes Soares (1999)] assume significant wave heights to follow Beta distributions, [Ferreira and Guedes Soares (1998)] use the Generalized Pareto densities for excedences of wave heights and significant wave heights, or [Ochi (1992)] proposes the Generalized Gamma distribution for significant wave heights.

Another field with a wide variety of stochastic models is Transportation. Some examples are [Lo et al. (1996)] who propose independent Poisson link counts or [Castillo et al. (2012)] who develop a bayesian network considering that the different traffic variables follow a generalized beta distribution. Multinomial models were assumed by [Clark and Watling (2005)] for route flows and shifted-gamma distribution was used by [Castillo et al. (2013)] for modelling the traffic flows.

From the list of publications above we can realize that a large set of distributions has been used. Detected inconsistencies in some of the proposed models motivates the current paper, which presents essential aspects to be considered when building statistical models.

Some considerations on units of the random variable and their moments

One common mistake when building statistical models is to ignore that not all families of distributions are valid for all types of variables. We need to be aware that parameters of statistical families have units. In particular, the mean has the same dimension as the random variable and the variance the squared dimensions.

*Example 3.1 (**Exponential distribution**) For the exponential distribution Exp(λ) we have:*

$$E[X]=1/\lambda; \qquad Var[X]=1/\lambda^2. \qquad (1)$$

Since the dimension of $1/\lambda^2$ is the square of the dimension of 1/λ, the inverse of the variable unit, the dimensions are consistent in this case.

*Example 3.2 (**Beta distribution**) For the beta distribution, X~Beta(a,b) we have:*

$$E[X]=\frac{a}{a+b}; \qquad Var[X]=\frac{ab}{(a+b)^2(a+b+1)}. \qquad (2)$$

This implies that X must be dimensionless, because in the term a+b+1 a and b must be dimensionless; otherwise they cannot be added to 1 (dimensionless). Once that a and b have been recognized as dimensionless, E[X] and Var[X] are also dimensionless (see (2)).

*Example 3.3 (**Weibull distribution**) For the Weibull distribution, X~W(λ,k) we have:*

$$E[X]=\lambda\Gamma\left(1+\frac{1}{k}\right); \qquad Var[X]=\lambda^2\left[\Gamma\left(1+\frac{2}{k}\right)-\Gamma^2\left(1+\frac{1}{k}\right)\right], \qquad (3)$$

which implies that k must be dimensionless and λ must have the same dimensions as X, and that the Weibull model can be made consistent for variables of any dimensions.

*Example 3.4 (**Gamma distribution**) If the random variable X is Gamma X~W(λ,k), the random variable X+a with a≠0 is not gamma any more. This means that the gamma family is not stable with respect to changes in location and has important consequences, because the gamma distribution cannot be used for location variables, such as temperatures. More precisely, if a random temperature is gamma measured in Celsius degrees, it is not gamma when measured in Farenheit or Reamur degrees. Thus, using the gamma family for temperatures is inadequate and misleading.*

Other examples of dimensionless families are the binomial, negative binomial and Poisson. Contrary, normal distributions are examples of statistical families compatible with any dimension.

Parsimonious and dimensionless models: The Buckingham theorem

When a mathematical or statistical model is built, a dimensional analysis of the variables involved must be initially carry out as this allows us to understand some deep relations among these variables and help to avoid dimensional contradictions. Besides, it is recommendable to build a dimensionless model in order to prevent dimensional inconsistencies and in some cases to reduce the problems associated with precision in numerical evaluations. Finally, it is important to work with parsimonious models, that is, the simplest models explaining all the aspects to be considered. To these aims the Rayleigh method of dimensional analysis and its formalization proposed by [Buckingham (1915)] plays a fundamental role. To illustrate, we propose the following example.

*Example 4.1 (**Corbel Example. Dimensionless variables**) The example deals with a reliability analysis of a corbel by means of the strut-and-tie model represented in Figure 1. In this case we assume two possible failure modes, defined by the limit-state functions H_1 and H_2:*

$$H_1 \equiv f_s A_s - F_v \tan\theta - F_h = 0, \qquad (4)$$

$$H_2 \equiv f_c A_b - F_v = 0 \qquad (5)$$

where F_v and F_h are the applied vertical and horizontal forces, respectively, f_s and f_c are the strength of the steel and of the compressed concrete, θ is the angle between the compression strut and the tie, A_s is the cross sectional area of the passive reinforcement and A_b is the area where the action is applied.
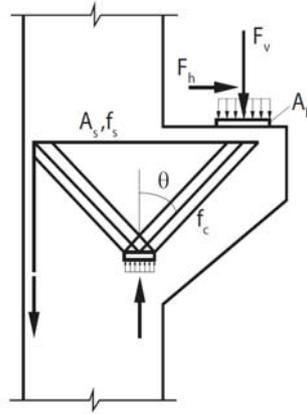
Figure 1: Corbel Example. Strut-and-tie model.

The failure curve can be expressed as the minimum value of the previous limit-state functions, that is,

$$H = min \{H_1, H_2\} \tag{6}$$

Based on the Buckingham Π Theorem, we get the dimensional decomposition shown in Table 2, where [F] and [L] denote force and length magnitudes, respectively. We see that the $n=7$ variables set up a dimensional matrix with rank $q=2$. Applying the Buckingham Π Theorem, we conclude that the model (6) is equivalent to another with $p=n-q=5$ dimensionless parameters (ratios).

Table 2: Corbel Example. Dimensional decomposition.

	F_v	F_h	f_s	f_c	θ	A_s	A_b
[F]	1	1	1	1	0	0	0
[L]	0	0	-2	-2	0	2	2

If we use f_s and A_s as reference or normalizing variables, we obtain the following new dimensionless variables:

$$F_v^* = \frac{F_v}{f_s A_s}; \quad F_h^* = \frac{F_h}{f_s A_s}; \quad f_c^* = \frac{f_c}{f_s}; \quad \theta^* = \theta; \quad A_b^* = \frac{A_b}{A_s}, \tag{7}$$

and the new mathematical expression for the model (6) becomes:

$$H^* = \frac{H}{f_s A_s} = min\left[1 - F_v^* \tan \theta^* - F_h^*, \ f_c^* A_b^* - F_v^*\right], \tag{8}$$

where the asterisks refer to dimensionless variables. The main advantages of using the Buckingham theorem are:

1. The model presents $p=5$ variables instead of $n=7$, which implies a reduction in the problem complexity.

2. The variables are independent from any units being considered, avoiding possible dimensional mistakes. Moreover, the normalization modifies the variable ranges and reduces possible numerical precision problems.

3. The variable H^* becomes more meaningful than the associated dimensional variable H because its value can be compared for different cases when the steel characteristics (area and yield strength) are kept constant.

We point out that the dimensional results can be recovered at the end of the process, undoing the change proposed in (7).

Extreme values and probability papers

In the engineering design of structures we need to deal with extreme values, that is, maxima (for example, loads, moments, etc.) or minima (for example, strength properties). In such cases, a careful selection of extreme value distributions to approximate the distribution of extremes is required. In this paper we deal only with maxima, but the minimum problem is similar. In order to see if a cdf $F(x)$ can be approximated for maxima by a reverse Weibull, Gumbel or Frechet distribution we can use the following theorem by [Castillo (1988)].

Theorem 1 *If F(x) is the cumulative distribution function of a random variable and*

$$\lim_{\varepsilon \to 0} \frac{F^{-1}(1-\varepsilon) - F^{-1}(1-2\varepsilon)}{F^{-1}(1-2\varepsilon) - F^{-1}(1-4\varepsilon)} = 2^c \tag{9}$$

then F(x) can be approximated in its right tail by a Frechet distribution if c>0, a Gumbel distribution if c=0 and a Weibull distribution if c<0.

In particular, if the range of $F(x)$ is limited it cannot be approximated by a Frechet distribution and if it is unlimited, we cannot use a Weibull distribution.

As some interesting examples, Table 3 shows the corresponding approximating distributions of some of the most common distributions for maxima and minima.

Table 3: Corresponding approximating distributions for maxima and minima of the most common distributions.

Distribution	Domain of Attraction		Distribution	Domain of Attraction	
	Maximal	Minimal		Maximal	Minimal
Normal	Gumbel	Gumbel	Uniform	Weibull	Weibull
Exponential	Gumbel	Weibull	Weibull	Weibull	Gumbel
Log-normal	Gumbel	Gumbel	Weibull	Gumbel	Weibull
Gamma	Gumbel	Weibull	Cauchy	Fréchet	Fréchet
Gumbel	Gumbel	Gumbel	Pareto	Fréchet	Weibull
Gumbel	Gumbel	Gumbel	Fréchet	Fréchet	Gumbel
Rayleigh	Gumbel	Weibull	Fréchet	Gumbel	Fréchet
M = maxima m = minima			M = maxima m = minima		

The previous method permits determining the extreme value distributions associated with a given one $F(x)$. However, in practice we do not have this information but only data. In this case we can plot this data on a Maximal Gumbel probability paper, as shown in Figure 2. Then, looking to its right tail and determining whether the data trend is straight or has positive or negative curvature, we can decide about Gumbel, Weibull or Frechet as approximating distributions, respectively.

Building multivariate statistical models

In this section we deal with the problem of defining the joint multivariate density of all the variables which are relevant to the problem under consideration.

There are several ways to define the joint density of a multivariate model. These methods can be classified as underdetermined, overdetermined and uniquely determined methods, depending of the number of imposed conditions.
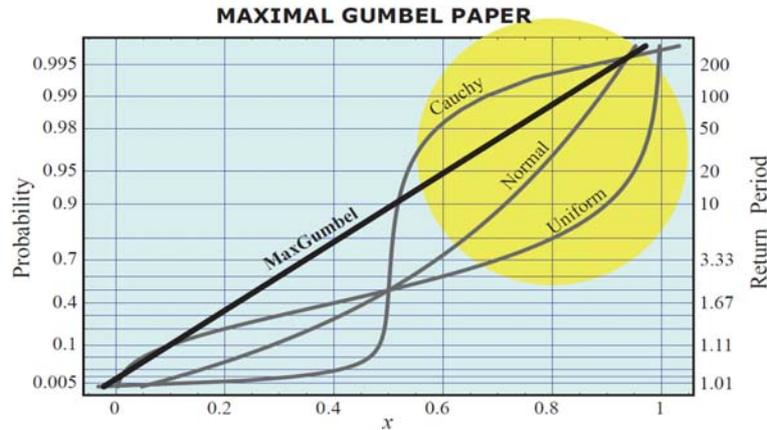
Figure 2: Maximal Gumbel probability paper illustrating the relevant zone.

In order to uniquely determine a multivariate model with an underdetermined method, we have to add some extra conditions. In the case of the overdetermined methods, the solution is not guarantied. For a more detailed description of these methods, see [Arnold et al. (1992,1999,2001)] and [Castillo et al (2014)].

The following example illustrates the cases of overdetermined and underdetermined methods.

*Example 6.1 (**Normal conditionals model**) [Arnold et al (1999)] demonstrate that there are two families of bivariate distributions with normal conditionals, that is, with conditionals X|Y=y and Y|X=x which are normals: (a) the normal and (b) a family with regression lines and conditional variances given by:*

$$E\left(X|Y=y\right) \; = \; \mu_1(y) = -\frac{m_{12}y^2 + m_{11}y + m_{10}}{2\left(m_{22}y^2 + m_{21}y + m_{20}\right)}, \tag{10}$$

$$var\left(X|Y=y\right) \; = \; \sigma_1^2(y) = \frac{-1}{2\left(m_{22}y^2 + m_{21}y + m_{20}\right)}, \tag{11}$$

$$E\left(Y|X=x\right) \; = \; \mu_2(x) = -\frac{m_{21}x^2 + m_{11}x + m_{01}}{2\left(m_{22}x^2 + m_{12}x + m_{02}\right)}, \tag{12}$$

$$var\left(Y|X=x\right) \; = \; \sigma_2^2(x) = \frac{-1}{2\left(m_{22}x^2 + m_{12}x + m_{02}\right)}, \tag{13}$$

where the m's are constants.

One example of a normal density is shown in the left plot of Figure 3, where the linear regression lines are shown on the top projection and the normal marginals in the left and right projections. Similarly, the right plot corresponds to a non-normal family, which shows projected non-linear regression lines and non-normal marginals.

If we assume normal conditionals alone, the resulting model is undefined, but if in addition we assume that the X|Y regression line is a proper third degree polynomial, we are in front of an inexistent or impossible model as we can conclude from Equation (10).

The simplest method to define a joint density corresponds to the independent model in which all variables are independent, so that it is sufficient to define the univariate marginals. However, when variables are dependent, the model complicates.
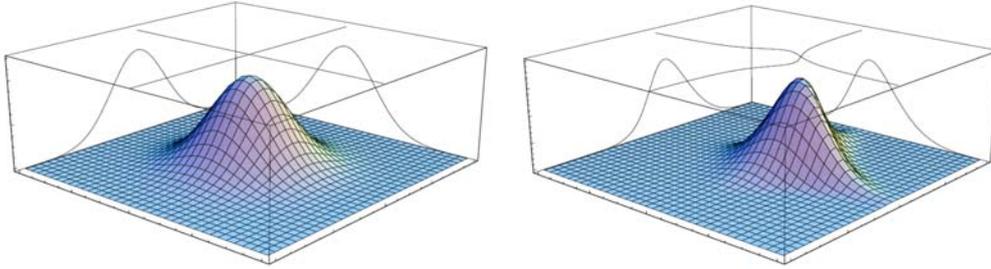
Figure 3: Two illustrative examples of bivariate densities with normal conditionals: normal (left figure) and non-normal (right figure).

Example 6.2 (Corbel Example. Selecting probability distribution families) In this example we select and discuss the probability distribution families associated with the variables involved in the corbel example introduced previously. To simplify and because the steel manufacture companies are very reliable, we can assume f_s and A_s as deterministic. Thus, we only have the random variables F_v, F_h, f_c, A_b and θ. In addition we can assume that all variables are independent. With the exception of F_v and F_h, this is a reasonable assumption because they involve forces, material strengths, areas and a design angle, whose values are undoubtedly independent. Furthermore, we assume the independence of F_v and F_h. This implies that only marginal distributions are needed in order to build the statistical multidimensional model.

Table 4 shows the selected marginal distributions, the associated parameters and the corresponding ranges.

Table 4: Corbel Example. Marginal distribution functions for the dimension variables.

Dimension Variable	Variable Type	Lower Bound	Upper Bound	Assumed Distribution	Assumed Parameters		
					Shape	Scale	Location
f_s	Deterministic	455124 kN/m²	455124 kN/m²				
A_s	Deterministic	5.92e-4 m²	5.92e-4 m²				
F_v	Random: Extreme Value	positive value	positive value	3P Max- Weibull	0.21	36.209 kN	68.846 kN
F_h	Random: Extreme Value	negative value	positive value	3P Max- Weibull	0.236	9.776 kN	−4.225kN
f_c	Random: General	positive value	positive value	Gamma	149.50	142.5 kN/m²	
A_b	Random: General	positive value	positive value	Gamma	1.45	0.031 m²	
θ	Random: General	$\pi/6$	$\pi/3$	Generalized Beta	$\alpha=2$ $\beta=15$	$\pi/6$	−1

Now, given these selected distributions, we can obtain the distributions for the associated dimensionless problem. With this aim, first we represent A_s and f_s by their expected values, i.e., $A_s \equiv E[A_s] = \mu_{A_s}$ and $f_s \equiv E[f_s] = \mu_{f_s}$, obtaining the following new dimensionless variables:

$$F_v^* = \frac{F_v}{\mu_{f_s}\mu_{A_s}}; \quad F_h^* = \frac{F_h}{\mu_{f_s}\mu_{A_s}}; \quad f_c^* = \frac{f_c}{\mu_{f_s}}; \quad A_b^* = \frac{A_b}{\mu_{A_s}},$$

and the limit-state function:

$$H^* = \frac{H}{\mu_{f_s}\mu_{A_s}} = min\left\{1 - F_v^*\tan\theta - F_h^*, \quad f_c^* A_b^* - F_v^*\right\}. \tag{14}$$

For the new non-dimensional variables appearing in equation (14), Table 5 shows the corresponding distributions and associated dimensionless parameters. The values have been obtained using $\mu_{f_s} = 455124\ kN/m^2$ *and* $\mu_{A_s} = 5.92\text{e-}4\ m^2$.

Table 5: Corbel Example. Marginal distribution functions for the dimensionless variables.

Dimensionless Variable	Assumed Distribution	Assumed Parameters		
		Shape	Scale	Location
F_v^*	3P Maximum Weibull	0.21	0.134	0.256
F_h^*	3P Maximum Weibull	0.236	0.036	−0.016
f_c^*	Gamma	149.5	3.13e−04	
f_s	Gamma	1.45	52.365	
θ^*	Generalized Beta	$\alpha=2,\ \beta=15$	$\pi/6$	−1

Table 5 shows that only the scale and location parameters are affected by the normalization. Moreover, the parameters of the Generalized Beta distribution remain constant because they are associated with the dimensionless variable $\theta^* = \theta$. *Finally, the statistical families in Table 4 remain in Table 5 because all of them are stable with respect to scale changes.*

However, there exists another way to deal with the dimensionless problem without using scale-stable distributions. The process consists of obtaining the dimensionless sample data before fitting the distribution parameters.

One of the most important methods to define dependent multivariate models is Bayesian networks, which are defined by means of a directed acyclic graph *G* together with the conditional distributions of each of the involved variables given their parents, as follows:

$$f(x_1; x_2; \ldots; x_n) = f_1(x_1) f_2(x_2|x_1) f_3(x_3|x_1; x_2) \ldots f_n(x_n|x_1; x_2; \ldots; x_{n-1}) = \prod_{i=1}^{n} f_i(x_i|\pi_i), \tag{15}$$

where π_i are the parents of the variable X_i in the directed acyclic graph *G*. Bayesian networks are the simplest way to reproduce complicated multidimensional families of distributions avoiding incompatibilities.

Example 6.3 (Corbel Example. A multivariate model) *In this example we determine a multivariate model associated to the previously dimensionless corbel example.*

From equations (4) and (5) we know that

$$H_1^* = f_1\left(F_h^*, F_v^*, f_s^*, A_s^*, \theta\right) \tag{16}$$

$$H_2^* = f_2\left(F_v^*, f_c^*, A_b^*\right), \tag{17}$$

where H_1^* and H_2^* are the dimensionless limit-state functions, using μ_{A_s} and μ_{f_s}.

In order to correctly define the joint distribution of these limit-state functions, we have to carry out an analyses of the dependence relation among the involved variables. In this example, we assume independence among all variables, except between F_v and F_h, because these forces are usually related.

Due to the fact that variables involved in H_2^* are independent we can compute the joint probability by means of the set of all marginal, that is,

$$f(H_2^*) = f(F_v^*, f_c^*, A_b^*) = f(F_v^*) f(f_c^*) f(A_b^*),\qquad (18)$$

However, in the case of the , H_1^* applying equation (15) to determine this joint probability, we get:

$$
\begin{aligned}
f(H_1^*) &= f(F_v^*, F_h^*, f_s^*, A_s^*, \theta) = \\
&= f(F_h^*) f(F_v^*|F_h^*) f(f_s^*|F_h^*, F_v^*) f(A_s^*|F_h^*, F_v^*, f_s^*) f(\theta|F_v^*, F_h^*, f_s^*, A_s^*) = \\
&= f(F_h^*) f(F_v^*|F_h^*) f(f_s^*) f(A_s^*) f(\theta),
\end{aligned}\qquad (19)
$$

which requires to know the conditional distribution of F_v^* given F_h^*.

With this aim, we represent the data (F_h^*, F_v^*) (see left Figure 4) and observe that they exhibit the following linear regression:

$$F_v^* = 3.99\, F_h^* + 0.31.\qquad (20)$$

Next, we find that the residuals follow a maximal Weibull model (see right Figure 4):

$$F_R(r) = \exp\left\{-\left[1 - k\left(\frac{r-\lambda}{\sigma}\right)\right]^{1/k}\right\}; \quad 1 - k\left(\frac{r-\lambda}{\sigma}\right) \geqslant 0.\qquad (21)$$
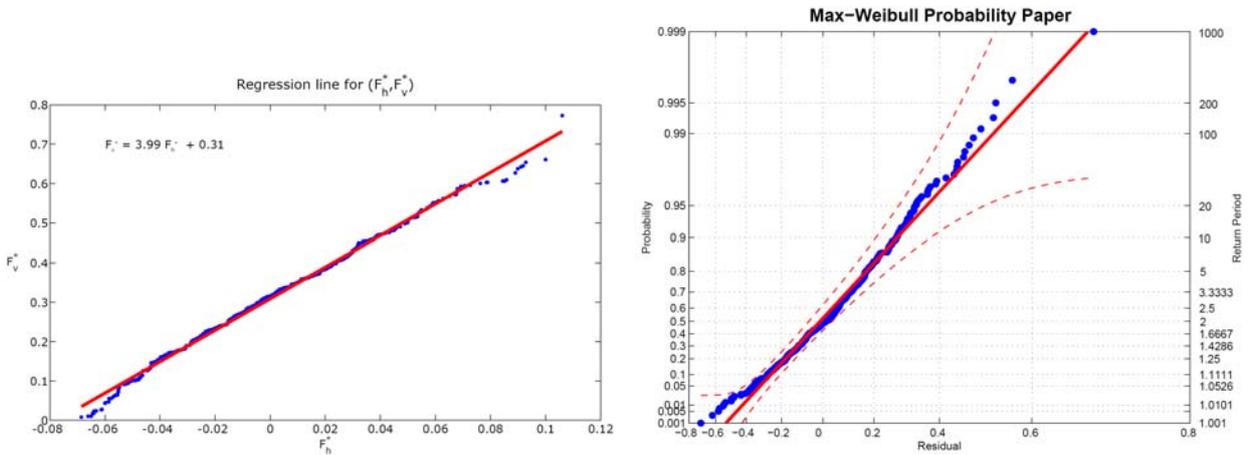
Figure 4: Data and regression line for the Corbel example and residuals given on a Normal probability plot.

Combining this expression with the regression equation (20) leads to the final model for $F_v^*|F_h^*$

$$F_{F_v^*|F_h^*}(f_v^*|f_h^*) = \exp\left\{-\left[1 - k\left(\frac{F_v^* - 3.99\, F_h^* - 0.31 - \lambda}{\sigma}\right)\right]^{1/k}\right\}\qquad (22)$$

only valid for

$$1 - k\left(\frac{F_v^* - 3.99\, F_h^* - 0.31 - \lambda}{\sigma}\right) \geqslant 0\qquad (23)$$

Then, the estimation of the Weibull parameters using the maximum likelihood method leads to

$$k=0.262; \qquad \sigma=0.218; \qquad \mu=-0.0789.$$

In this way, the joint probability of this multivariate model becomes defined and we can evaluate the failure probabilities.

Conclusions

The following conclusions can be drawn from this paper:

1. Random variables and the parameters of statistical distributions are dimensional. These must be taken into consideration when statistical models are selected, otherwise, inadequate models can be obtained leading to important dimensional problems.

2. A previous dimensional analysis of the variables involved must be performed before building a model. This leads to a deep understanding of the relations among the involved variables, avoids dimensional inconsistencies and reduces numerical precision problems. In this direction, the Π Buckingham theorem is the most convenient and recommendable tool to be used.

3. Identification of the adequate extreme value distribution is very important in real practice. There are theorems that allow us to decide which of the Weibull, Gumbel or Frechet distributions or their reverse versions corresponds to a given cdf $F(x)$.

4. We must be aware of the fact that different probability papers exists. With respect to extreme value analysis there are two Gumbel probability papers, one for maxima and one for minima. It is important to realize that only the tail of interest must be plotted and fitted.

5. Care must be taken in selecting the adequate multivariate joint density functions. In this line, we must be aware that an excess of assumptions leads to impossible models, and a lack of them, to undefined models. Finally, Bayesian networks is the most adequate method to define the joint distributions, based on a directed acyclic graph and the conditional distributions of each of the random variable given their parents.

Acknowledgements

References

O'Connor, A. J. and Kenshel, O. Experimental evaluation of the scale of fluctuation for spatial variability modeling of chloride-induced reinforced concrete corrosion, Journal Of Bridge Engineering 18(1) (2013) 3–14.

O'Connor, A. J. andEnevoldsen, I. Probability-based assessment of highway bridges according to the new danish guideline, Structure and Infrastructure Engineering 5(2) (2009) 157–168.

Simiu, E., Changery, M. and Filliben, J. J. Extreme wind speeds at 129 stations in the contiguous united states, Journal of the structural division-ASCE 106(4) (1980) 809–817.

Pourzeynali, S. and Datta, T. Reliability analysis of suspension bridges against fatigue failure from the gusting of wind, Journal of Bridge Engineering 10(3) (2005) 262–271.

Castillo E. and Fernández-Canteli, A. A Unified Statistical Methodology for Modelling Fatigue Damage, Springer, Netherlands, 2009.

Koller, R., Ruiz-Ripoll M. L., García A., Fernández-Canteli A. and Castillo, E., Experimental validation of a statistical model for the wöhler field corresponding to any stress level and amplitude, International Journal of Fatigue 31(1) (2009) 231–241.

Przybilla, C., Fernández-Canteli, A. and Castillo, E. Deriving the primary cumulative distribution function of fracture stress for brittle materials from 3- and 4-point bending tests, Journal of the European Ceramic Society 31 (2011) 451–60.

Ferreira, J. A., Guedes Soares, C. Modelling the long-term distribution of significant wave height with the Beta and Gamma models, Ocean Engineering 26(8) (1999) 713–725.

Ferreira, J., and Guedes Soares, C. An application of the peaks over threshold method to predict extremes of significant wave height, Journal of Offshore Mechanics and Arctic Engineering 120 (3) (1998) 165–176.

Ochi, M. New approach in estimating the severest sea state from statistical data, in: Proceedings on the 23rd Coastal Engineering Conference, ASCE, New York, 1992, pp. 512—523.

Lo, H. P., Zhang, N. and Lam, W.Estimation of an origin-destination matrix with random link choice proportions: a statistical approach., Transportation Research, Part B 30:4 (1996) 309–324.

Castillo, E., Nogal, M., Menéndez, J., Sánchez-Cambronero, S. and Jiménez, P. Stochastic demand dynamic traffic models using generalized beta-gaussian Bayesian networks, IEEE Transactions on Intelligent Transportation Systems 13 (2) (2012) 565–581.

Clark, S. andWatling, D. Modelling network travel time reliability under stochastic demand, European Journal of Operational Research 39 (2005) 119–140.

Castillo, E., Jiménez, P., Menéndez, J. and Nogal, M. A Bayesian method for estimating traffic flows based on plate scanning, Transportation 40 (2013) 173–201.

Buckingham, E. The principle of similitude, Nature 96 (1915) 396–397.

Castillo, E. Extreme Value Theory in Engineering, Academic Press, New York, 1988.

Arnold, B. C., Castillo, E. and Sarabia, J. M. Conditionally specified distributions, Springer-Verlag, 1992.

Arnold, B. C., Castillo, E. and Sarabia, J. M. Conditional specification of statistical models, Springer, 1999.

Arnold, B. C., Castillo, E. and Sarabia, J. M. Conditionally specified distributions: an introduction (with comments and a rejoinder by the authors), Statistical Science 16 (3) (2001) 249–274.

Castillo, E., O'Connor, A. J., Nogal, M. and Calviño, A. On the physical and probabilistic consistency of civil engineering random models, Structural Safety (2014) Submitted.

Dynamic Modeling of Tensegrity Robots Rolling over the Ground

*Shinichi Hirai and Ryo Imuta

School Department of Robotics, Ritsumeikan University, Kusatsu, Shiga 525-8577, Japan

*Corresponding author: hirai@se.ritsumei.ac.jp

Abstract

This paper presents dynamic modeling of tensegrity robots rolling over the ground. We have developed a 6-strut tensegrity robot that deforms its body for rolling locomotion over the ground. Designing tensegrity structures and control laws appropriate to locomotion experimentally has consumed much time and labor. Dynamic simulation of tensegrity robot rolling is thus required to reduce time and labor in experimental trials.

We have formulated a set of dynamic equations of motion of tensegrity robots. Our tensegrity robots consist of rigid struts and elastic actuators. Elastic actuators, which act as tensile elements, shrink by applying air pressure into the actuators. Applying air pressure to designated actuators deforms the tensegrity structure, which allows the structure roll over the ground. We have simulated the rolling of two icosahedron tensegrity robots; one consists of 24 actuators while the other consists of 12 actuators. Experimental evaluation validated our dynamic simulation.

Keywords Tensegrity, Rolling, Dynamics, Modeling, Simulation

Introduction

Locomotion has been a main research issue in robotics and many robots have been proposed in the past decade including wheel robots, crawler robots, and legged robots. Recently, much attention has been paid to soft-bodied robots, which employ deformable bodies consisting of soft materials. Such soft-bodied robots can deform their body for locomotion. Deformable bodies are useful for obstacle avoidance and narrow passage locomotion. On the other hands, it is difficult to build larger bodies since soft materials deform naturally under gravity. To cope with this problem, we have proposed to introduce tensegrity structures into robot bodies.

Tensegrity structures consists of a set of rigid elements connected by elastic elements. Rigid elements, which are referred to as struts, act as bones of a robot while elastic elements, which are referred to as tensile elements, provide softness to the robot. Tenesgrity structures have been applied robot locomotion [Aldrich *et al.* (2003); Paul *et al.* (2006); Arsenault and Gosselin (2008)]. Most tensegrity robots employ crawling for locomotion. For dynamic locomotion, we have proposed tensegtity robots that roll over the ground [Shibata *et al.* (2009)] and developed a six-strut tensegrity robot driven by pneumatic McKibben actuators [Koizumi *et al.* (2012)]. Activating an appropriate set of actuators in sequence, a tensegrity robot rolls over the ground. Unfortunately, determining a sequence of appropriate actuators for locomotion requires much time since it is performed using a real robot in trial and error manner. Additionally, we have many choices in tensegrity structures. We have to select structures appropriate to rolling locomotion. This selection would require much time.

Determining actuator sequence and selecting tensegrity structures would be performed on a computer once we have establish a dynamic simulation of rolling tensegrity structures. Thus, we will establish dynamic modeling of rolling tensegrity structures and perform simulation of rolling of tensegrity robots.

<center>(a)   (b)</center>
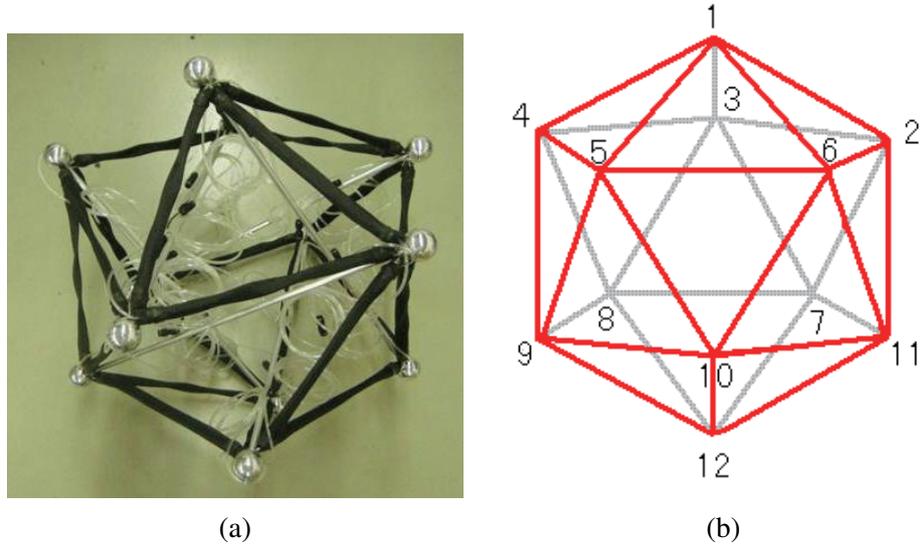
Figure 1: Prototype of six-strut tensegrity robot

Tesegrtity Robots

Figure 1-(a) shows a prototype of six-strut tensegrity robots. This prototype consists of 6 rigid struts and 24 pneumatic McKibben actuators. The struts are made of aluminum and are 570 mm in length. Two rigid balls of diameter 45mm are attached to the both ends of each strut to help the rolling of a tensegrity robot. McKibben actuators shrink by applying air pressure and extend via external forces. Namely, McKibben actuators act as elastic elements. The actuators can generate force of 800 N at air pressure of 0.5MPa. Contraction ratio is almost 34% without load and 20% under the load of 3N by at air pressure of 0.5MPa. Air pressure to the actuators is applied externally through air hoses.

Figure 1-(b) shows geometric description of a six-strut tensegrity robot. Let us attach numbers 1 through 12 to individual vertices of the tensegrity robot. Then, each strut or each actuator is specified by a pair of numbers corresponding to its both ends. A six-strut tensegrity forms an icosahedron, consisting of eight regular triangles and twelve non-regular isosceles triangles. One triangle is contacting to the ground when this tensegrity robot is in equilibrium, implying that each equilibrium can be specified by its corresponding triangle.

Figure 2 describe successive rolling of a six-strut tensegrity robot. The prototype can perform a successive rolling over a flat ground by applying air pressure to a sequence of actuator pairs.

Dynamic Modeling of Tensegrity Rolling

Let us summarize the dynamic of a rigid body. Let us attach body coordinate system $C - \xi\eta\zeta$ to the body while fix space coordinate system $O - xyz$ to space. Orientation of a rigid body is described by rotation between the two coordinate systems. Let us introduce quaternion $\boldsymbol{q} = [\, q_0, \, q_1, \, q_2, \, q_3 \,]^{\mathrm{T}}$ to describe the orientation of a rigid body [Kuipers 2002]. This quaternion must satisfy $\boldsymbol{q}^{\mathrm{T}}\boldsymbol{q} = q_0^2 + q_1^2 + q_2^2 + q_3^2 = 1$. The orientation matrix of a rigid body is then given as

$$R(\boldsymbol{q}) = \begin{bmatrix} 2(q_0^2 + q_1^2) - 1 & 2(q_1q_2 - q_0q_3) & 2(q_1q_3 + q_0q_2) \\ 2(q_1q_2 + q_0q_3) & 2(q_0^2 + q_2^2) - 1 & 2(q_2q_3 - q_0q_1) \\ 2(q_1q_3 - q_0q_2) & 2(q_2q_3 + q_0q_1) & 2(q_0^2 + q_3^2) - 1 \end{bmatrix}. \tag{1}$$
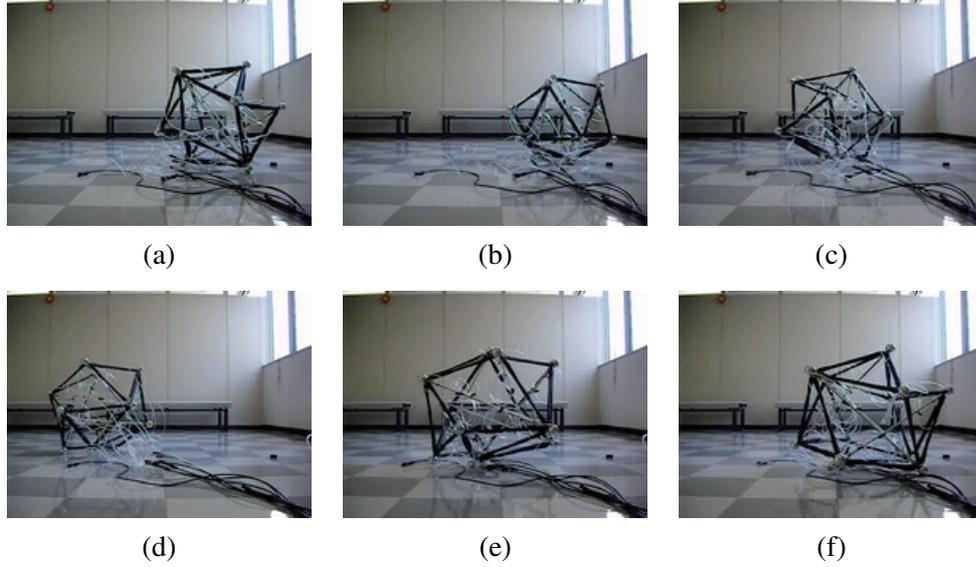
Figure 2: Successive rolling of a six-strut tensegrity robot

The first, second, and third columns of the above matrix correspond to unit vectors along ξ-, η-, and ζ-axes. Angular velocity vector of a rigid body is described as

$$\boldsymbol{\omega} = 2H(\boldsymbol{q})\,\dot{\boldsymbol{q}} = -2H(\dot{\boldsymbol{q}})\,\boldsymbol{q}, \tag{2}$$

where

$$H(\boldsymbol{q}) = \begin{bmatrix} -q_1 & q_0 & q_3 & -q_2 \\ -q_2 & -q_3 & q_0 & q_1 \\ -q_3 & q_2 & -q_1 & q_0 \end{bmatrix}.$$

Let J be inertia matrix of a rigid body and $\boldsymbol{\tau} = [\,\tau_\xi,\,\tau_\eta,\,\tau_\zeta\,]^{\mathrm{T}}$ be a set of external moments around ξ-, η-, and ζ-axes applied to the body. Then, dynamic equation of rigid body rotation is formulated as:

$$\ddot{\boldsymbol{q}} = -r(\boldsymbol{q},\dot{\boldsymbol{q}})\,\boldsymbol{q} - 2H^{\mathrm{T}}(\boldsymbol{q})\,J^{-1}\left((H(\boldsymbol{q})\dot{\boldsymbol{q}}) \times (JH(\boldsymbol{q})\dot{\boldsymbol{q}}) - \frac{1}{4}\boldsymbol{\tau}\right), \tag{3}$$

where

$$r(\boldsymbol{q},\dot{\boldsymbol{q}}) = \dot{\boldsymbol{q}}^{\mathrm{T}}\dot{\boldsymbol{q}} + 2\nu\boldsymbol{q}^{\mathrm{T}}\dot{\boldsymbol{q}} + \frac{1}{2}\nu^2(\boldsymbol{q}^{\mathrm{T}}\boldsymbol{q} - 1) \tag{4}$$

with positive constant ν. This $r(\boldsymbol{q},\dot{\boldsymbol{q}})$ originates from stabilization of holonomic constraint $\boldsymbol{q}^{\mathrm{T}}\boldsymbol{q} - 1 = 0$ [Baumgarte (1972)]. Denoting the right hand of Eq.3 by $\boldsymbol{h}(\boldsymbol{q},\dot{\boldsymbol{q}},\boldsymbol{\tau})$, dynamic equation of rigid body rotation is simply described as $\ddot{\boldsymbol{q}} = \boldsymbol{h}(\boldsymbol{q},\dot{\boldsymbol{q}},\boldsymbol{\tau})$.

Let us formulate the motion of the i-th strut. Let $2L$ be the length of the struts. Assume that each strut is uniform with its mass m and inertia matrix J. Let C_i be the center of motion of the i-th strut and $\boldsymbol{x}_i$ denote its position. Let us attach coordinate system $C_i - \xi_i\eta_i\zeta_i$ to the i-th strut. Assume that ζ_i-axis lies on the line between the both end of the strut and $\boldsymbol{\zeta}_i$ be unit vector along ζ_i-axis. Letting
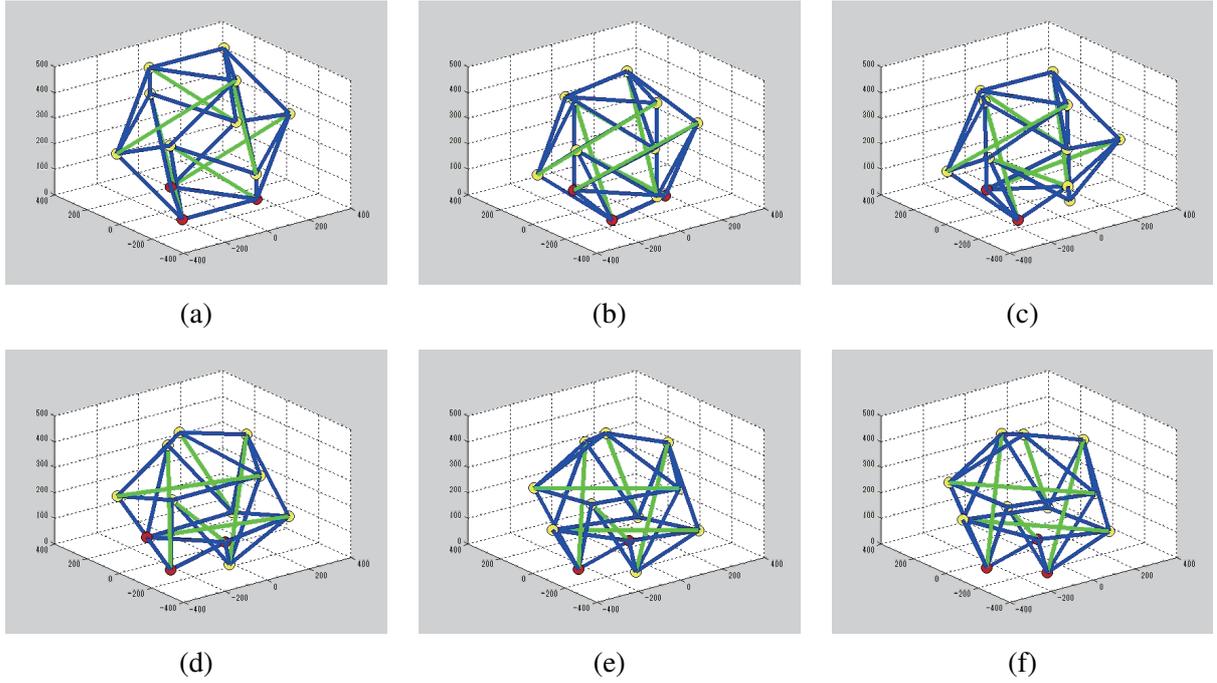
Figure 3: Simulation result of rolling of six-strut tensegrity robot

$\boldsymbol{f}_i$ and $\boldsymbol{\tau}_i$ be external force and moment applied to the i-th strut, equations of motion of the strut are given by

$$m\ddot{\boldsymbol{x}}_i = \boldsymbol{f}_i, \qquad \ddot{\boldsymbol{q}}_i = \boldsymbol{h}(\boldsymbol{q}_i, \dot{\boldsymbol{q}}_i, \boldsymbol{\tau}_i). \tag{5}$$

Recall that vertices of a six-strut tensegrity robot have their own numbers. Let R_l be a set of numbers adjacent to vertex l via elastic elements. Let $\boldsymbol{y}_l$ be the position vector of vertex l. Let j and k be vertex numbers at both end points of the i-th strut. Position vectors of the points are given by $\boldsymbol{y}_j = \boldsymbol{x}_i + L\boldsymbol{\zeta}_i$ and $\boldsymbol{y}_k = \boldsymbol{x}_i - L\boldsymbol{\zeta}_i$. Let $\boldsymbol{f}_{\mathrm{ela}}(\boldsymbol{y}_l, \boldsymbol{y}_n, \dot{\boldsymbol{y}}_l, \dot{\boldsymbol{y}}_n)$ be viscoelastic force generated by an elastic element connecting vertices l and n. Then, the resultant force applied to vertex j is formulated as

$$\boldsymbol{f}_i^+ = \sum_{l \in R_j} \boldsymbol{f}_{\mathrm{ela}}(\boldsymbol{y}_j, \boldsymbol{y}_l, \dot{\boldsymbol{y}}_j, \dot{\boldsymbol{y}}_l). \tag{6}$$

Similarly, the resultant force applied to vertex k is given by

$$\boldsymbol{f}_i^- = \sum_{l \in R_k} \boldsymbol{f}_{\mathrm{ela}}(\boldsymbol{y}_k, \boldsymbol{y}_l, \dot{\boldsymbol{y}}_k, \dot{\boldsymbol{y}}_l). \tag{7}$$

Additionally, we will apply penalty method to formulate contact forces from the ground. Assuming that the ground is specified by $z \leq 0$, contact force applied to a vertex of which position is represented as $\boldsymbol{x} = [\, x,\, y,\, z \,]^{\mathrm{T}}$ is given by

$$\boldsymbol{f}_{\mathrm{con}}(\boldsymbol{x}) = \begin{cases} 0 & (z > 0) \\ -Kz - C\dot{z} & (z \leq 0) \end{cases}, \tag{8}$$

(a-1) start   (a-2) end

(a) experiment

(b-1) start   (b-2) end

(b) simulation
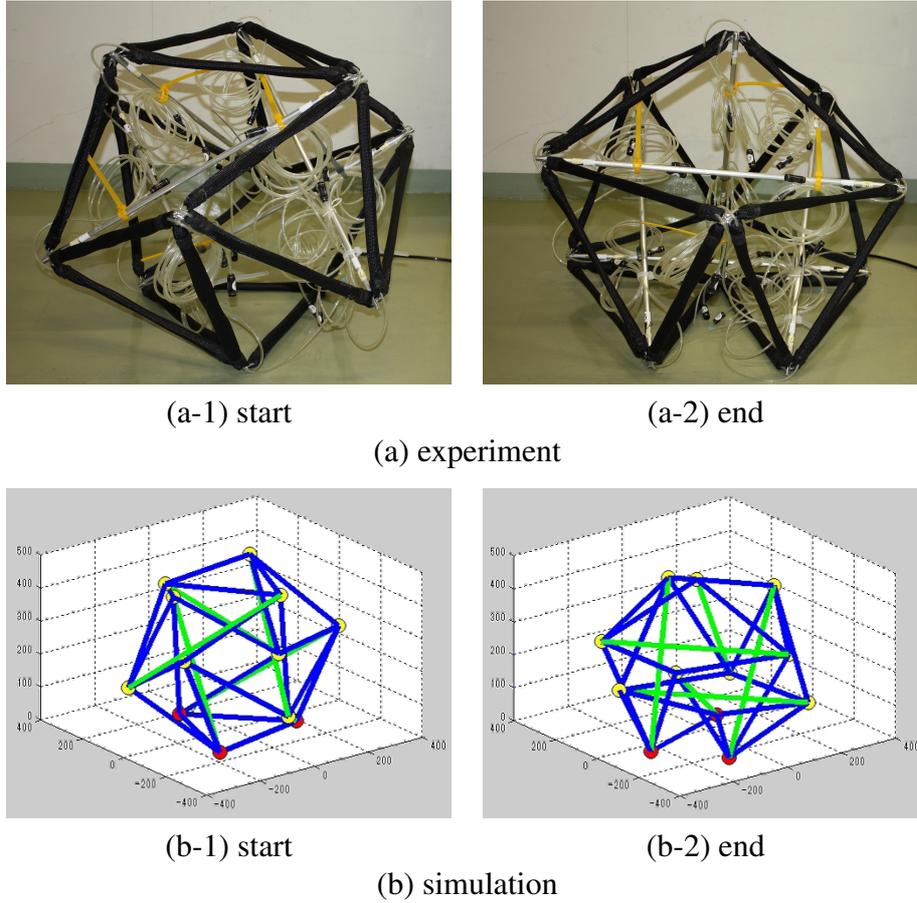
Figure 4: Six-strut tensegrity robot rolling from planar symmetric contact

where K and C represent elastic and viscous coefficients of the ground. Contact forces applied to vertices j and k is then formulated as $\boldsymbol{f}_{\mathrm{con}}(\boldsymbol{y}_j)$ and $\boldsymbol{f}_{\mathrm{con}}(\boldsymbol{y}_k)$.

Consequently, the resultant force and moment applied to the i-th strut are formulated as:

$$\boldsymbol{f}_i = \boldsymbol{f}_i^+ + \boldsymbol{f}_i^- + \boldsymbol{f}_{\mathrm{con}}(\boldsymbol{y}_j) + \boldsymbol{f}_{\mathrm{con}}(\boldsymbol{y}_k) + m\boldsymbol{g}, \tag{9}$$

$$\boldsymbol{\tau}_i = (L\boldsymbol{\zeta}_i) \times \left(\boldsymbol{f}_i^+ - \boldsymbol{f}_i^- + \boldsymbol{f}_{\mathrm{con}}(\boldsymbol{y}_j) - \boldsymbol{f}_{\mathrm{con}}(\boldsymbol{y}_k) \right), \tag{10}$$

where $\boldsymbol{g}$ represents the acceleration of gravity. Solving equations of motion of all struts numerically, we can simulate the motion and deformation of a tensegrity robot.

Simulation Results

We have performed dynamic simulation of rolling of a six-strut tensegrity robot. Figure 3 shows a sequence of snapshots of a result. Red circles describe vertices contacting to the ground while yellow ones are not in contact with the ground. At first, a regular triangle is in contact with the ground (Figure 3-(a)). Then, the body deforms (Figure 3-(b)) and one vertex of the regular triangle loses its contact (Figure 3-(c)). Namely, the tensegrity robot is out of equilibrium, yielding rotation around the line between the two contacting points (Figure 3-(d) and (e)). Finally, one vertex contacts to the ground, resulting that a non-regular isosceles triangle is in contact with the ground (Figure 3-(f)).

(a-1) start   (a-2) end

(a) experiment

(b-1) start   (b-2) end

(b) simulation
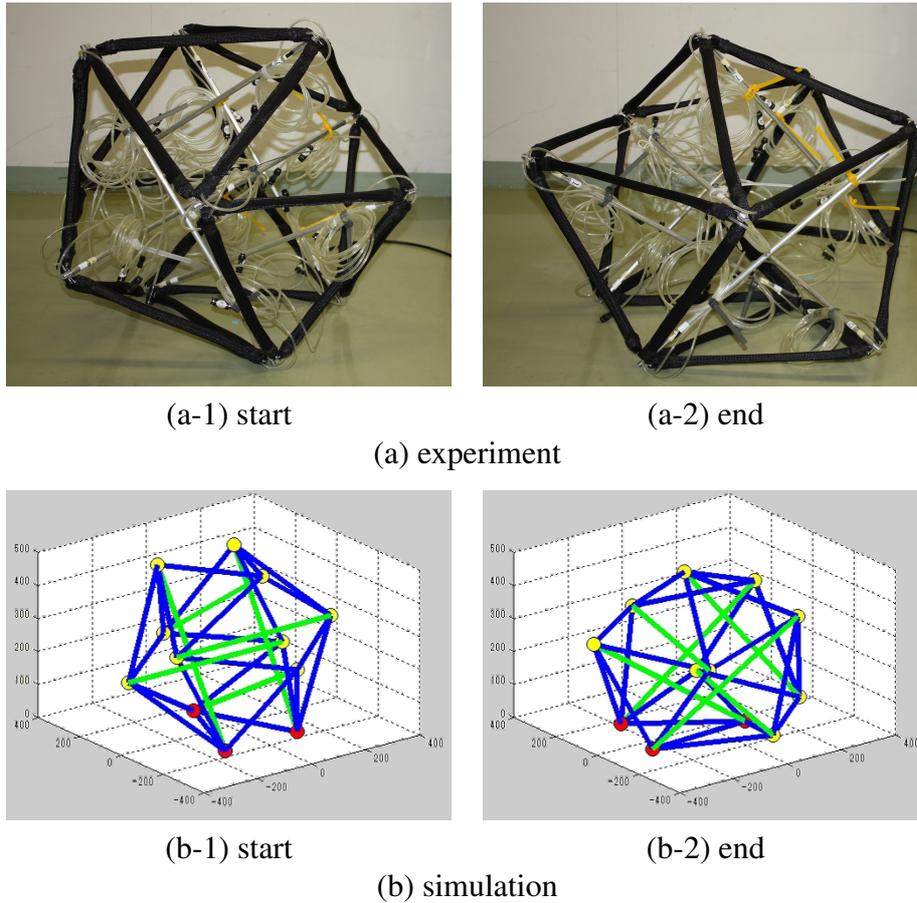
Figure 5: Six-strut tensegrity robot rolling from planar symmetric contact

Contact between a six-strut tensegrity robot and the ground can be specified by the triangle contacting to the ground. Contact specified by a regular triangle is referred to as *axial symmetric contact* while contact represented by a non-regular isosceles triangles is referred to as *planar symmetric contact*. Note that we have eight axial symmetric contacts and twelve planar symmetric contacts. Rolling of a six-strut tensegrity robot corresponds to a sequence of transitions among the twenty contacts.

We have found that driving a pair of pneumatic McKibben actuators yields 1) transition from axial symmetric contact to its neighboring planar symmetric contact, or 2) transition from planar symmetric contact to its neighboring planar symmetric contact [Koizumi *et al.* (2012)]. Let us examine if the above two transitions can be simulated. Figure 4 shows experimental and simulation results of transition from axial to planar symmetric contacts. We have found that experimental and simulation results meet well. Figure 5 shows experimental and simulation results of transition from planar contact to its neighboring planar symmetric contact. The simulation result agrees with the experimental result. As a result, we conclude that dynamic simulation of rolling tensegrity robots works well.

Let us simulate the rolling of another tensegrity structure. Figure 6 shows a star-shaped tensegrity structure. This structure consists of six struts and twelve actuators. Each actuator connects one end point of a strut and the center of another strut. Note that no actuators contact to the ground during rolling. We have simulated transitions from axial and planar symmetric contacts. Figure 7
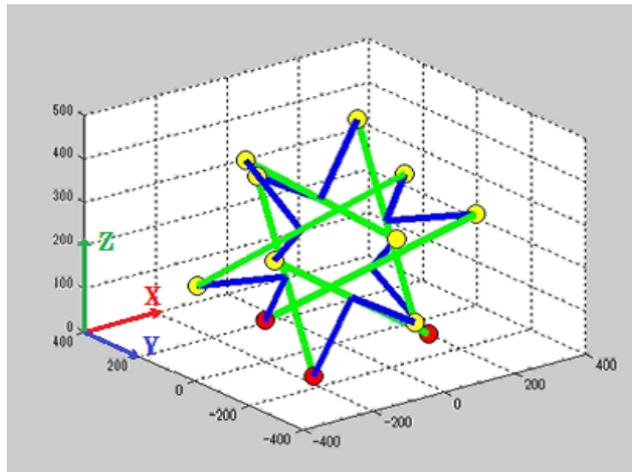
Figure 6: Star-shaped tensegrity robot
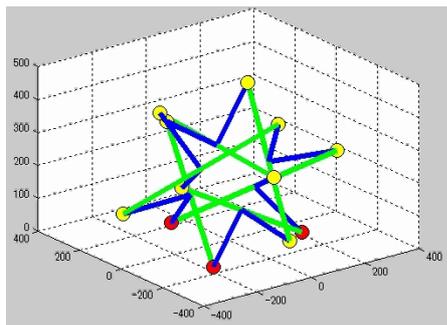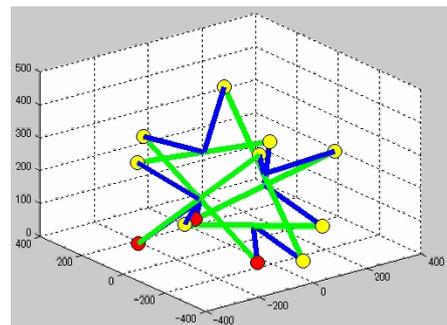
(a-1) start   (a-2) end

(a) transition from axial symmetric contact
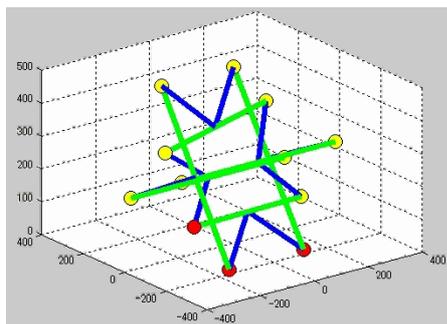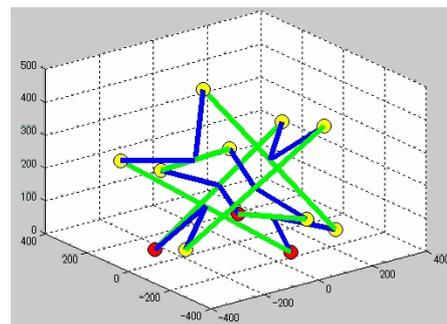
(b-1) start   (b-2) end

(b) transition from planar symmetric contact

Figure 7: Simulation results of rolling of star-shaped tensegrity robot

shows simulation results. Figure 7-(a) shows an axial symmetric contact transits to its neighboring planar symmetric contact. Figure 7-(b) describes a planar symmetric contact transits to its neighboring axial symmetric contact. These results suggest that this star-shaped tensegrity robot can perform rolling from any contact to another.

Conclusion

We have established dynamic simulation of tensegrity robot rolling. It turns out that rolling of a six-strut tensegrity robot can be simulated and simulation results agree with experimental results. Additionally, we have simulated the rolling of a star-shaped tensegrity robot. Through simulation, we have found that this tensegrity robot can perform rolling locomotion.

Acknowledgments

References

Aldrich, J. B., Skelton, R. E., and Kreutz-Delgado, K. (2003) Control Synthesis for a Class of Light and Agile Robotic Tensegrity Structures, Proc. American Control Conference, pp. 5245–5251.

Paul, C., Valero-Cuevas, F. J., and Lipson, H. (2006) Design and Control of Tensegrity Robots for Locomotion, *IEEE Trans. on Robotics*, Vol. 22, No. 5, pp.944–957.

Arsenault M. and Gosselin, C. M. (2008) Kinematic and Static Analysis of a Three-degree-of-freedom Spatial Modular Tensegrity Mechanism, *Int. Journal of Robotics Research*, Vol. 27, No. 8, pp.951–966.

Shibata, M. Saijyo, F. and Hirai, S. (2009) Crawling by Body Deformation of Tensegrity Structure Robots, *Proc. IEEE Int. Conf. on Robotics and Automation*, pp.4375–4380, Kobe, May 12–17.

Koizumi, Y., Shibata, M., and Hirai, S. (2012) Rolling Tensegrity Driven by Pneumatic Soft Actuators, *Proc. IEEE Int. Conf. on Robotics and Automation*, pp.1988–1993, St. Paul, U.S.A., May 14–18.

Kuipers, J. B. (2002) Quaternions and Rotation Sequences, Princeton University Press.

Baumgarte, J. (1972) Stabilization of constraints and integrals of motion in dynamical systems, *Computer Methods in Applied Mechanics and Engineering*, 1:1–16, .

Solving 2D multi-crack problems with arbitrary distribution by virtual boundary meshless least squares method

*Qiang Xu¹

¹Department of Building Engineering, Tongji University, Shanghai 200092, China.

*Corresponding author: xuqiang@tongji.edu.cn

Abstract

This paper is about how solving two dimensional multi-crack problems with arbitrary distribution by the virtual boundary meshless least squares method. In this article, the local domain where a single crack is contained would be treated as twain subdomain when solving multi-crack problem. And this method incorporates the point interpolation method (PIM) with the compactly supported radial basis function (CSRBF) often used in boundary-type meshless methods to approximately construct the virtual source function on the virtual boundary corresponding to each subdomain. According to the definition about sub-domain in this paper, the added extra sub-domains on the boundary extended along the crack surface as "conventional sub-domain method" in the direct boundary element method do not have to be considered, thereby reducing the computational, especially avoiding this calculation error caused due to inadequate number of the elements or with the collocation points configured on the boundary of the additional sub-domains and its improper configuration. In addition, since the configuration of virtual boundary has a certain preparability, the integration along the virtual boundary can be carried out over the smooth simple curve that can be structured beforehand (for 2D problems) to reduce the complicity and difficulty of calculus without loss of accuracy, while ''Vertex Question'' existing in BEM can be avoided.

Keywords: Virtual boundary, Meshless, Least squares, Radial basis function, Multi-crack

1. Introduction

Generally speaking, crack, multi-crack or micro-crack are pre-existed in engineering components [1] and structures [2], brittle or quasi-brittle materials, and so on. As is known to all, the stress intensity factor can be used to describe the stress field of the crack tip and predict crack growth in fracture mechanics. So the stress intensity factor for the calculation of a crack or multi-crack analysis also is very important. In fact, the equation with solving crack problem is easily established, but the exact solution is quite difficultly obtained, especially in multi-crack problems. That is, the analytic methods, such as the westergaard method [3], the complex variable function method [4], conformal mapping [5] and so on, can only solve the simple or regular crack problem, and complex or irregular crack problems need resort to numerical methods.

The boundary element method (BEM) is an important kind of numerical methods, and it is suitable for analyzing a large field gradient function of the problem and also can better calculate the stress concentration. Some scholars solve the crack problems by the BEM, such as Z.H. Yao, P.B. Wang and H.T. Wang et al. [6-7] use dual BEM to analyze the numerous micro-cracks, Q.H. Qin and Y.W. Mai [8] employ the BEM for crack-hole problems in thermopiezoelectric materials, E.D. Leonel and W.S. Venturini [9] use the dual boundary element formulation to analysis of multi-fractured domains, X.Q. Yan [10] analyzes the stress intensity factors of multiple circular arc cracks in a plane elasticity plate by employing the BEM, and so on. But the BEM still has its own drawbacks. It is shown that the coefficient matrix is asymmetric and its construction is time consuming in the process of calculation. In addition, the treatment of singular integration is inconvenient and takes much more time, and there is mimicry singular integral whilst reducing calculating precision, especially when solving the related physics quantities on the boundary, which

is called "Boundary Layer Effect". Fortunately, the virtual boundary element least square method proposed by the authors in the literature [11-19] can avoid above described drawbacks in BEM, since its coefficient matrix is symmetric and it does not involve singular integral.

According to the definition about sub-domain in this paper, the added extra sub-domains on the boundary extended along the crack surface as "conventional sub-domain method" in the direct boundary element method do not have to be considered, thereby reducing the computational, especially avoiding this calculation error caused due to inadequate number of the elements or with the collocation points configured on the boundary of the additional sub-domains and its improper configuration. In addition, since the configuration of virtual boundary has a certain preparability, the integration along the virtual boundary can be carried out over the smooth simple curve that can be structured beforehand (for 2D problems) to reduce the complicity and difficulty of calculus without loss of accuracy, while ''Vertex Question'' existing in BEM can be avoided.

The rest of the paper is organized as follows. In Section 2, the radial point interpolation method and the stress intensity factor are described in brief. In Section 3, the calculation scheme for solving multi-crack problems by virtual boundary meshless least squares method is derived in detail, and the related processing technologies in the calculation of multi-crack problems are introduced. In Section 4, numerical examples are presented to demonstrate the efficiency and validity of the method proposed in the paper. Finally, a summary is given in Section 5 to conclude this paper.

2. The radial point interpolation method and the stress intensity factor

2.1 The radial point interpolation method

Consider a scalar function $u(\boldsymbol{x})$ defined in the problem domain Ω, the approximation function $u^h(\boldsymbol{x})$ of u(x) is represented by a set of scattered interpolation nodes, and $u^h(\boldsymbol{x})$ can approximates $u(\boldsymbol{x})$ at a point of interest $\boldsymbol{x}$. By the reference [16], the RPIM function interpolation expression augmented with polynomials can be rewritten as [20-27]

$$u^h(\boldsymbol{x}) = \sum_{i=1}^{n} R_i(\boldsymbol{x})a_i + \sum_{j=1}^{m} p_j(\boldsymbol{x})b_j = \mathbf{R}^{\mathrm{T}}(\boldsymbol{x})\mathbf{a} + \mathbf{p}^{\mathrm{T}}(\boldsymbol{x})\mathbf{b} \tag{1}$$

In which, $R_i(\boldsymbol{x})$ is the compactly supported radial basis function (CSRBF) proposed by Wu [28], n is the number of RBFs in the defined domain of the calculation point $\boldsymbol{x}$, namely

$$R_i(\boldsymbol{x}) = \tilde{r}^4(4 + 16r + 12r^2 + 3r^3) \tag{2}$$

where $r = \|\boldsymbol{x} - \boldsymbol{x}_i\|/d_i$, in which d_i is the dimension of the local support domain for CSRBF (shown in Fig. 1), and

$$\tilde{r} = \begin{cases} (1-r) & \text{for } 0 \leq r \leq 1 \\ 0 & \text{other} \end{cases} \tag{3}$$

and $p_j(\boldsymbol{x})$ is polynomial basis functions in two-dimensional coordinates $\boldsymbol{x}^{\mathrm{T}} = [x, y]$, m is the number of polynomial basis functions. The vector $\mathbf{a}$ of coefficients for RBFs is $\{a_1 \quad a_2 \quad \cdots \quad a_n\}^{\mathrm{T}}$, and the vector $\mathbf{b}$ of coefficients for polynomial is $\{b_1 \quad b_2 \quad \cdots \quad b_m\}^{\mathrm{T}}$. Then, coefficients a_i and b_j are constants yet to be determined.

In order to determine a_i and b_j in Eq. (1), a support domain is formed for the point of interest at $\boldsymbol{x}$, and n field nodes are included in the support domain. Coefficients a_i and b_j in Eq. (1) can be determined by enforcing Eq. (1) to be satisfied at these n nodes surrounding the point of interest $\boldsymbol{x}$. Then, the $n+m$ unknown coefficients a_i and b_j in Eq. (1) can be obtained by the following equations

$$\sum_{i=1}^{n} R_i(\boldsymbol{x}_I)a_i + \sum_{j=1}^{m} p_j(\boldsymbol{x}_I)b_j = u_s(\boldsymbol{x}_I) \qquad I = 1,2,\cdots,n \tag{4}$$

$$\sum_{I=1}^{n} p_j(\boldsymbol{x}_I)a_I = 0 \qquad j = 1,2,\cdots,m \tag{5}$$

where Eq. (5) is m constraint conditions added using the orthogonality between $p_j(\boldsymbol{x}_I)$ ($I=1,2,\cdots,n$) and $\mathbf{a}$ to solve $n+m$ variables in Eq. (1). Combing Eqs. (4) and (5) yields the following set of equations in the matrix form

$$\mathbf{B}\begin{Bmatrix}\mathbf{a}\\\mathbf{b}\end{Bmatrix} = \begin{Bmatrix}\mathbf{u}_s\\\mathbf{0}\end{Bmatrix} \tag{6}$$

where the moment matrix $\mathbf{B}$ is

$$\mathbf{B} = \begin{bmatrix} R_1(\boldsymbol{x}_1) & \cdots & R_n(\boldsymbol{x}_1) & p_1(\boldsymbol{x}_1) & \cdots & p_m(\boldsymbol{x}_1) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ R_1(\boldsymbol{x}_n) & \cdots & R_n(\boldsymbol{x}_n) & p_1(\boldsymbol{x}_n) & \cdots & p_m(\boldsymbol{x}_n) \\ p_1(\boldsymbol{x}_1) & \cdots & p_1(\boldsymbol{x}_n) & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ p_m(\boldsymbol{x}_1) & \cdots & p_m(\boldsymbol{x}_n) & 0 & \cdots & 0 \end{bmatrix} \tag{7}$$

Using Eq. (6) and by Eq. (1), we can obtain

$$u^h(\boldsymbol{x}) = \left\{\mathbf{R}^{\mathrm{T}}(\boldsymbol{x}) \quad \mathbf{p}^{\mathrm{T}}(\boldsymbol{x})\right\}\mathbf{B}^{-1}\begin{Bmatrix}\mathbf{u}_s\\\mathbf{0}\end{Bmatrix} = \tilde{\mathbf{N}}^{\mathrm{T}}(\boldsymbol{x})\begin{Bmatrix}\mathbf{u}_s\\\mathbf{0}\end{Bmatrix} \tag{8}$$

where the vector $\mathbf{u}_s$ of function values is $\{u_1 \quad u_2 \quad \cdots \quad u_n\}^{\mathrm{T}}$, and the RPIM shape functions can be expressed as

$$\tilde{\mathbf{N}}^{\mathrm{T}}(\boldsymbol{x}) = \left\{\mathbf{R}^{\mathrm{T}}(\boldsymbol{x}) \quad \mathbf{p}^{\mathrm{T}}(\boldsymbol{x})\right\}\mathbf{B}^{-1}$$
$$= \left\{N_1(\boldsymbol{x}) \quad N_2(\boldsymbol{x}) \quad \cdots \quad N_n(\boldsymbol{x}) \quad N_{n+1}(\boldsymbol{x}) \quad \cdots \quad N_{n+m}(\boldsymbol{x})\right\} \tag{9}$$

And the RPIM shape functions $\mathbf{N}(\boldsymbol{x})$ corresponding to the nodal displacements vector $\mathbf{u}_s$ are obtained as

$$\mathbf{N}^{\mathrm{T}}(\boldsymbol{x}) = \left\{N_1(\boldsymbol{x}) \quad N_2(\boldsymbol{x}) \quad \cdots \quad N_n(\boldsymbol{x})\right\} \tag{10}$$

Therefore, Eq. (8) can be rewritten as

$$u^h(\boldsymbol{x}) = \mathbf{N}^{\mathrm{T}}(\boldsymbol{x})\mathbf{u}_s = \sum_{i=1}^{n} N_i u_i \tag{11}$$

In this paper, in order to further improve the accuracy and computational efficiency, the idea of RPIM with CSRBF is incorporated to approximately construct the virtual source function $\varphi_k(\boldsymbol{\xi})$ ($k=1,2$) in VBEM.
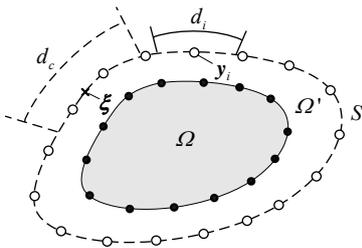
| **Figure 1.** Computing model of single domain problem for the meshless VBEM | **Figure 2.** Local coordinate description of the crack tip displacement | **Figure 3.** Diagrammatic sketch of multi- |
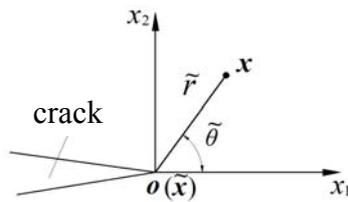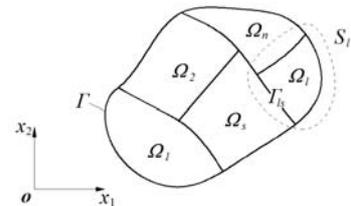
2.2 The stress intensity factor

In a given 2D Cartesian coordinate system, for the composite crack problems with containing Ⅰ and Ⅱ type, the displacement u_1 and u_2 along x_1、x_2 direction at crack tip point $\tilde{x}$ can be expressed respectively as (shown in Fig. 2)

$$u_1 = \frac{K_\mathrm{I}}{2G}\sqrt{\frac{\tilde{r}}{2\pi}}\cos\frac{\tilde{\theta}}{2}\left(\kappa-1+2\sin^2\frac{\tilde{\theta}}{2}\right) + \frac{K_\mathrm{II}}{2G}\sqrt{\frac{\tilde{r}}{2\pi}}\sin\frac{\tilde{\theta}}{2}\left(\kappa+1+2\cos^2\frac{\tilde{\theta}}{2}\right) \tag{12}$$

$$u_2 = \frac{K_\mathrm{I}}{2G}\sqrt{\frac{\tilde{r}}{2\pi}}\sin\frac{\tilde{\theta}}{2}\left(\kappa+1-2\cos^2\frac{\tilde{\theta}}{2}\right) + \frac{K_\mathrm{II}}{2G}\sqrt{\frac{\tilde{r}}{2\pi}}\cos\frac{\tilde{\theta}}{2}\left(-\kappa+1+2\sin^2\frac{\tilde{\theta}}{2}\right) \tag{13}$$

In which, K_I and K_II are the stress intensity factors respectively corresponding to Ⅰ and Ⅱ type; $\tilde{r}$ is the distance between the calculation point x and the crack tip point $\tilde{x}$, $\tilde{\theta}$ is the angle between the radial vector $\tilde{r}$ and the axis x_1; in addition, $G = E/2(1+v)$, and $\kappa = 3 - 4v$ (Plane strain problem), $\kappa = (3-v)/(1+v)$ (Plane stress problem).

By Eqs. (12) and (13), when $\tilde{\theta}$ turn counterclockwise an angle π, the displacement along x_1、x_2 direction at the crack tip point $\tilde{x}$ can be denoted respectively as

$$u_{1+} = \frac{K_\mathrm{II}}{2G}\sqrt{\frac{\tilde{r}}{2\pi}}(\kappa+1), \quad u_{2+} = \frac{K_\mathrm{I}}{2G}\sqrt{\frac{\tilde{r}}{2\pi}}(\kappa+1) \tag{14}$$

and, when $\tilde{\theta}$ turn clockwise an angle π, the corresponding displacement is

$$u_{1-} = -\frac{K_\mathrm{II}}{2G}\sqrt{\frac{\tilde{r}}{2\pi}}(\kappa+1), \quad u_{2-} = -\frac{K_\mathrm{I}}{2G}\sqrt{\frac{\tilde{r}}{2\pi}}(\kappa+1) \tag{15}$$

By Eqs. (14) and (15) simultaneous solution, K_I and K_II can be obtained, namely

$$K_\mathrm{I} = \frac{(u_{2+}-u_{2-})G}{\kappa+1}\sqrt{\frac{2\pi}{\tilde{r}}}, \quad K_\mathrm{II} = \frac{(u_{1+}-u_{1-})G}{\kappa+1}\sqrt{\frac{2\pi}{\tilde{r}}} \tag{16}$$

For ease of comparison to the stress intensity factor K by different loads and under different geometric scales, need to introduce dimensionless stress intensity factor F. From the reference [29], $F = K/(\sigma\sqrt{\pi a})$, in which σ is a reference stress and a is crack half-length.

3. The calculation scheme for solving multi-crack problems

The idea for solving the multi-domain problems (shown in Fig. 3) by virtual boundary meshless least squares method has been given in the reference [16], and the virtual source function on the virtual boundary about each subdomain is approximately constructed by the boundary-type radial point interpolation method [20]. Here, according to basic idea of the literature [16], the idea for solving multi-crack problems is presented. According to the literature [16], there is

$$\boldsymbol{J}[\varphi^{(1)}(\boldsymbol{\xi}), \varphi^{(2)}(\boldsymbol{\xi}), \cdots, \varphi^{(n)}(\boldsymbol{\xi})] =$$

$$\sum_{l=1}^{n}\sum_{i=1}^{2}\left\{\sum_{j=1}^{N_u^l}\alpha_l\left[u_i^{(l)}(\boldsymbol{x}_j)-\overline{u}_i^{(l)}(\boldsymbol{x}_j)\right]^2\bigg|_{\boldsymbol{x}_j\in\Gamma_l^u\cap\Gamma} + \sum_{j=N_u^l+1}^{N_u^l+N_P^l}\beta_l\left[p_i^{(l)}(\boldsymbol{x}_j)-\overline{p}_i^{(l)}(\boldsymbol{x}_j)\right]^2\bigg|_{\boldsymbol{x}_j\in\Gamma_l^p\cap\Gamma}\right\} +$$

$$\sum_{G(\Gamma_{ks})}\sum_{i=1}^{2}\left\{\sum_{j=1}^{M_G}\alpha_J\left[u_i^{(k)}(\boldsymbol{x}_j)-u_i^{(s)}(\boldsymbol{x}_j)\right]^2\bigg|_{\boldsymbol{x}_j\in\Gamma_{ks}^G} + \sum_{j=1}^{M_G}\beta_J\left[p_i^{(k)}(\boldsymbol{x}_j)+p_i^{(s)}(\boldsymbol{x}_j)\right]^2\bigg|_{\boldsymbol{x}_j\in\Gamma_{ks}^G}\right\} \tag{17}$$

In which, $\boldsymbol{J}[\varphi^{(1)}(\boldsymbol{\xi}), \varphi^{(2)}(\boldsymbol{\xi}), \cdots, \varphi^{(n)}(\boldsymbol{\xi})]$ is the square deviation functional of the multi-domain composite problem about virtual source function $\varphi_k^{(i)}(\boldsymbol{\xi})$ ($k = 1, 2$; $i = 1, 2, \cdots, n$), and the above equation established must satisfies the given actual boundary conditions of the original problem and the corresponding connection conditions between the adjacent subdomains at the same time. n is

the number of subdomains within the problem given. $\Gamma_l^u \cap \Gamma$ and $\Gamma_l^p \cap \Gamma$ are called the exterior displacement and traction boundary of Ω_l, respectively, N_u^l is the number of real boundary nodes whose displacement is known on the boundary $\Gamma_l \cap \Gamma$ (namely $\Gamma_l^u \cap \Gamma$), while N_p^l the nodes number of known traction with outward normal direction on the boundary $\Gamma_l^p \cap \Gamma$; $\bar{u}_i^{(l)}(x_j)$ and $\bar{p}_i^{(l)}(x_j)$ are the known displacement and traction values along i^{th} direction at the jth boundary point x_j on the exterior boundary $\Gamma_l^u \cap \Gamma$ and $\Gamma_l^p \cap \Gamma$ of lth subdomain, respectively; α_l and β_l are weight coefficient about the displacement and traction of the lth subdomain, respectively; G is the overall serial number of the current internal boundary Γ_{ks} (namely Γ_{ks}^G), M_G is the over-fulfilled collocation number on Γ_{ks}^G, α_J and β_J are weight coefficient about the displacement and traction on Γ_{ks}^G, respectively. By solving Eq. (17), the virtual source function $\varphi_k^{(l)}(\xi)$ ($k=1,2$; $l=1,2,\cdots,n$) can be obtained. As soon as the virtual source function $\varphi_k^{(l)}(\xi)$ is known, the corresponding physical value at x about each subdomain can be calculated through the following equations, namely

$$u_i^{(l)}(x) = \int_{S_l} u_{ik}^*(x,\xi)\varphi_k^{(l)}(\xi)dS \tag{18}$$

$$\varepsilon_{ij}^{(l)}(x) = \int_{S_l} \varepsilon_{ijk}^*(x,\xi)\varphi_k^{(l)}(\xi)dS \tag{19}$$

$$\sigma_{ij}^{(l)}(x) = \int_{S_l} \sigma_{ijk}^*(x,\xi)\varphi_k^{(l)}(\xi)dS \tag{20}$$

$$p_i^{(l)}(x) = \int_{S_l} \sigma_{ijk}^*(x,\xi)n_j\varphi_k^{(l)}(\xi)dS \tag{21}$$

It can be known from Eqs. (18) and (21), Eq. (17) is the square deviation functional about virtual source function $\varphi_k^{(i)}(\xi)$ ($k=1,2$; $i=1,2,\cdots,n$) that is the unknown function on the virtual boundary S_i. Unlike the conventional VBEM, 'virtual boundary meshless' mentioned in the paper means that the virtual source approximate functions $\tilde{\varphi}_k^{(i)}(\xi)$ ($k=1,2$; $i=1,2,\cdots,n$) constructed in the meshless VBEM are not dependent upon the geometric mesh generation of the computing element. That is to say, the idea of RPIM with CSRBF is incorporated to approximately construct the virtual source function. And a kind of background-mesh can be employed about the numerical integration of virtual source function in the method. Such as Eqs. (18) and (21), the virtual boundary S_l will be separated into m_e elements and there are e_g Gauss spots within each element, then their Gauss numerical integral of $u_i^{(l)}(x)$ and $p_i^{(l)}(x)$ can be expressed respectively as

$$u_i^{(l)}(x) \approx \int_{S_l} u_{ik}^*(x,\xi)\tilde{\varphi}_k^{(l)}(\xi)dS = \sum_{e=1}^{m_e^l}\sum_{g=1}^{e_e^l} \tilde{w}(\xi_g^e)u_{ik}^*(x,\xi_g^e)\tilde{\varphi}_k^{(l)}(\xi_g^e)$$
$$= \sum_{e=1}^{m_e^l}\sum_{g=1}^{e_e^l} \tilde{w}(\xi_g^e)u_{ik}^*(x,\xi_g^e)\mathbf{N}^{(l)\mathrm{T}}(\xi_g^e)\cdot {}_s\tilde{\boldsymbol{\varphi}}_k^{(l)} \tag{22}$$

$$p_i^{(l)}(x) \approx \int_{S_l} \sigma_{ijk}^*(x,\xi)n_j\tilde{\varphi}_k^{(l)}(\xi)dS = \sum_{e=1}^{m_e}\sum_{g=1}^{e_g} \tilde{w}(\xi_g^e)\sigma_{ijk}^*(x,\xi_g^e)n_j\tilde{\varphi}_k^{(l)}(\xi_g^e)$$
$$= \sum_{e=1}^{m_e}\sum_{g=1}^{e_g} \tilde{w}(\xi_g^e)\sigma_{ijk}^*(x,\xi_g^e)n_j\mathbf{N}^{(l)\mathrm{T}}(\xi_g^e)\cdot {}_s\tilde{\boldsymbol{\varphi}}_k^{(l)} \tag{23}$$

Substitute Eq. (22) and Eq. (23) into Eq. (17), then we can see that $J[\varphi^{(1)}(\xi),\cdots,\varphi^{(n)}(\xi)]$ will change as the vector of the entire virtual source function values ${}_s\tilde{\boldsymbol{\varphi}}^{(l)}$ ($l=1,2,\cdots,n$), so $J[\varphi^{(1)}(\xi),\cdots,\varphi^{(n)}(\xi)]$ can be expressed as $J[{}_s\tilde{\boldsymbol{\varphi}}^{(1)}, {}_s\tilde{\boldsymbol{\varphi}}^{(2)},\cdots, {}_s\tilde{\boldsymbol{\varphi}}^{(n)}]$. In order to obtain the solution ${}_s\tilde{\boldsymbol{\varphi}}^{(l)}$ ($l=1,2,\cdots,n$), the variation is being made for Eq. (25), namely

$$\delta J = 0 \tag{24}$$

Let Ns be node number on all virtual boundaries of multi-domain complex issues, then the

unknown vector of entire node function values A can be recorded as a unified manner, namely

$$A = \left\{ {}_s\tilde{\boldsymbol{\varphi}}^{(1)T}, {}_s\tilde{\boldsymbol{\varphi}}^{(2)T}, \cdots, {}_s\tilde{\boldsymbol{\varphi}}^{(n)T} \right\}^T = \left\{ {}_1\tilde{\varphi}_1 \quad {}_1\tilde{\varphi}_2 \quad {}_2\tilde{\varphi}_1 \quad {}_2\tilde{\varphi}_2 \quad \cdots \quad {}_{Ns}\tilde{\varphi}_1 \quad {}_{Ns}\tilde{\varphi}_2 \right\}^T$$

$$= \left\{ \tilde{\varphi}_1 \quad \tilde{\varphi}_2 \quad \cdots \quad \tilde{\varphi}_{2Ns} \right\}^T \tag{25}$$

From Equation (24), we get the governing equation for solving the unknown virtual source function values at all nodes on all virtual boundaries. Equation (24) can be rewritten in matrix form as follows [16]

$$KA = B \tag{26}$$

In the above equation, $K = [k_{st}]_{2Ns \times 2Ns}$ is a symmetric coefficient matrix; A is the vector of the unknown virtual source function values $(2Ns \times 1)$; B is the right term that can be obtained based on the boundary conditions.

Definition of crack subdomain

Suppose that is the number m of cracks contained in whole domain Ω, and the local domain of each crack is treated as two sub-domains. So whole domain is artificially divided into $2m+1$ sub-domains, in which the boundary Γ_+^i and Γ_-^i of each crack is respectively referred to as the upper and lower boundaries of the crack (or left and right boundaries), as shown in Figure 4. Respective sub-domain can be artificially defined corresponding to the boundary Γ_+^i or Γ_-^i of each crack, namely sub-domain Ω_+^i corresponding to Γ_+^i. That is, there is twin sub-domain corresponding to each crack. And sub-domain Ω^* is considered as "Substrate domain" in addition to the crack sub-domains defined.

On numerical implementation of this method terms, the definition of the configuration shape of the crack has a certain degree of arbitrariness, such as rectangular or semi-circular or semi-elliptical and so on. And the selection of its shape and size hardly affects the result of the calculation. Compared with "conventional sub-domain method" in the direct boundary element method [30], the added extra sub-domains on the boundary extended along the crack surface do not have to be considered according to the definition about sub-domain in the paper, thereby reducing the computational, especially avoiding this calculation error caused due to inadequate number of the elements or with the collocation points configured on the boundary of the additional sub-domains and its improper configuration.

Element division near the crack tip by equal proportions

Due to stress gradient with larger changes on the vicinity of the crack tip, therefore, how the elements are reasonably distributed on the actual boundary near the crack tip to get a better numerical solution will be very important. Through numerical integration practice, we decorate the desired nodes in order to define the elements required near the crack tip for the numerical integration based on "proportional" mathematical ideas in the paper, thereby decide the division and distribution of element near the crack tip. The ideas of the specific implementation: suppose that a is crack half-length, and denote $a = \overline{BA}$. Here, point A is regarded as the crack tip, the desired nodes a_i are arranged from point B to point A according to the calculation formula $a_i = (1 - q^i)a/(1 - q^n)$ $(i = 1, 2, \cdots, n; \ q \neq 1)$, in which q is scale factor. When $i = n$, the position corresponding to a_n is the crack tip A.

4. Numerical Examples

4.1 There is a through-wall crack with crack half-length $a = 1 \text{ mm}$ at the center position of the plate of side length $l = 100 \text{ mm}$, as shown in Figure 5. The load of perpendicular to upper and lower plate

edge is uniform distribution of tension, and it is expressed as $\sigma = 1\,\text{MPa}$. Young's modulus of elasticity $E=2.1\times10^6\,\text{MPa}$, Poisson's ratio v=0.3.

Now, employ the method proposed in this article to calculate the above issues. The problem can be artificially divided into three sub-domains, namely one referred to as the subdomain Ω^* of "Substrate domain" and the other two sub-domains Ω^1_+ and Ω^1_- belonging to crack definition, as shown in Figure 6. For crack sub-domain Ω^1_+, 22 elements are divided on the crack boundary Γ^1_+ based on "proportional" mathematical ideas, and three internal boundaries corresponding to Ω^1_+ take straight line edge and its each one is divided into 20 elements. Then there are a total of 104 elements on all actual boundaries corresponding to Ω^1_+. However, for the virtual boundary S^1_+ of Ω^1_+, employ a circle configuration of radius $r_1 = 1.22\,\text{mm}$, and 45 elements are evenly distributed on S^1_+. However, for sub-domain Ω^1_-, its actual and virtual boundaries are divided by using the same method as the processing subdomain Ω^1_+. For the subdomain Ω^* of "Substrate domain", its outer boundary Γ_S is formed by the four straight line edges and 20 elements on each one are evenly distributed, then the element implementation of its internal boundary, namely element number and distribution, should be the same as the division corresponding to crack subdomain. However, for Ω^*, exist simultaneously inner virtual boundary S^*_{in} and external one S^*_{ex}, and their configuration all are a circle with the radius $r_3 = 72\,\text{mm}$ and $r_4 = 1.9\,\text{mm}$, respectively. And 45 elements are evenly distributed on S^*_{in} or S^*_{ex}. All nodes on inner and external virtual boundaries add up to 180, namely DOF number is 360.
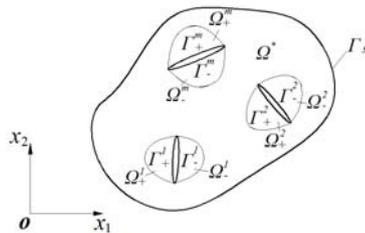
Figure 4. Diagrammatic sketch of containing multiple cracks
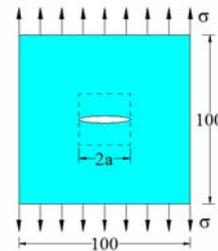
Figure 5. Calculation diagram of single crack
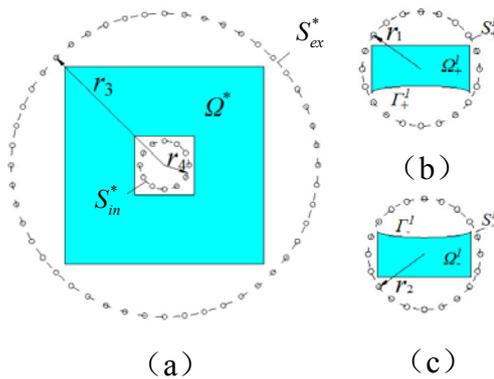
Figure 6. Discretization of single crack
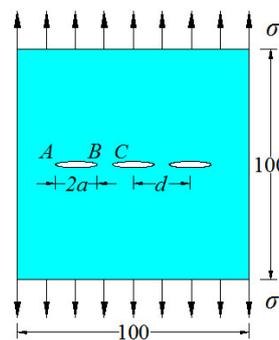
Figure 7. Calculation figure of three collinear cracks

With employing Westergaard' stress function method [29] to calculate the stress intensity factor of this example, the analytical solutions obtained is $\sqrt{\pi}\,\text{MPa}$ ($=1.7725\,\text{MPa}$), and the numerical result in accordance with the method proposed in the paper for solving the stress intensity factor is $1.7758\,\text{MPa}$, its numerical error is 1.86‰. However, under the same degree of freedom, the comparison of the numerical results of the literature [31] about BEM J integral method and

displacement discontinuity method with the analytical solutions and the numerical solutions in the paper is shown in Table 1.

4.2 There are three collinear through-wall cracks with crack half-length $a = 1$ mm of each crack for the plate with side length $l = 100$ mm, as shown in Figure 7. The load of perpendicular to upper and lower plate edge is uniform distribution of tension, and it is expressed as $\sigma = 1$ MPa. Young's modulus of elasticity E=2.1×10^6 MPa, Poisson's ratio v=0.3. In addition, the center distance of adjacent cracks denotes d, and the distance between adjacent crack tips is $2a/d$.

Now, employ the method proposed in this article to calculate the above issues. The problem can be artificially divided into seven sub-domains, namely one referred to as the subdomain Ω^* of "Substrate domain" and the other six sub-domains Ω^i_+ and Ω^i_- (i=1,2,3) belonging to crack definition. Then the whole region Ω is divided into seven sub-domains, namely Ω_i ($i = 1, 2, \cdots, 7$), and the discrete processing program on actual and virtual boundary of each subdomain is same as one of the previous example. However, the comparison of the numerical results of calculating the normalized stress intensity factor K_A, K_B, K_C at crack tip *A*, *B*, *C*. by the method proposed in the paper with them of employing the BEM with 45 displacement discontinuity element[32] and based on the stress function method[33] is shown in Table 2. And by comparison with the literature [32], the method proposed has fewer degrees of freedom for calculating the same problem, thus calculation efficiency can be improved. And the results of the method proposed is numerically more to be close to them of the literature [33] by comparison with the literature [32].

Tab.1 the stress intensity factor for the single crack

K	Analytical solution	J integral method [31]	Displacement discontinuity method [31]	The method of this paper
result (MPa)	1.7725	1.7867	1.9303	1.7758
error	——	8 ‰	8.9 %	1.86 ‰

Tab.2 the normalized stress intensity factor for three collinear cracks

F / $2a/d$	K_A			K_B			K_C		
	The paper	[32]	[33]	The paper	[32]	[33]	The paper	[32]	[33]
0.05	0.99885	0.9961	1.00083	0.99851	0.9961	1.00040	0.99876	0.9963	1.00063
0.1	0.99951	0.9972	1.00150	0.99959	0.9973	1.00164	1.00058	0.9982	1.00252
0.2	1.00370	1.0015	1.00585	1.00483	1.0026	1.00702	1.00799	1.0059	1.01030
0.3	1.01067	1.0085	1.01296	1.01480	1.0126	1.01710	1.02170	1.0195	1.02407
0.4	1.02058	1.0184	1.02297	1.03115	1.0288	1.03353	1.04285	1.0405	1.04529
0.5	1.03381	1.0317	1.03631	1.05666	1.0540	1.05913	1.07405	1.0714	1.07663
0.6	1.05118	1.0490	1.05383	1.09643	1.0932	1.09915	1.12046	1.1171	1.12316
0.7	1.07447	1.0722	1.07724	1.16144	1.1571	1.16456	1.19261	1.1881	1.19558
0.8	1.10743	1.1049	1.11032	1.27891	1.2724	1.28348	1.31668	1.3104	1.32136
0.9	1.16105	1.1581	1.16439	1.55727	1.5405	1.56454	1.59952	1.5835	1.60685

5. Conclusions

1. The ideas of the virtual boundary meshless least squares method with radial point interpolation have been formulated for solving multi-crack problems. However, the given numerical examples indicate its high accuracy and high efficiency.

2. The point interpolation scheme with compactly supported radial basis function is introduced into the method so that no element mesh is required in this method. Consequently, this method has the advantages of boundary-type meshless methods. It can be used for the calculation and analysis of complex question.

3. By comparison of "conventional sub-domain method" in the direct boundary element method, the added extra sub-domains on the boundary extended along the crack surface do not have to be considered according to the definition about sub-domain in the paper, thereby reducing the computational, especially avoiding this calculation error caused due to inadequate number of the elements or with the collocation points configured on the boundary of the additional sub-domains and its improper configuration.

References

[1] Z.H. Shi, M. Ohtsu, M. Suzuki, Y. Hibino, Numerical analysis of multiple cracks in concrete using the discrete approach, J. Struct. Eng-Asce, 127(9) (2001) 1085-1091.

[2] H. Qing, W. Yang, J. A. Lu, Numerical Simulation of Multiple Cracking in ANSI 304 Stainless Steel Under Thermal Fatigue, Int. J. Damage Mech, 19(7) (2010) 767-785.

[3] H.M. Westergaard, Bearing pressures and cracks, J. Appl. Mech, 6(2) 1939 49-53.

[4] N.I. Muskhelishvili, J. Radok, Some basic problems of the mathematical theory of elasticity, Cambridge Univ Press, 1953.

[5] O.L. Bowie, Analysis of an infinite plate containing radial cracks originating at the boundary of an internal circular hole, J. Math. Phys, 35(1) (1956) 60-71.

[6] P.B. Wang, Z.H. Yao, T. Lei, Analysis of solids with numerous microcracks using the fast multipole DBEM, Cmc-Comput Mater Con, 3(2) (2006) 65-75.

[7] P.B. Wang, Z.H. Yao, H.T Wang, Fast multipole BEM for simulation of 2-D solids containing large numbers of cracks, Tsinghua Sci \& Technol, 10(1) (2005) 76-81.

[8] Q.H. Qin, Y.W. Mai, BEM for crack-hole problems in thermopiezoelectric materials, Eng. Fract. Mech, 69(5) (2002) 577-588.

[9] E.D. Leonel, W.S. Venturini, Dual boundary element formulation applied to analysis of multi-fractured domains, Eng Anal Bound Elem, 34(12) (2010) 1092-1099.

[10] X. Q. Yan, A boundary element analysis for stress intensity factors of multiple circular arc cracks in a plane elasticity plate, Appl. Math. Model, 34(10) (2010) 2722-2737.

[11] Q. Xu, H.C. Sun, Virtual boundary element least square method for solving three dimensional problems of thick shell. J Dalian Univ Technol 36 (1996) 413-418. in Chinese.

[12] Q. Xu, H.C. Sun, Virtual boundary element least-square collocation method, Chin J Comput Mech,14 (1997) 166-73. in Chinese.

[13] H.C Sun, L.Z Zhang, Q. Xu, Y.M Zhang, Nonsingularity boundary element methods, Dalian: Dalian Univ. Techno. Press, (1999). in Chinese.

[14] Q. Xu, Three-Dimension virtual Boundary element method for solving the problems of thin shells, Chin J Appl Mech, 17 (2000) 111-115. in Chinese.

[15] Q. Xu, H.C Sun, Unified way for dealing with three-dimensional problems of solid elasticity, Appl Math Mech, 22 (2001) 1357-1367.

[16] Q. Xu, Z.J. Zhang, W. Si, Virtual boundary meshless least square collocation method for calculation of 2D multi-domain elastic problems, Eng Anal Bound Elem, 36(5) (2012) 696-708.

[17] D.S. Yang, Q. Xu, Virtual boundary meshless least square integral method with moving least squares approximation for 2D elastic problem, Eng Anal Bound Elem, 37(3) (2013) 616-623.

[18] Q. Xu, A new high-precision boundary-type meshless method for calculation of multidomain steady-state heat conduction problems, Numerical Heat Transfer, Part B, 64 (2013) 388-406.

[19] Q. Xu, D.S. Yang, Solving multi-domain 2D heat conduction problems by the least squares collocation method with RBF interpolation on virtual boundary, Eng Anal Bound Elem, 42 (2014) 37-44.

[20] G. R. Liu, Y. T. Gu, Boundary meshfree methods based on the boundary point interpolation methods, Eng Anal Bound Elem, 28(5) (2004) 475-487.

[21] Gu Y.T., Liu G.R., A boundary point interpolation method for stress analysis of solids, Computational Mechanics. 28, 2002, pp. 47-54.

[22] Gu, Y.T., Liu G.R., A boundary radial point interpolation method (BRPIM) for 2D meshless structure analysis, Structural engineering and mechanics, vol. 15, pp. 535-550, 2003.

[23] Liu G.R., A Point Assembly Method for Stress Analysis for Solid, in Impact Response of Materials & Structures, V. P. W. Shim et al. eds, Oxford, pp. 475-480, 1999.

[24] Wang J.G., Liu G.R., Radial point interpolation method for elastoplastic problems. Proc. of the 1st Int. Conf. On Structural Stability and Dynamics, Dec. 7-9, 2000, Taipei, Taiwan, pp. 703-708, 2000.

[25] Liu G.R., Mesh Free Methods: Moving Beyond the Finite Element Method. CRC press, Boca Raton, USA, 2002.

[26] Gu Y.T., Liu G.R., Using radial function basis in a boundary-type meshless method, boundary point interpolation method (BPIM), Inter. Conf. On Sci. & Engr. Comput., March 19-23, 2001a, Beijing. 68.

[27] Gu Y.T., Liu G.R., A boundary point interpolation method (BPIM) using radial function basis, First MIT Conference on Computational Fluid and Solid Mechanics, June 2001b, MIT, pp. 1590-1592.

[28] Wu Z.M., Compactly supported positive definite radial function, Advance in computational mathematics, vol. 4, pp. 283-292, 1995.

[29] China Aviation Research Institute, Handbook of stress intensity factor, Beijing: Science Press, 1993: 3, 68,134. in Chinese.

[30] LI Zailiang, WANG Cheng. Higher Boundary Element Method, Beijing: Science Press, 2008: 262-263. in Chinese.

[31] Tuhkuri J. Dual boundary element analysis of closed cracks[J]. International Journal for Numerical Methods in Engineering, 1997, 40(16): 2995-3014.

[32] Yan X., Interaction of arbitrary multiple cracks in an infinite plate, The Journal of Strain Analysis for Engineering Design, 2004, 39(3): 237-244.

[33] Chen Y Z., Stress intensity factors for curved and kinked cracks in plane extension, Theoretical and Applied Fracture Mechanics, 1999, 31(3): 223-232.

Hydro-geochemical Model for Treating Acidic Groundwater using a Permeable Reactive Barrier

Udeshini Pathirage[1], and *Buddhima Indraratna[2]

[1]Associate Research Fellow, Centre for Geomechanics and Railway Engineering, University of Wollongong, Wollongong City, NSW 2522, Australia.
[2]Professor of Civil Engineering and Research Director, Centre for Geomechanics and Railway Engineering; Program Leader, ARC Centre of Excellence for Geotechnical Science and Engineering; University of Wollongong, Wollongong City, NSW 2522, Australia.

*Corresponding author: indra@uow.edu.au

Abstract

This study couples geochemistry with geo-hydraulics to enable time-dependent modelling for the treatment of acidic groundwater using an alkaline permeable reactive barrier (PRB). This is the first such model developed for acidic groundwater generated from acid sulfate soil which is useful to address the adverse effects such as massive fish kills, corrosion of concrete and steel structures and unfavourable conditions for vegetation. The remediation process has been successful to date with a slight decrease in efficiency caused due to chemical clogging by secondary mineral precipitates, which reduces the porosity and hydraulic conductivity of the reactive medium. To predict these changes numerically, governing equations were incorporated into finite different codes, MODFLOW and RT3D. An original geochemical algorithm was developed for RT3D to simulate chemical reactions occurring in the PRB. The results and the model predictions are in agreement, confirming the hydraulic conductivity reduction due to mineral precipitation was only 3% at the entrance phase of PRB.

Keywords: Geochemical algorithm, Finite different modeling, Hydraulic conductivity reduction, Permeable reactive barrier, Acid sulfate soil

Introduction

The Australian coastal floodplains have been undergoing environmental and socio-economic problems due to the acid sulfate soils (ASS), which are spread over 3 million hectares of land (White et al., 1997). The oxidation of pyritic soil is increased due to the growing population and resulting changes in land use such as construction of deep flood mitigation drains which promote the generation of sulfuric acid. One of the injurious processes resulting is the leaching of heavy metals such as aluminium (Al) and iron (Fe) to adjacent water bodies. The acid sulfate soil research team at the University of Wollongong has been examining several engineering solutions such as weirs and floodgates, which have been mounted near Broughton Creek, Shoalhaven Floodplain, New South Wales (NSW) (Indraratna et al., 2001). These methods were capable to prevent pyrite oxidation, but were unable to treat prevailing acidity deposited in the soil (Indraratna et al., 2005). A noteworthy progress was achieved through a pilot-scale permeable reactive barrier (PRB) installed at Manildra Group's Environmental Farm in October 2006. Recycled concrete aggregates were utilised as the reactive material to neutralise the acidic groundwater (Golab et al., 2006). This PRB is a promising technique for sustaining a groundwater pH from alkaline to neutral (pH 10.0-7.2) and removing dissolved Al and total Fe below 2 and 0.5 mg/L, respectively inside the PRB. Even though the total performance was acceptable, a slight decrease in pH inside the PRB was seen

due to armouring of the reactive media by the Al- and Fe precipitates in the forms of oxy/hydroxide minerals. According to Regmi et al., (2011), laboratory column experiments have indicated a 50% reduction in the actual acid neutralisation capacity (ANC) of the recycled concrete compared to its theoretical ANC.

This study presents the model developed to couple geochemistry with geo-hydraulics in the remediation of acidic groundwater, which has not been carried out in the past, especially for time-dependent modeling and performance verification. The commercial numerical codes, MODFLOW and RT3D were used to couple groundwater flow with reaction kinetics. An original geochemical algorithm was developed to feed into RT3D which captures the geochemical reactions taking place in the PRB (Indraratna et al., 2014). This model is advantageous to practicing environmental engineers and scientists who have to work with the acid sulfate soil related problems.

Methodology

The pilot-scale PRB

In October 2006, the pilot-scale PRB (17.7 m long, 1.2 m wide and 3 m deep) was mounted at the study site, such that it intersects the region of maximum groundwater flow. Geo-textile material was used to shield the reactive media from physical clogging by soil and other fine debris inflowing the barrier. There are 36 observation wells and 15 piezometers installed inside, up-gradient and down-gradient of the PRB (Figure 1) to observe the phreatic surface variations, groundwater chemistry and hydraulic gradients.
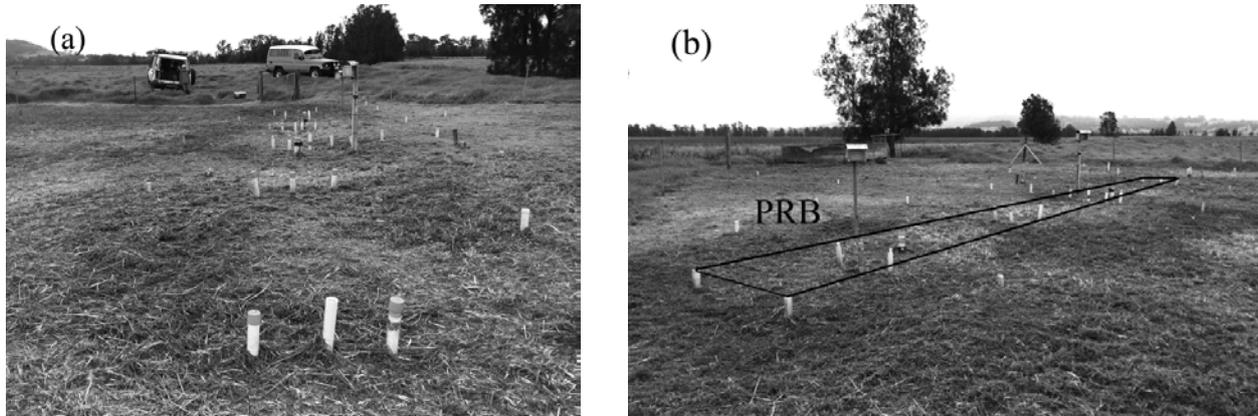
Figure 1. (a) and (b) Pilot-scale PRB and the observation wells and piezometers at study site (Pathirage et al., 2014)

Monthly field visits were carried out to monitor the water quality parameters in monitoring wells such as pH, ORP, and temperature and groundwater elevation in piezometers. These were directly measured onsite every month from the installation time to date. Groundwater samples were analysed according to APHA (1998) guidelines for acidic cations (total Fe, Al^{3+}), major cations (Ca^{2+}, Mg^{2+}, Na^+, K^+), anions (Cl^- and SO_4^{2-}), acidity and alkalinity. More elaborated details of the contents discussed in the paper can also be found in previous publications of the second Author and his research students in Computers and Geotechnics and ASCE Journals. Authors acknowledge Elsevier Publication for allowing permission to re-use some of technical content published in Computers and Geotechnics.

Mathematical model

The transition state theory (Eqn. 1) was used to develop the geochemical algorithm. This algorithm comprised of most leading reactions taking between acidic groundwater and the alkaline minerals existing in the recycled concrete.

$$r = -k_{eff}\left(1 - \frac{IAP}{K_{eq}}\right)$$
(1)

where, r is the rate of mineral precipitation ($r > 0$) or dissolution ($r < 0$), k_{eff} is the effective rate coefficient, IAP is the ion activity product, and K_{eq} is the solubility constant for the reaction.

PRHEEQC software was used to calculate the saturation indices (SI) useful to compute the values for IAP/K_{eq} as shown in Equation 2. The effective rate coefficient (k_{eff}) for each substance in Equation 1 was a calibrated value and expected to be spatially invariant and time independent during the simulation (Table 1) (Indraratna et al., 2014).

$$SI = \log(IAP) - \log(K_{eq})$$
(2)

Table 1. Calibrated k_{eff} values from the data provided by Regmi et al., (2011)

Mineral phase	k_{eff} (mol/L.s)
Ca^{2+}	2.27×10^{-7}
Al^{3+}	6.86×10^{-8}
Total Fe (Fe^{2+} and Fe^{3+})	5.87×10^{-8}

The precipitated secondary minerals which accumulate on the recycled concrete were assumed to be immobile. The pore space occupied by each mineral was calculated from their molar volume. Hence, the total volume change was calculated by the volume employed by mineral precipitates minus the volume gained by the dissolution of alkaline minerals in recycled concrete. The resulting porosity change was calculated using Equation 3 (Indraratna et al., 2014).

$$n_t = n_0 - \sum_{k=1}^{N_m} M_k R_k t$$
(3)

where, n_0, n_t are the porosities initially and at time t respectively, M_k is the mineral molar volume, R_k is the overall reaction rate of the mineral, N_m is the number of minerals from 1 to k.

The associated change in hydraulic conductivity was calculated by the normalised Kozeny Carmen equation (Eqn. 4) (Pathirage et al., 2012, Li et al., 2006).

$$K = K_0 \left[\frac{n_0 - \Delta n_t}{n_0}\right]^3 / \left[\frac{1 - n_0 + \Delta n_t}{1 - n_0}\right]^2$$
(4)

where, K_0 and K are hydraulic conductivities initially and at time t, Δn_t is the change in porosity at time t.

The finite different numerical codes: MODFLOW and RT3D were coupled to simulate the transport of the main cations in the field. MODFLOW did not update the associated change in hydraulic conductivity due to mineral precipitation/dissolution at each time step. In order to enable that, a

mathematical model was established to capture the changing hydraulic conductivity via the change in head (h) (Eqn. 5). The methodology for obtaining Equation 5 is elaborated in Indraratna et al., (2014).

$$h = \left(\exp\left[-\frac{\mu^2 b K_0}{S \sum_{k=1}^{N_m} M_k R_k} \frac{(1-n_0)^2}{n_0^3} \left\{ \begin{array}{c} \alpha^2 \left(1.5 + \frac{1}{\beta}\right) \\ -3(\alpha + \ln \beta) \end{array} \right\} \right] \right) \cdot (C \sin \mu x + D \cos \mu x) \tag{5.a}$$

$$\alpha = n_0 + \sum_{k=1}^{N_m} M_k R_k t \tag{5.b}$$

$$\beta = 1 - n_0 - \sum_{k=1}^{N_m} M_k R_k t \tag{5.c}$$

where, b is the aquifer thickness, S is the storage co-efficient, and μ, C, D are integral constants.

It was important to update the porosity and hydraulic conductivity changes at each time step due to the changes in volume fractions of precipitated and dissolved minerals. For instance, when the simulation was carried out for the 1st time step, the resulting porosity and hydraulic conductivity should be updated for the 2nd time step. This was capable through Equation 5 which determines the resulting head as that was a vital input for MODFLOW to carry on the simulation for subsequent time steps. The developed geochemical algorithm fed into RT3D was coupled with MODFLOW by the advection, diffusion and dispersion equation (Eqn. 6). MODFLOW and RT3D were run contemporarily to obtain the concentrations of reactants at each time step. In this study, dissolution/precipitation reactions were taken into account by replacing λ with, r in Equation 1 multiplied by M.

$$R \frac{\partial C}{\partial t} = D \frac{\partial^2 [C]}{\partial x^2} - v \frac{\partial [C]}{\partial x} - \lambda C \tag{6}$$

where, R is the retardation coefficient, C is the aqueous species concentration, D is the hydrodynamic dispersion coefficient, v is the seepage velocity and λ is the first-order decay constant.

Model application to PRB

The reactive transport analysis was carried out along the centreline of the PRB representing the whole PRB. The width of the PRB (1.2 m) was discretised into 0.1 m intervals (Figure 2). The acidic groundwater inflowing the PRB was expected to be in chemical equilibrium. The flow domain was a fully saturated system with specified head boundaries and a mean hydraulic gradient of 0.006, observed according to the field data from 2006 to 2012. The top, bottom and lateral faces of the flow domain were no-flow boundaries.
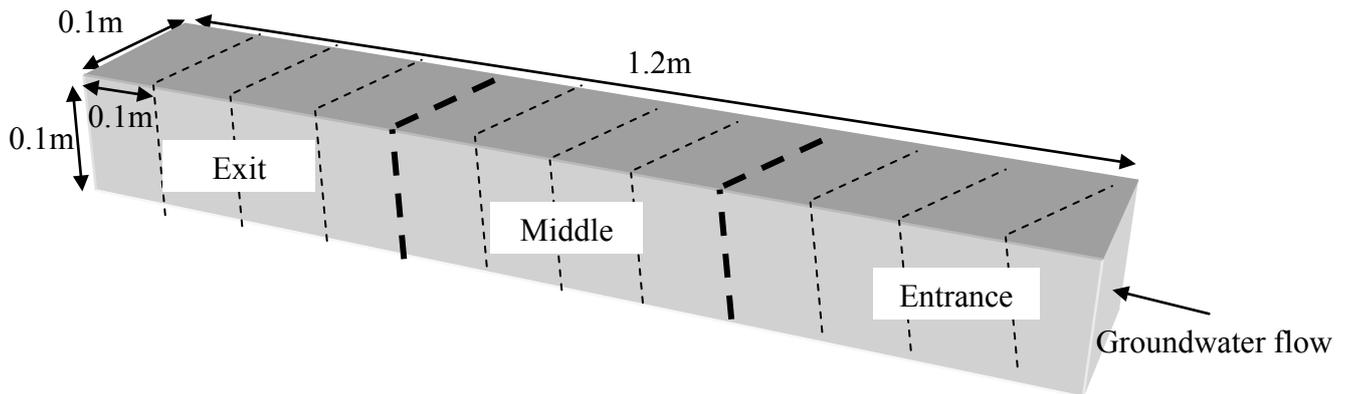
Figure 2. Discretisation of the centreline of the PRB

Results and Discussion

The model output presented the profiles of pH, Al and total Fe concentrations. Figure 3 shows the favourable comparisons found between the model predictions and field measurements for pH, Al and total Fe concentrations for 2012, which is after 6 years of installation of the PRB. In 2012, the up-gradient groundwater pH fluctuated between 3.2 and 4.1 with an average of 3.6, while the pH inside the PRB was neutral, ranging from 6.7 to 7.4 with an average of 7. It is obvious from both field measurements and simulation results that the pH at the entrance zone of the PRB is lower than that of at the middle and exit zones. This is possibly because of the exhaustion of alkaline material of the reactive media during the neutralisation of acidic groundwater. Additionally, the armouring and clogging of the reactive media due to the secondary mineral precipitation is also accountable for the small decrease in pH at the entrance of the PRB.

The highest Al^{3+} concentration detected up-gradient of the PRB was 32 mg/L for 2012. Both field monitoring results and simulation results show that the Al^{3+} reduced promptly within the PRB to less than 1 mg/L. Similarly, high concentrations of total Fe in the up-gradient of the PRB were also detected with a rapid decrease to less than 0.5 mg/L within the PRB. The rapid decrease in Al and total Fe shows that they are precipitated in their oxy/hydroxide forms.

The porosity and hydraulic conductivity would be decreased as a disadvantage of the secondary minerals precipitation inside the PRB. Nevertheless, the computed decrease in hydraulic conductivity for last six years was only 3%, which is not substantial because of the granular sized recycled concrete aggregates (d_{50}=40 mm) used in the PRB. Moreover, this slight decrease is supported by the detected stable piezometric heads within the PRB over the last 6 year monitoring period, which shows no risk of clogging prolong to failure of the PRB (Figure 4).
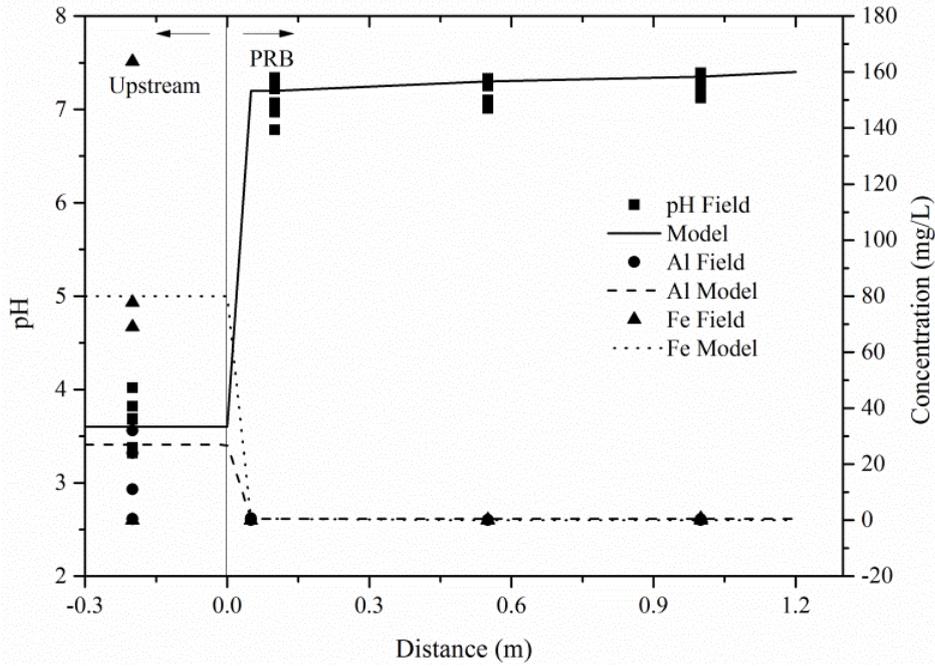
Figure 3. Field observed results and model predictions of pH, Al and Fe concentrations in the upstream and PRBn for 2012
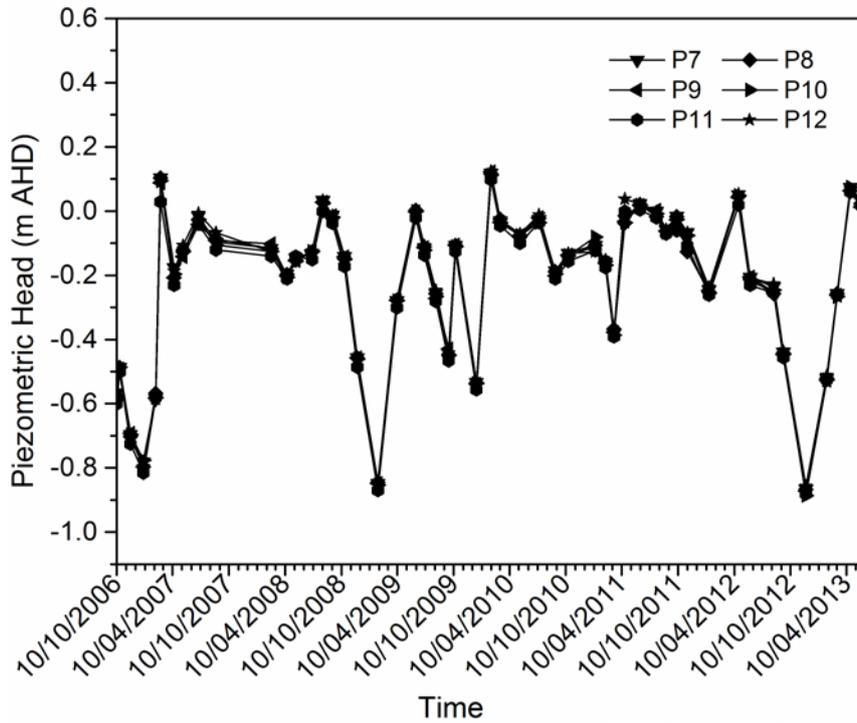
Figure 4. Groundwater elevations inside the PRB with respect to time (P7-P12 are the six piezometers inside the PRB) (updated after Regmi 2012)

Conclusion

MODFLOW and RT3D were run simultaneously to simulate flow and the reactive transport of dominant mineral components. The simulated pH and the Al^{3+} and total Fe concentrations were in good agreement with the observed field data. The developed mathematical model captures the

change in hydraulic conductivity due to mineral precipitation/dissolution. The associated hydraulic conductivity reduction after six years of operation is only 3% in the PRB. Moreover, this is confirmed by the steady piezometric heads inside the PRB. The performance of the PRB for the last six years confirms that recycled concrete is a low cost suitable reactive material for using in PRBs for the remediation of acidic groundwater in typical acid sulfate soil terrain. The average pH within the PRB was around 7. The pH of the entrance zone of the PRB has been decreasing slowly, compared to that of the middle and exit zones. This is attributed to hindrance of the alkalinity generating materials in recycled concrete as well as by secondary mineral precipitates accumulating on the reactive surface and in pore spaces of the materials.

Acknowledgment

References

APHA 1998. *Standard Methods for the Examination of Water and Wastewater,* Washington, 20th Edition, American Public Health Association.

Indraratna, B., Glamore, W., Blunden, B., Downey, J. & Tularam, A. 2001. Engineering strategies for controlling problems of acid sulphate soils in low-lying coastal areas. *In:* Smith, D. W., Fityus, S. G. & Allman, M. A., eds. *Proc. 2nd Australia and New Zealand Conference on Environmental Geotechnics*, 2001 Newcastle, Australia. Not Collected: Australian Geomechanics Society Inc., 133-146.

Indraratna, B., Golab, A. & Glamore, W. 2005. Acid sulphate soil remediation techniques on the Shoalhaven River Floodplain, Australia. *Quarterly Journal of Engineering Geology and Hydrogeology,* 38, 129-142.

Indraratna, B., Pathirage, P. U., Kerry, R. & Banasiak, L. 2014. Coupled hydro-geochemical modelling of a permeable reactive barrier for treating acidic groundwater. *Computers and Geotechnics Journal,* 55, 429-439.

Li, L., Benson, C. H. & Lawson, E. M. 2006. Modeling porosity reductions caused by mineraql fouling in continuous-wall permeable reactive barriers. Journal of Contaminants Hydrology, 83, 89-121.

Pathirage, P. U., Indraratna, B., Nghiem, L. D., Banasiak, L. & Regmi, G. 2012. Armouring by precipitates and the associated reduction in hydraulic conductivity of recycled concrete aggregates used in a novel PRB for the treatment of acidic groundwater. 11th Australia - New Zealand Conference on Geomechanics: Ground Engineering in a Changing World (pp. 828-833). ANZ Geomechanics. Melbourne, Australia.

Pathirage, P. U., Indraratna, B., McIntosh, G. & Banasiak, L. 2014. Modeling of mineral fouling in an alkaline permeable reactive barrier in Australia. 14th International Conference of the International Association for Computer Methods and Advances in Geomechanics. Kyoto, Japan. (Accepted)

Regmi, G., Indraratna, B., Nghiem, L. D. & Golab, A. N. 2011. Treatment of Acidic Groundwater in Acid Sulphate Soil Terrain Using Recyced Concrete: Column Experiments *Journal of Environmental Engineering, ASCE,* 137, 433-443.

Regmi, G. 2012. Performance Validation of a Permeable Reactive Barrier (PRB) for Treating Acidic Groundwater. Degree of Doctor of Philosophy, University of Wollongong.

White, I., Melville, M. D., Wilson, B. P. & Sammut, J. 1997. Reducing acidic discharges from coastal wetlands in eastern Australia. *Wetlands Ecology and Management,* 5, 55-72.

Thermo-Mechanical Model for Wheel Rail Contact using Coupled

Point Contact Elements

*J. Neuhaus[1] and W. Sextro[1]

[1]Chair of Mechatronics and Dynamics, University of Paderborn, Pohlweg 47-49, 33098 Paderborn, Germany

*Corresponding author: jan.neuhaus@upb.de

Abstract

A model to calculate the locally resolved tangential contact forces of the wheel rail contact with respect to contact kinematics, material and surface properties as well as temperature is introduced. The elasticity of wheel and rail is modeled as an elastic layer consisting of point contact elements connected by springs to each other and to the wheel. Each element has two degrees of freedom in tangential directions. The resulting total stiffness matrix is reduced to calculate only the position of the elements in contact. Friction forces as well as contact stiffnesses are incorporated by a nonlinear force-displacement characteristic, which originates from a detailed contact model. The contact elements are transported through the contact zone in discrete time steps. After each time step an equilibrium is calculated. For all elements, their temperature and its influence on local friction are regarded by calculating friction power and temperature each time step.

Keywords: Rolling Contact, Discrete Elements, Contact Stiffness, Temperature

Introduction

To calculate the tangential contact forces in wheel rail systems for e.g. vehicle dynamics, a profound knowledge of the creep force characteristic is necessary. To achieve this, the creep dependent distribution of tangential forces inside the contact zone is necessary. Especially at higher slip the temperature in the contact zone becomes high enough to effect significantly on the tangential forces.

Many approaches have been made to calculate the tangential force distribution. Kalker [Kalker (1967); Kalker (1990)] developed the program CONTACT to calculate contact forces. This program assumes Hertzian contact and halfspace assumption to calculate the traction forces. Due to halfspace assumption the computation times are very high. Approaches have been made to reduce computation time compared to CONTACT [Kalker (1982); Polach (2000)], however these do not include temperature effects.

In [Sextro (2007)] a numerical model for the wheel rail contact is developed. The contact zone is discretized and each partial area is described by a point contact element. A single point contact element includes differential contact stiffness as well as a nonlinear partial friction force. By applying deformations on the point contact elements, which are deduced from the kinematics of wheel and rail, the tangential force distribution in the contact zone can be calculated. Also temperature and its influence on the friction coefficient are regarded in this approach.

In [Tomberger et. al (2011)] a model for the wheel rail contact comprising roughness interfacial fluids and temperature is shown. The tangential contact is discretized using independent contact stiffnesses as suggested in [Sextro (2007)]. The modeling method is based on FASTSIM, but allows varying coefficients of friction. In a micro contact model the effects of interfacial fluids and temperature on the friction coefficient are calculated with respect to roughness, contact kinematics and material properties.

Also, the Finite Element Method has been used to calculate forces in the rolling contact. Nackenhorst [Nackenhorst (2004)] used an arbitrarian lagrangian eulerian formulation for calculating contact forces using the Finite Element Method for stationary rolling contact. By splitting the motion into rigid body motion and elastic deformation and solving the transportation problem of wheel and rail elements through the contact zone similar to transport problems in fluid dynamics. The contact forces can be calculated by this method retaining a fine mesh in the contact zone and a large mesh outside of it.

In [Wen et al. 2011] and [Zhao and Li 2011] the tangential forces in wheel rail contact where calculated using Finite Elements models. However, temperature effects were not taken into account and the computation time for these models is very high.

The model presented here is based on the approach by Sextro [Sextro (2007)]. Instead of using independent point contact elements, the elasticity of wheel and rail is modeled by an elastic layer, to better model the elastic deformation inside the contact zone. The algorithm used in this model computes the tangential contact forces at stationary rolling by computing equilibrium of the forces caused by the elastic layer and the tangential contact forces. Additionally, the influence of roughness and temperature can be investigated with this model.

Contact Kinematics

The kinematics of a rigid wheel rolling over a rigid rail under slip angle α and angular velocity ω are depicted in Fig. 1.
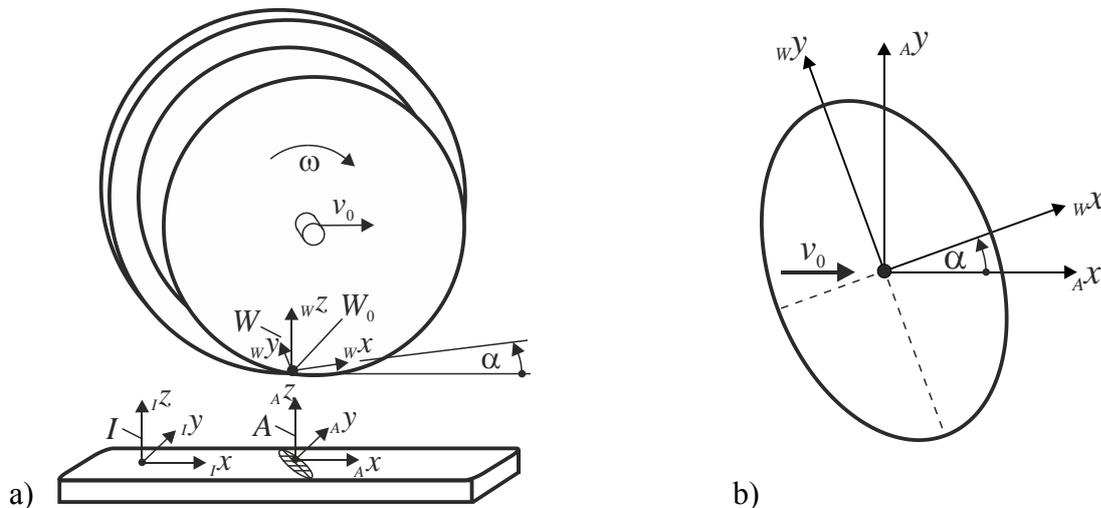
Figure 1, Wheel and Rail Schematic

The wheel shown in Fig. 1 a) is moving relative to the inertia frame I in $_I x$-direction with constant velocity v_0. The contact patch shown in Fig. 1b) is described by the reference system A, which is also moving with the wheel center in $_I x$-direction with velocity v_0. A point on the wheel's surface in contact is described by the wheel fixed coordinate system W, which is rotated relative to the inertia system around the $_I z$-axis by a constant angle α. The angular velocity of the wheel is defined relative to the wheel fixed coordinate system W with

$$_{W}^{IW}\underline{\omega} = \begin{bmatrix} 0 & \omega & 0 \end{bmatrix}^T .$$ (1)

The inertia system *I* is fixed on the rail. The tangential velocity of the point W_0 on the wheel in the contact is dependent on the angular velocity and the slip angle α and can be calculated by

$$_I^I \underline{v}_{W_0} = \begin{bmatrix} v_0 - \omega R \cos\alpha & -\omega R \sin\alpha & 0 \end{bmatrix}^T .$$
(2)

The longitudinal slippage *s* can be calculated with

$$s = \frac{v_0 - \omega R}{v_0} .$$
(3)

In the following, all vectors will be given in coordinates of the *W*-system, unless noted otherwise.

Contact Model

Modelling approach

The creep force characteristic of wheel and rail is highly dependent on the elastic deformations inside the contact area. The elasticity of wheel and rail is modeled as an elastic layer consisting of discrete elements having an elasticity equivalent to the combined elasticity of wheel and rail. A detail of the elastic layer is shown in Fig. 2 a). The layer consists of coupled massless point contacts *P* connected by springs with spring stiffnesses $\Delta c_{k,x}$ and $\Delta c_{k,y}$ in tangential direction to each other. The layer is also coupled by springs with stiffnesses Δc_x and Δc_y to the rigid wheel in ${}_W x$- and ${}_W y$-direction respectively.
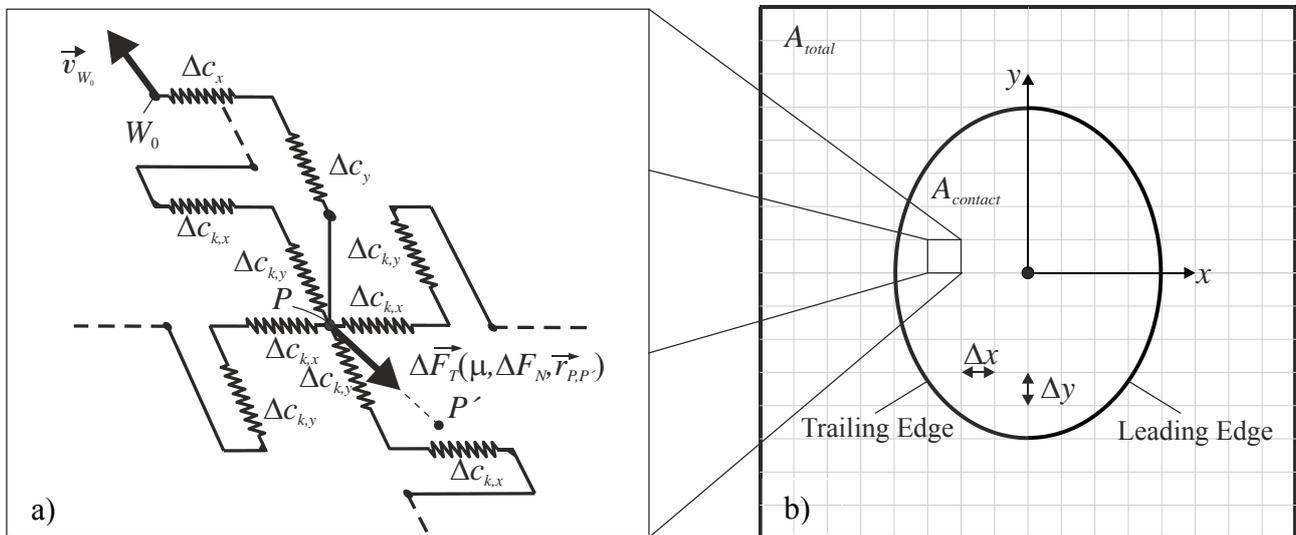
Figure 2, Elastic Layer and Contact Area

If the point *P* is inside the contact area, a nonlinear tangential force $\vec{F}_T(\mu, \Delta F_N, \vec{r}_{P,P'})$ applies, which is described in detail later in this paper. *P'* is a point which is fixed on the rail, so $\vec{r}_{P,P'}$ is the sliding distance of *P*, which is also the displacement of *P* relative to the rail. The tangential forces cause deformations which influence the neighborhood around the contact area and vice versa. Therefore, not only the contact area but also the surrounding area A_{total} is modeled with the elastic layer. The size of the total area A_{total} is not as large as the total wheel's surface but large enough, that deformations at the edge of the total area are diminishing small compared to the deformations inside the contact area. The discretised elliptical contact area with the surrounding total area is shown in

Fig. 2 b). The total area is discretised in rectangular partial areas with the dimensions Δx and Δy. The position of the contact area border in drive direction is referred in the following as leading edge. The elastic layer can be described by a set of linear coupled equations. The relation between force and displacement with respect to the wheel coordinate system W is:

$$\begin{bmatrix} \underline{F}_T \\ \underline{0} \end{bmatrix} = \begin{bmatrix} \underline{C}_{ii} & \underline{C}_{io} \\ \underline{C}_{oi} & \underline{C}_{oo} \end{bmatrix} \begin{bmatrix} \underline{r}_{P,i} \\ \underline{r}_{P,o} \end{bmatrix}. \tag{4}$$

The force and displacement vectors in this equation have been sorted for point elements inside the contact area with index i and outside the contact area with index o. The displacement vectors can be written as

$$\underline{r}_{P,i} = \begin{bmatrix} r_{P1,x,i} & r_{P1,y,i} & r_{P2,x,i} & r_{P2,y,i} & \cdots \end{bmatrix}^T \text{ and } \underline{r}_{P,o} = \begin{bmatrix} r_{P1,x,o} & r_{P1,y,o} & r_{P2,x,o} & r_{P2,y,o} & \cdots \end{bmatrix}^T. \tag{5}$$

$r_{P,x}$ and $r_{P,y}$ are the displacements of the point P in x- and y-direction respectively.

In order to reduce computational effort, the equation system is reduced to calculate only the forces and displacements inside the contact area. The displacements of the points outside the contact area can be calculated by the second row of the block matrix equation Eq. (4):

$$\underline{r}_{P,o} = -\underline{C}_{oi} \underline{C}_{oo}^{-1} \underline{r}_{P,i}. \tag{6}$$

Inserting Eq. (6) in the first row of Eq. (4) leads to

$$\underline{F}_T = \underbrace{\left(\underline{C}_{ii} - \underline{C}_{oi} \underline{C}_{oo}^{-1} \right)}_{\underline{C}_{red}} \underline{r}_{P,i}. \tag{7}$$

So, the relation between tangential forces and tangential displacements in contact can be described by a linear equation system with reduced stiffness matrix $\underline{C}_{red}$.

Nonlinear Tangential Force

The elasticity of wheel and rail is modeled as described above. Due to the roughness of surfaces in contact a normal pressure dependent contact stiffness develops. Furthermore, the tangential force for sticking and sliding conditions has to be applied at point P. This is regarded by a nonlinear force displacement characteristic, derived from a detailed micro contact model [Neuhaus and Sextro (2013)]. Using this model, the tangential force including the transition from sticking to sliding can be computed using measured rough surfaces for different nominal normal pressures. The curves achieved from this simulation can be approximated by adequate analytical functions for efficient use in the rolling contact model. Figure 3 shows a result using this micro contact model. The development of the normalized force while moving two surfaces tangentially against each other is shown. The slope at normalized tangential displacement of zero can be interpreted as contact stiffness while sticking. Full sliding is indicated by a gradient of zero of the force displacement curve.

Pure sticking only exists when no tangential force applies; afterwards the curve rises nonlinear and continuously goes over to sliding. This curve can be approximated well by an exponential function. In case of monotone relative movement in constant direction this force displacement characteristic can be used to model sticking and sliding, because the displacement vector always points in the same direction as the velocity vector.
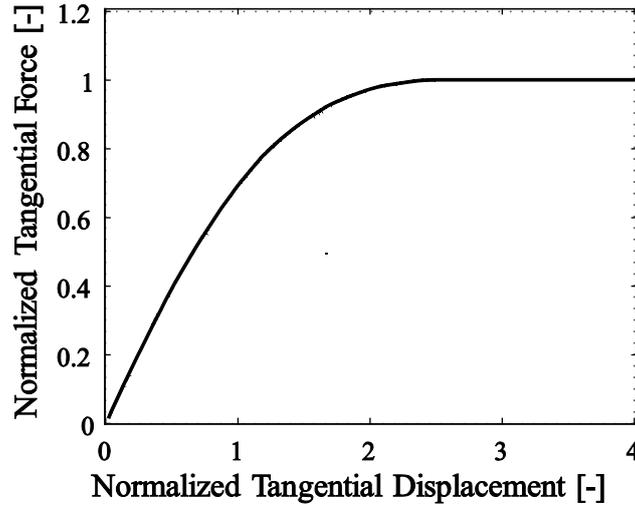
Figure 3, Tangential Force

The tangential force on a single point contact P is dependent on the sliding distance of P and thus of the magnitude of $\vec{r}_{P,P'}$ and its direction. As mentioned above, it can be calculated by an exponential law which approximates the force displacement characteristic seen in Fig. 3 with

$$\Delta \underline{F}_T = \mu \Delta F_N \left(1 - e^{-k_c \left| r_{P,P'} \right|} \right) \frac{{}^W \underline{r}_{P,P'}}{\left| \underline{r}_{P,P'} \right|}$$ (8)

where $\dfrac{{}^W \underline{r}_{P,P'}}{\left| \underline{r}_{P,P'} \right|}$ is the unit vector pointing opposite to the sliding direction.

The corresponding differential normal force is calculated according to a given normal pressure distribution. The normal pressure distribution can be achieved for example by using the Hertzian Theory, Finite element calculations or measurement results.

Simulation procedure

The rolling contact is simulated by transporting the elastic layer as described above through the contact area in discrete time steps Δt. The transporting velocity relative to the contact coordinate system A given in coordinates of the wheel system W is

$$^A\underline{v}_t = \begin{bmatrix} -\omega R & 0 & 0 \end{bmatrix}^T.$$ (9)

In order to maintain a constant grid, the value of Δt is chosen in a way that an element is transported the distance Δx in one time step. So, the position $\vec{r}_{W_0}$ of point W_0 is equal to the position of its successor in negative x -direction after one time step if

$$\Delta t = \frac{\Delta x}{\omega R}.$$ (10)

Because P' is fixed to the rail a relative differential displacement $\Delta \vec{r}_{W_0,P}$ between W_0 and P' applies after each time step Δt. It is the displacement between a point on the rigid wheel and a point on the

rigid rail after one time step and can be expressed relative to the contact coordinate system A assuming small angles α by

$$\Delta_A \underline{r}_{W_0, P'} = \begin{bmatrix} v_0 - \omega R & -\omega R \alpha & 0 \end{bmatrix}^T \Delta t \tag{11}$$

or with respect to the wheel coordinate system W by

$$\Delta \underline{r}_{W_0, P'} = \begin{bmatrix} v_0 - \omega R & -v_0 \alpha & 0 \end{bmatrix}^T \Delta t . \tag{12}$$

This differential relative displacement $\Delta \vec{r}_{W_0, P}$ is added to the position vector of P', $\vec{r}_{P'}$, after each time step.

For new elements entering the contact area at the leading edge, the position vector $\vec{r}_{P'}$ is set to their position vector $\vec{r}_P$ immediately before entering the contact, such that $\vec{r}_{P,P'} = \vec{0}$. After applying the relative differential deformation for one time step, an equilibrium between the nonlinear tangential force defined in Eq. 8 and the forces from the elastic layer is found according to Eq. 7 by

$$\underline{C}_{red\,W} \underline{r}_{P,i} - _W \underline{F}_T = \underline{0} \tag{13}$$

with

$$\underline{F}_R = \begin{bmatrix} \Delta F_{R,1}^T & \Delta F_{R,1}^T & \cdots & \Delta F_{R,nk}^T \end{bmatrix}^T \tag{14}$$

where n_k is the number of elements in contact.

The equilibrium for Eq. 13 is found by using the Newton Raphson method. Firstly, the gradient of the tangential force $\dfrac{\partial \underline{F}_T}{\partial \underline{r}_{P,i}}$ has to be calculated, which can be done due to the analytical description of the force displacement characteristic. Secondly, the Jacobian $\underline{J}$ is calculated by adding the force gradient of the nonlinear tangential force to the constant gradient emerging from the reduced stiffness matrix shown in Eq. 15:

$$\underline{J} = \underline{C}_{red} + \frac{\partial \underline{F}_T}{\partial \underline{r}_{P,i}^T} \tag{15}$$

Using the Jacobian from Eq. 15, a differential displacement can be calculated and added to the displacement vector $_W \underline{r}_{P,i}$ iteratively until the equilibrium condition of Eq. 13 is fulfilled within an relative error ε_{rel}.

$$\left| \underline{C}_{red} \underline{r}_{P,i} - \underline{F}_T \right| < \varepsilon_{rel} \mu F_N \tag{16}$$

After finding the equilibrium for a defined tolerable error ε a new time step is calculated until a steady state is reached in the contact area. This is usually the case, when an element entering the contact area at simulation start has crossed and left it at the trailing edge.

Temperature Model

Caused by high contact pressure in the wheel rail contact and consequently high friction power, the temperature in the contact zone plays an important role to describe the friction behavior correctly. Therefore, this effect is considered in this model. The model used here is adapted from the temperature model described in [Sextro (2006)]. The friction power for a single element can be calculated by

$$\Delta P_R = \Delta \left| \underline{F}_R \right| \left| \underline{v}_P \right| . \tag{17}$$

The heat source q_H can be calculated relating the friction power to a partial area

$$q_H = \frac{\Delta P_R}{\Delta A}, \quad \Delta A = \Delta x \Delta y . \tag{18}$$

Assuming a high Peclet number, Knothe and Liebelt [Knothe and Liebelt (1990)] showed, that the three dimensional heat transfer problem can be reduced to a two dimensional problem for a strip in x-direction using the approximated heat transfer equation

$$v_0 \frac{\partial T}{\partial x} = \kappa \frac{\partial^2 T}{\partial z^2} \tag{19}$$

with the thermal diffusivity defined as

$$\kappa_W = \frac{\lambda_W}{\rho_W c_W} , \tag{20}$$

where λ_W denotes the heat conductivity, ρ_W the density and c_W the specific heat capacity of the wheel [Sextro (2006)]. Assuming a constant heat source, the temperature distribution in x-direction, the temperature at time step j can then be calculated using the temperature at the element at time step j-1 by

$$T_j^2 = T_{j-1}^2 + \Delta T^2 . \tag{21}$$

The temperature difference ΔT is computed as

$$\Delta T = \frac{2}{\sqrt{\pi}} \sqrt{\frac{\kappa_W}{v_0}} \frac{\alpha_W q_H}{\lambda_W} \sqrt{\Delta x} \tag{22}$$

with α_W as the heat partitioning factor between wheel and rail. For low slippage the factor can be assumed to be 0.5, this means heat is equally distributed between wheel and rail. Otherwise it can be calculated from the velocities of the contact relative to wheel and rail. For details see [Sextro (2006)]. Using this approach, the maximum temperature occurs at the trailing edge due to the assumption of constant heat distribution, but in fact the temperature reaches its maximum shortly before the trailing edge due to the not-constant heat distribution. However this procedure can be used to regard temperature influence because the temperature distribution in the main contact area where most of the forces are transmitted is approximated well.

In principle, due to the mutual interaction between friction power, temperature and friction coefficient, an iterative loop inside the computation of a single time step is necessary. This can be skipped, because the temperature and thus the temperature dependent friction coefficient converge within the time step simulation. The temperature dependent friction coefficient at time step *j+1* is assumed to be approximate equal to the temperature dependent friction coefficient at time step *j*. Using this procedure usually a steady state is reached after an element is transported through the complete contact area.

Thus the temperature dependent friction coefficient of time step *j+1* is calculated using a linear relationship between the friction coefficient μ_0 and temperature at time step *j*

$$\mu_{j+1} \approx \mu_j = \mu_0 \left(1 - \frac{T_j}{T_E} \right) \tag{23}$$

where T_E defines the slope of the temperature dependency.

Results

The equations above have been implemented in MATLAB to model the rolling contact. The results shown in the following are normalized and therefore have qualitative character. This will be done in future by comparing the deformations of the elastic layer with the deformation computed by a Finite Element model and adapting the spring stiffnesses to minimize the deformation difference. Nonetheless the results show that the effects in rolling contact can be modeled plausibly. For the simulations, the normal pressure has been calculated according to Hertzian theory. The normalized normal pressure distribution

$$p^* = \frac{p_N}{p_{N_{max}}} \tag{24}$$

is shown in Fig. 4. The *x** and *y** coordinates have been normalized to the ellipsis half axes.
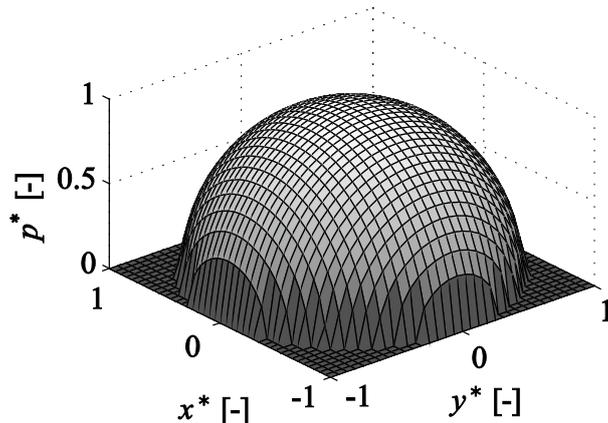
Figure 4 Normal Pressure Distribution

Fig. 5 shows the tangential force in *x*-direction normalized to the maximal transmittable tangential force with

$$\Delta F_T^* = \frac{\Delta F_T}{\mu_0 p_{N_{max}} \Delta A}. \tag{25}$$

The simulation has been performed at pure longitudinal slip s of 0.008. The leading edge is seen front right. A zone with sticking friction, behind the leading edge can be identified with a linear rising tangential force towards the trailing edge shown in Fig. 5a) where almost no sliding velocity exists, as shown in Fig. 5b). Also a region with sliding friction can be seen towards the trailing edge. Here, the shape of the tangential force distribution equals the normal pressure distribution and the sliding velocities rise towards the trailing edge. The step in the shown tangential force at the leading edge and at the transition from sticking to sliding is caused by the coupling stiffnesses. Qualitatively the shape of the tangential force distribution matches well with Kalker's theory and Finite element calculations.
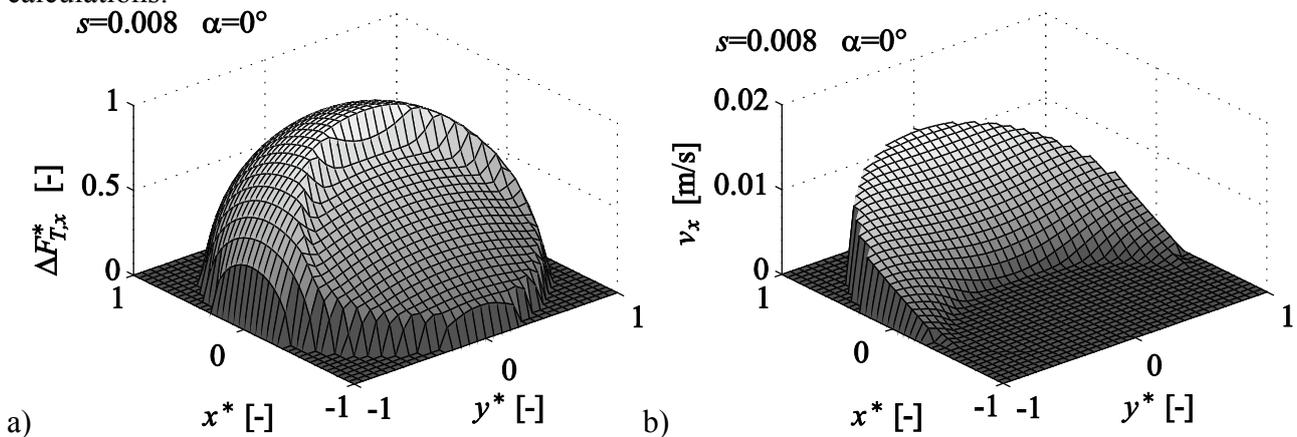
Figure 5 Tangential Force Distribution and Sliding Velocity

In Fig. 6 the tangential forces in x- and y-direction for a longitudinal slip s of 0.008 and slip angle α of 0.5 degrees are shown.
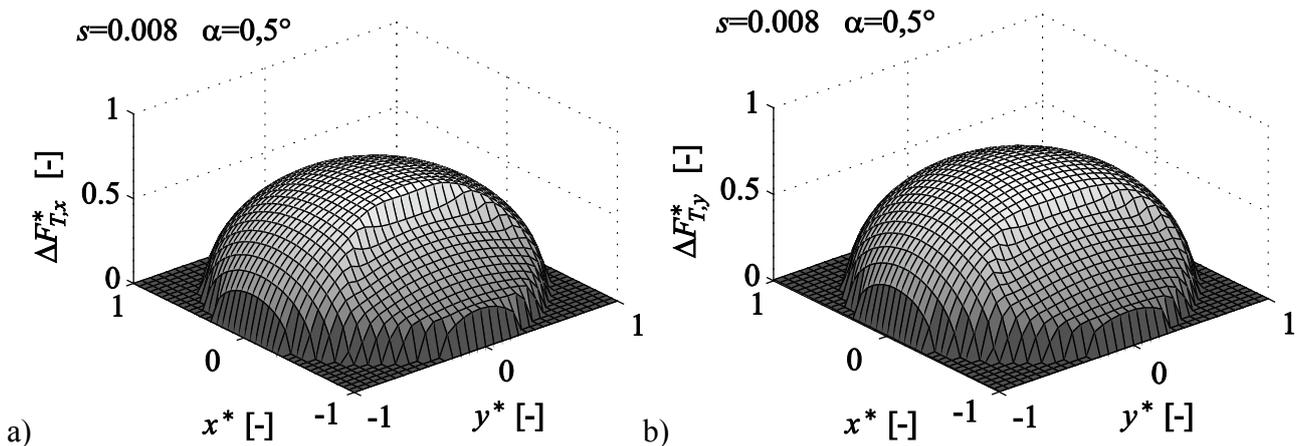
Figure 6 Tangential Force Distribution in x- and y-Direction

Compared to pure longitudinal slip, the total tangential force in x-direction is significantly lower, because lateral forces in y-direction occur at this slip angle, which is shown in Fig. 6b). Also the area with sliding friction is larger compared to the case with pure longitudinal slip in Fig. 5.

The friction power for each partial area ΔP_R and the temperature difference ΔT relative to the surrounding are shown in Fig. 7 a) and b) respectively for combined slip. The friction power reaches its maximum just before the trailing edge due to the tangential force and sliding velocity

distribution. In the sticking region, the friction power is zero. As described at the temperature model, temperature reaches its maximum at the trailing edge.
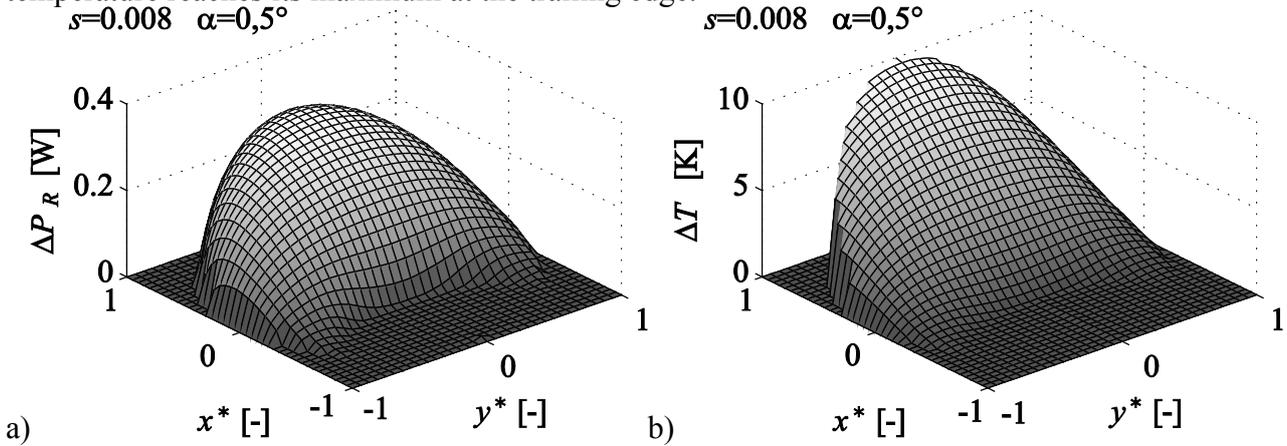
Figure 7 Friction Power and Temperature

The computing time for one simulation was 19 seconds on an Intel i7 processor using 2 of 4 cores. This is quite low compared to computationally intensive Finite Element models.

Conclusions

A model for calculating the tangential forces in the wheel rail contact comprising contact stiffness and temperature has been developed. The model is based on the rolling contact model of Sextro [Sextro (2006)]; however the point contacts are coupled with each other to model an elastic layer which represents the elasticity of wheel and rail. Also, a nonlinear tangential force is applied to the coupled massless points, to model contact stiffness as well as sticking and sliding friction. The set of equations for describing contact and surrounding area can be reduced to compute only forces and displacements inside the contact area which reduces computation time. The simulation is carried out in discrete time steps, in which the elastic layer is moved through the contact area and an equilibrium is calculated after each time step. Temperature and its influence on the friction coefficient are calculated as in the model of Sextro [Sextro (2006)] assuming a high Peclet number and constant heat source distribution. This is a satisfying approximation for the exact solution of elliptical heat source distribution.

The simulation produces plausible results for tangential force distribution under pure longitudinal and combined slip. Friction power and temperature distribution match qualitatively well compared to other modeling methods.

References

Kalker, J. J. (1967). *On the rolling contact of two elastic bodies in the presence of dry friction* (Doctoral dissertation, Delft University)

Kalker, J. J. (1990). *Three-dimensional elastic bodies in rolling contact* (Vol. 2). Springer.

Kalker, J. J. (1982). A fast algorithm for the simplified theory of rolling contact. *Vehicle System Dynamics*, *11*(1), 1-13.

Polach, O. (2000). A fast wheel-rail forces calculation computer code. *Vehicle System Dynamics*, *33*, 728-739.

Sextro, W. (2007). *Dynamical contact problems with friction: models, methods, experiments and applications*. Springer.

Tomberger, C., Dietmaier, P., Sextro, W., & Six, K. (2011). Friction in wheel–rail contact: a model comprising interfacial fluids, surface roughness and temperature. *Wear*, *271*(1), 2-12.

Nackenhorst, U. (2004). The ALE-formulation of bodies in rolling contact: theoretical foundations and finite element approach. *Computer Methods in Applied Mechanics and Engineering*, *193*(39), 4299-4322.

Wen, Z., Wu, L., Li, W., Jin, X., & Zhu, M. (2011). Three-dimensional elastic–plastic stress analysis of wheel–rail rolling contact. *Wear*, *271*(1), 426-436.

Zhao, X., & Li, Z. (2011). The solution of frictional wheel–rail rolling contact with a 3D transient finite element model: Validation and error analysis. *Wear*, *271*(1), 444-452.

Neuhaus, J.; Sextro, W. (2013): A Discrete 2D Model for Dynamical Contact of Rough Surfaces. *3rd International Conference on Computational Contact Mechanics 2013, 2013*

Knothe, K., & Liebelt, S. (1995). Determination of temperatures for sliding contact with applications for wheel-rail systems. *Wear*, *189*(1), 91-99.

A New Method for Hybrid of Probability and Interval Uncertainty Analysis

*J. L. Wu[1], Y. Q. Zhang[1], and Z. Luo[2]

[1]National Engineering Research Center for CAD, Huazhong University of Science and Technology, China
[2]School of Mechanical and Mechatronic Systems, The University of Technology, Sydney, NSW, Australia

*Corresponding author: wu_jinglai@hust.edu.cn

Abstract

This paper will propose a new uncertain analysis method for dynamics problems involving hybrid uncertain parameters. The Polynomial Chaos (PC) theory is systematically integrated with the Chebyshev inclusion function theory to deliver a new hybrid uncertain analysis approach termed as PCCI method, in which the former is applied to solve the random uncertainty and the latter is used to account for the interval uncertainty. The PCCI method is non-intrusive, which does not require the amendment of the original solver for different and complicated dynamics problems. As a result, the PCCI method can be implemented easily. The interval mean (IM) and the other is interval variance (IV) are proposed as the evaluation indexes. The proposed hybrid uncertain analysis method may produce the similar accuracy of the combination of Monte Carlo method and scanning method, and it saves the computational cost much.

Keywords: Hybrid uncertainty, Polynomial Chaos, Chebyshev inclusion function

1. Introduction

In the field of uncertain research, the probabilistic method has been widely studied with the development of a series of techniques. The probabilistic methods can be classified into two types: statistic methods and non-statistic methods. Statistic methods have been derived to include a variety of approaches. The Monte Carlo method is the widely used statistic method for uncertain analysis due to the easiness of its implementation. However, the large number of sampling points of the Monte Carlo method limits its scope of applications, especially for complicated or time-consuming problems. Thus, the Monte Carlo method is often used as the reference of other probability methods. The non-statistic methods mainly include differential analysis approaches and the spectral-based stochastic finite element methods.The spectral-based stochastic finite element method employs a series expansion, like Karhunen-Loeve expansion, to represent stochastic processes, in which the Galerkin method is used to transform the original control equation with uncertain parameters to several equations without uncertain parameters. This is the main process of Polynomial Chaos (PC) expansion, and also the basis for Stochastic Response Surface Methodology (SRSM) (Isukapalli 1999). The (Xiu and Karniadakis 2003) presented an algorithm to model the input uncertainty and its propagation in incompressible flows, where the stochastic input was represented spectrally via an orthogonal polynomial functional from the Askey scheme. Compared with the solutions obtained by the Monte Carlo simulation, the generalized PC method shows better efficiency.

In the non-probabilistic uncertain methods, the interval model is experiencing popularity, because it makes it possible to measure the uncertainties for uncertain-but-bounded parameters, without requiring complete information of the system and only with knowing lower and upper bounds of an uncertain parameter. After the appearance of Moore's work (Moore 1966), several interval methods have been proposed to solve the static problems (Moore 1966; Ishibuchi and Tanaka 1990). Based on the interval arithmetic, the interval method can directly calculate the upper and lower bounds of the response, but one of the shortcomings is the overestimation caused by the wrapping effect, which is inherent in interval computation.

In interval analysis, how to reduce the overestimation becomes one of the key issues in interval analysis (Wu, Luo et al. 2013; Wu, Zhang et al. 2013). Some special techniques should be contained in interval computation to control the overestimation, such as the interval Taylor series

method (Alefeld and Mayer 2000; Jackson and Nedialkov 2002) and Taylor model method (Berz and Makino 1999; Wu, Zhao et al. 2005), or a combination of these two methods (Lin and Stadtherr 2007). However, the Taylor series and model based interval methods generally require the explicit expression of the govern equations. Recently, Wu et al (Wu, Zhang et al. 2013) proposed a Chebyshev inclusion function for ODEs with interval parameters, without requiring the explicit expression of governing equations. For nonlinear ODEs, the Chebyshev inclusion function-based method can control overestimation better than the Taylor model method. This method has also been applied to solve the DAEs with interval parameters (Wu, Luo et al. 2013) for multi-body dynamics. The Chebyshev inclusion-based method (Wu, Luo et al. 2013; Wu, Zhang et al. 2013) has shown several merits in solving dynamic problems with uncertainty, including effective control of interval overestimation, and non-intrusive characteristic which can also be applied as a general method to solve black-box type problems.

As aforementioned, it can be seen that most works in this field were mainly focused on either the random parameters or the interval variables. The studies including both types of uncertainties are relatively small, although many engineering problems in nature involve both types of uncertainties simultaneously. The research about the mixed uncertainties is mainly focused on the reliability-based design (RBD). In (Qiu and Wang 2003; Du, Sudjianto et al. 2005), the authors attempted to deal with variables characterized by a mixture of probability distributions and interval uncertainty, where the optimization method was used to find the values of random and interval variables while the reliability index was the worst scenario. It can be found that these RBD methods are mainly focused on static problems, while vehicle analysis problems often involve dynamics, which requires the solution of the differential equations with longer computational time. In dynamics problem, the simulation period is often divided to many discrete time steps, so the computational cost will be prohibitive, if the optimization is directly incorporated in each discrete time step.

This paper will mainly focus on the dynamics problems with hybrid uncertainties of random parameters and interval parameters. Due to the complexity and high computational cost of vehicle dynamics, this study will propose a more effective and efficient PCCI method to solve the dynamic problems with hybrid uncertainties. The PC theory is applied to solve the random uncertainty and the Chebyshev inclusion function theory is used to handle the interval uncertainty.

2. Polynomial Chaos theory for random parameters

The fundamental idea of polynomial chaos is that the random process of interest can be approximated by sums of orthogonal polynomials of random independent variables (Xiu and Karniadakis 2003). For a deterministic model with random inputs, if the inputs are represented in terms of the set $\left\{\xi_i\right\}_{i=1}^{n}$, the output metrics can also be represented with the same set, as the uncertainty of the outputs is solely because of the uncertainty of the inputs (Isukapalli 1999). A random process $Y(\kappa)$, viewed as a function of the random event κ, can be expanded in terms of the orthogonal polynomial chaos as:

$$Y(\kappa) = \sum_{j=0}^{\infty} y_j \phi_j \left(\xi(\kappa) \right) \tag{1}$$

Here y_j represents the deterministic coefficients to be estimated, $\phi_j(\xi)$ are the generalized Askey-Wiener polynomial chaos of order j, according to the multi-dimensional random variable $\xi = (\xi_1, ..., \xi_n)$ (Xiu and Karniadakis 2003). For uniformly distributed random variables the basis are Legendre polynomials, for Gaussian random variables the basis are Hermite polynomials, and more basis for other random variables can be find in (Xiu and Karniadakis 2003). In this paper, only the uniformly distribution random variables are considered, and other random variables can be dealt with in the same way. In the numerical implementation, we have to employ finite terms to approximate the accuracy value. If we remain s terms, $Y(\kappa)$ can be expressed by

$$Y(\kappa) = \sum_{j=0}^{s-1} y_j \phi_j \left(\xi(\kappa) \right) \tag{2}$$

The Legendre polynomial forms a complete orthogonal basis in the L_2 space consisting of the uniformly random variables, i.e.

$$\left\langle \phi_i, \phi_j \right\rangle = \left\langle \phi_i^2 \right\rangle \delta_{ij} \tag{3}$$

where δ_{ij} is the Kronecker delta, and $\langle .,. \rangle$ denotes the ensemble average inner product.

$$\langle f(\xi), g(\xi) \rangle = \int f(\xi) g(\xi) w(\xi) d\xi \qquad (4)$$

Here $w(\xi) = (1/2)^n$ is the weighting function of Legendre polynomials. With the orthogonality, the coefficient y_i in Eq. (2) can be obtained via the following expression

$$y_i = \frac{\langle Y, \phi_i \rangle}{\langle \phi_i^2 \rangle} = \frac{1}{\langle \phi_i^2 \rangle} \int Y \phi_i(\xi) \omega(\xi) d\xi \qquad (5)$$

Once getting the coefficients, the statistics characteristics can be obtained. The mean of Y is given by the 0th order term in the stochastic expansion, and the variance of Y can be expressed by the sum of square of other terms multiplying with $\langle \phi_i^2 \rangle$

$$\mu = y_0, \quad \sigma^2 = \left\langle \left(Y - \overline{Y}\right)^2 \right\rangle = \sum_{i=1}^{s-1} y_i^2 \langle \phi_i^2 \rangle \qquad (6)$$

The coefficients of the PC expansion can also be obtained through the collocation method, by using the model outputs at some selected collocation points to regress the coefficients (Isukapalli 1999). The collocation points are selected from the roots of the polynomial, which is one order higher than the polynomial chaos expansion. Once the collocation points selected, the least square method can be used to produce the coefficients, i.e.

$$\mathbf{y} = \left(\mathbf{X}(\xi)^T \mathbf{X}(\xi)\right)^{-1} \mathbf{X}(\xi)^T \mathbf{Y}, \quad \mathbf{X}(\xi) = \begin{bmatrix} \phi_0(\xi_1) & \cdots & \phi_{s-1}(\xi_1) \\ \vdots & \ddots & \vdots \\ \phi_0(\xi_N) & \cdots & \phi_{s-1}(\xi_N) \end{bmatrix} \qquad (7)$$

where $\mathbf{Y} = \left[Y(\xi_1) ... Y(\xi_N) \right]^T$ denotes the model output vector at the collocation points, N denotes the number of collocation points, $\mathbf{y} = \left[y_0, y_1 ... y_{s-1} \right]^T$ is the coefficients vector of polynomial chaos expansions, $\mathbf{X}(\xi)$ is the transform matrix, and $\xi_1, ..., \xi_N$ denote the collocation points in a n-dimensional space.

3. Chebyshev method for interval parameters

The PC method expanses the function with random variables by the sum of some orthogonal polynomials, and the corresponding orthogonal polynomials are determined by the distribution of random variables. This section will consider the interval variables.

Define a real interval $[x]$ as a connected nonempty subset of real set R, expressed as

$$[x] = [\underline{x}, \overline{x}] = \{ x \in R : \underline{x} \leq x \leq \overline{x} \}, \qquad (8)$$

where $\underline{x}$ and $\overline{x}$ denotes the lower and upper bounds of $[x]$, respectively. Any interval $[x]=[a, b]$ can be transformed to the expression of $[\eta] = [-1, 1]$, so we only consider the interval $[\eta] = [-1, 1]$ in this paper. (Wu, Zhang et al. 2013) proposed the Chebyshev inclusion function to estimate the bounds for an interval function, and to control the overestimation in interval arithmetic. The Chebyshev inclusion function will be introduced here briefly, more information can be found in (Wu, Luo et al. 2013; Wu, Zhang et al. 2013).

Firstly, for one dimensional interval function $f([\eta])$, the pth order Chebyshev inclusion function is

$$[f]([\eta]) = \frac{1}{2} f_0 + \sum_{i=1}^{p} f_i C_i([\eta]) = \frac{1}{2} f_0 + \sum_{i=1}^{p} f_i \cos i[\theta] \qquad (9)$$

where $\theta = \arccos(\eta) \in [0, \pi]$, and $C_i(\eta) = \cos i\theta$ denotes the Chebyshev polynomial with order i. The coefficients f_i can be calculated by

$$f_i = \frac{2}{\pi} \int_{-1}^{1} \frac{f(\eta) C_i(\eta)}{\sqrt{1-\eta^2}} dx \approx \frac{2}{\pi} \frac{\pi}{p+1} \sum_{j=1}^{p+1} f(\eta_j) C_i(\eta_j) = \frac{2}{p+1} \sum_{j=1}^{p+1} f(\cos \theta_j) \cos i\theta_j \qquad (10)$$

where the interpolation points η_j are defined as the zeros of Chebyshev polynomial with order $p+1$:

$$\eta_j = \cos\theta_j, \ \ \theta_j = \frac{2j-1}{p+1}\frac{\pi}{2}, \ j=1,...,p+1 \tag{11}$$

Equation (10) is given with the Gaussian-Chebyshev interpolation integral formula.

Denote the Chebyshev polynomial $C_i(\eta)$ as $\psi_i(\eta)$, the Chebyshev inclusion function can be expressed as

$$[f]([\eta]) \approx \frac{1}{2}f_0 + \sum_{i=1}^{P} f_i C_i([\eta]) = \sum_{i=0}^{P}\gamma_i \psi_i([\eta]) \tag{12}$$

Similar to the SRSM, the interpolation points are chosen to build the Chebyshev inclusion function, and the least square method is used. The Legendre polynomials in the transform matrix are required to change to the Chebyshev polynomials as

$$\gamma = \left(X(\eta)^T X(\eta)\right)^{-1} X(\eta)^T f, \text{ where } X(\eta) = \begin{bmatrix} \psi_0(\eta_1) & \cdots & \psi_{k-1}(\eta_1) \\ \vdots & \ddots & \vdots \\ \psi_0(\eta_M) & \cdots & \psi_{k-1}(\eta_M) \end{bmatrix} \tag{13}$$

where M denotes the number of interpolation points, $f = [f(\eta_1)...f(\eta_M)]^T$ denotes the model output vector at the interpolation points, $\gamma = [\gamma_0,\gamma_1...\gamma_{k-1}]^T$ denotes the coefficients vector of Chebyshev polynomials. It is noted that $C_0([\eta])=1$, $C_i([\eta])=\cos i[\theta]=[-1,1]$, $i \geq 1$, so $\psi_0([\eta])=1$, and $\psi_i([\eta])=[-1,1]$, $i \geq 1$. Based on the interval arithmetic, we can calculate the bounds of the interval function as follows:

$$[f]([\eta])=\gamma_0+\left(\sum_{i=1}^{k-1}|\gamma_i|\right)[-1,1] \tag{14}$$

4. The hybrid uncertain analysis method

In this section, both the random and interval variables are considered in $F(\xi,[\eta])$. The function contains an n-dimensional random variable $\xi \in U(-1, 1)^n$ and an m-dimensional interval variable $[\eta]=[-1, 1]^m$. Hence, the output of the function will have the characteristics of both random and interval variables, and the PCCI method will integrated the PC method with the Chebyshev based interval method.

Consider the random variable ξ only, and use Eq. (2) to expand the function $F(\xi,[\eta])$

$$F(\xi,[\eta]) = \sum_{j=0}^{s-1} \beta_j \phi_j(\xi) \tag{15}$$

Here we use β_j denotes the PC coefficients. Since the left side of Eq. (15) contains both the interval variable $[\eta]$ and the random variable ξ, while $\phi_j(\xi)$ at the right side is the Legendre polynomials which is only the function of ξ, the coefficients β_j will be a function with respect to $[\eta]$, namely $\beta_j([\eta])$. Use the Chebyshev expansion Eq. (12) to the coefficients $\beta_j([\eta])$, obtaining its Chebyshev inclusion function

$$[\beta_j]([\eta])=\sum_{i=0}^{k-1}\beta_{j,i}\psi_i([\eta]) \tag{16}$$

Here $\beta_{j,i}$ denotes the element in the coefficient matrix β with k rows and s columns. Substitute Eq. (16) into Eq. (6), the mean and variance will be obtained as follows:

$$\mu([\eta]) = \beta_0([\eta])=\sum_{i=0}^{k-1}\beta_{0,i}\psi_i([\eta]), \ \sigma^2([\eta]) = \sum_{j=1}^{s-1}\beta_j^2([\eta])\langle\phi_j^2\rangle = \sum_{j=1}^{s-1}\left(\sum_{i=0}^{k-1}\beta_{j,i}\psi_i([\eta])\right)^2\langle\phi_j^2\rangle \tag{17}$$

Since the expression of the mean and variance contain interval variables, the two statistics will also be interval numbers: interval mean (IM) $[\mu]$ and interval variance (IV) $[\sigma^2]$, respectively. Based on the Chebyshev polynomials, the IM $[\mu]$ can be expressed as

$$[\mu]([\eta]) = [\beta_0]([\eta])=\sum_{i=0}^{k-1}\beta_{0,i}\psi_i([\eta])=\beta_{0,0}+\left(\sum_{i=1}^{k-1}|\beta_{0,i}|\right)[-1,1] \tag{18}$$

Similarly, the IV $[\sigma^2]$ may be expressed by

$$[\sigma^2]([\mathbf{\eta}]) = \sum_{j=1}^{s-1}\left(\left(\sum_{i=0}^{k-1}\beta_{j,i}\psi_i([\mathbf{\eta}])\right)^2\langle\phi_j^2\rangle\right) = \sum_{j=1}^{s-1}\left(\left(\beta_{j,0}+\left(\sum_{i=1}^{k-1}|\beta_{j,i}|\right)[-1,1]\right)^2\langle\phi_j^2\rangle\right) \tag{19}$$

Since $[\mu]$ and $[\sigma^2]$ are the functions with respect to the interval numbers [-1,1], the above equations (18) and (19) still involve the overestimation (Moore 1966) according to the interval arithmetic, particularly, when the evaluated functions are multimodal. Here, the bounds of IM and IV can be calculated respectively as

$$[\mu] = \left[\underline{\mu},\bar{\mu}\right] = \left[\min_{-1\le\eta\le1}\sum_{i=0}^{k-1}\beta_{0,i}\psi_i(\mathbf{\eta}),\ \max_{-1\le\eta\le1}\sum_{i=0}^{k-1}\beta_{0,i}\psi_i(\mathbf{\eta})\right] \tag{20}$$

$$\left[\sigma^2\right] = \left[\underline{\sigma}^2,\bar{\sigma}^2\right] = \left[\min_{-1\le\eta\le1}\sum_{j=1}^{s-1}\left(\left(\sum_{i=0}^{k-1}\beta_{j,i}\psi_i(\mathbf{\eta})\right)^2\langle\phi_j^2\rangle\right),\ \max_{-1\le\eta\le1}\sum_{j=1}^{s-1}\left(\left(\sum_{i=0}^{k-1}\beta_{j,i}\psi_i(\mathbf{\eta})\right)^2\langle\phi_j^2\rangle\right)\right] \tag{21}$$

In interval analysis, the scanning method or global optimization algorithms are often applied to the above equations, in order to solve the "min" and "max" problems to obtain the bounds. In this case, the overestimation of the interval computation can be well controlled.

Since the evaluated functions may be multimodal, the global optimization algorithms have to be used in order to find the minimum and maximum values for the lower bound and upper bound of IM and IV, respectively. Based on the explicit expressions of IM and IV, both the scanning method and the global optimization algorithm can efficiently find the bounds for IM and IV. If the dimension of the interval variables is less than 3 (m<3), the scanning method (Buras, Jamin et al. 1996)can directly produce accurate bounds. However, for the high dimensional problems, some global optimization algorithms, such as the genetic algorithm, particle swarm algorithm, and simulated annealing algorithm, may be more effective.

5. The uncertain analysis of vehicle dynamics

To demonstrate the effectiveness of the proposed PCCI method in engineering, the 4-DOF roll plan model of vehicles (Blanchard, Sandu et al. 2009) is studied in this section. The roll plan model is shown in Fig. 1.

There is an added mass M on the roll bar, which denotes the driver, the passenger, and other object in the vehicle. The d denotes the distance from the added mass position to the left end of the roll bar. The vehicle body is presented by a roll bar with mass m, length l, and inertia I. The mass of left tyre and right tyre is m_{t1} and m_{t2}, respectively, and the tyre stiffness is k_{t1} for the left side and k_{t2} for the right side. Considering the nonlinear stiffness of suspension, the linear stiffness is denoted by k_i and the nonlinear stiffness is represented by k_i', where $i=1$ for the left suspension and $i=2$ for the right suspension. The damping ratio for the left suspension and right suspension is noted as c_1 and c_2, respectively.
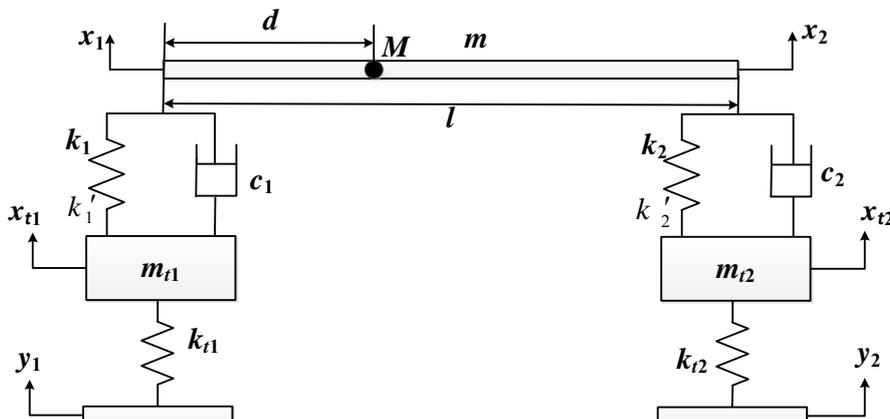
Figure. 1 The roll plan model of a vehicle

Assume that there are some uncertain parameters in this system, including the stiffness of the suspension k_1, k_2, k_1', k_2', the added mass M, and the position of added mass d. The stiffness parameters are considered as random variables, assuming that they satisfy the uniform distribution. For the added mass and its position, it is practically hard to obtain their probability distribution, but their variation ranges are limited inside some intervals. Therefore, the added mass M and its position d are described as interval parameters. The uncertain and other parameters are shown in Table 1.

Table 1 Parameters of the roll plan model

Parameters	m (kg)	m_{t1}, m_{t2} (kg)	c_1, c_2 (N/(m/s))	k_1, k_2 (N/m)	k_1', k_2' (N/m^3)
values	580	36.26	710.7	U(19000, 20000)	U(95000,105000)
Parameters	l (m)	I (kg.m^2)	k_{t1}, k_{t2} (N/m)	M (kg)	d (m)
values	1.524	63.3316	96319.76	[150, 250]	[0.5, 1]

The road input is given in Fig. 2, and the vehicle velocity is 16 km/h. The left tyre moves upgrade from 0m, and reaches the highest position 0.1m where the horizontal displacement is 1m. Keeping the height 0.1m unchanged until the left tyre goes downgrade, which is asymmetrical to the upward slope. The right tyre moves along a similar track to the left one, but its upgrade starts from 0.6m of the horizontal displacement, and its maximum height is 0.08m. The output of the roll plan model are defined as the deformation of the suspension, i.e. $z_1 = x_1 - x_{t1}$ and $z_2 = x_2 - x_{t2}$.
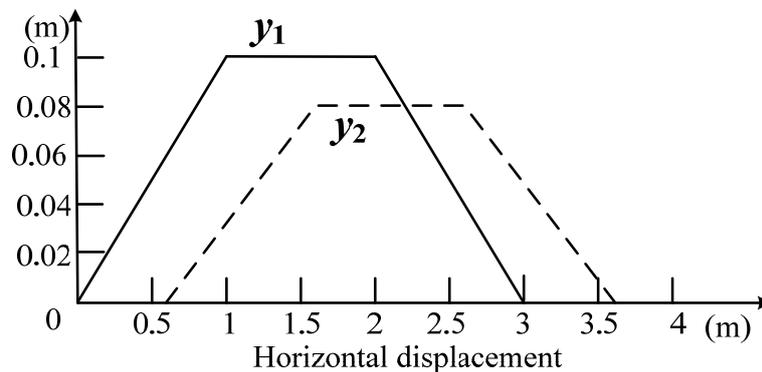
Figure 2 The road input

Due to the uncertainty of suspension stiffness, the added mass, and its position, the output should also be uncertain. The hybrid uncertain analysis method is used to solve this problem. Replacing the function $F(\xi, [\eta])$ in Fig. 1 by the output of this roll plan model, we can obtain the IM and IV of the output. In this paper, we choose the order of PC and Chebyshev inclusion function as 4 ($p=4$), and the scanning method with 20 symmetrical scanning points in each dimension of interval parameters is used to compute Eqs. (20) and (21), which provides accurate bounds information. To validate the proposed method, the Monte-Carlo-Scanning test is also performed, in which the number of Monte Carlo sampling points is 1000, and 20 scanning points are used in each dimension of interval parameters, so the total number of the system is $1000 \times 20^2 = 400,000$. The proposed method takes about 180.4s to obtain the results, while the Monte-Carlo-Scanning test takes 9915.3s, which is more than 50 times than that of the proposed method.

The IM and IV of the output are shown in Fig. 3-6. The results show that the IM of the PCCI method are close to the IM of the Monte-Carlo-Scanning test, and the test results of IM are contained in that of PCCI tightly. So the PCCI method can provide sufficient accuracy to the interval mean. For the IV, the intervals of the PCCI method do not contain all the intervals obtained by the reference test, but there is only small difference between them. Thus, the PCCI method can also provide good estimation for interval variance.
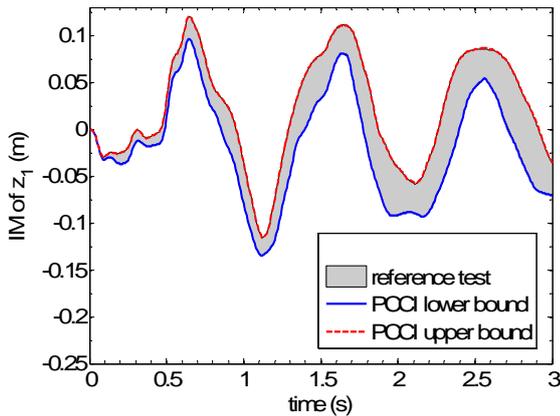
Figure 3   IM of z_1
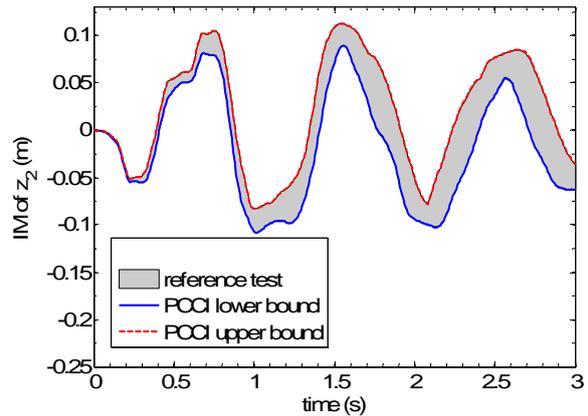
Figure 4   IM of z_2
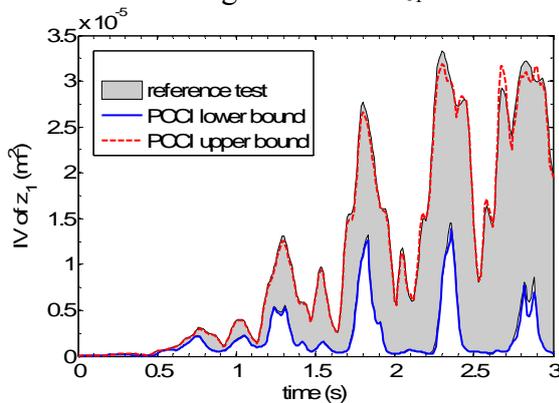
Figure 5   IV of z_1
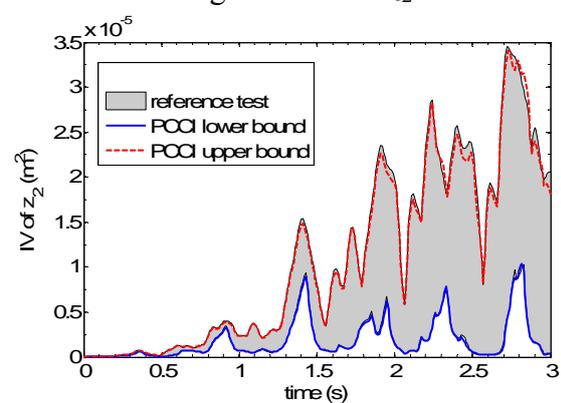
Figure 6   IV of z_2

6. Conclusions

This paper has proposed an uncertain analysis method, termed as PCCI, for systems involving hybrid uncertain parameters, namely the random parameters and interval variables. In this method, the PC method is applied to deal with the random uncertainty and the Chebyshev-based interval method is proposed to handle the interval uncertainty. The evaluation indexes are proposed, which include the interval mean (IM) and interval variance (IV). To validate the PCCI method, a Monte-Carlo-Scanning test scheme is proposed, by combining the Monte Carlo method and the scanning method to calculate the two types of evaluation indexes. A 4-DOF vehicle roll plan model is used to demonstrate the effectiveness of the proposed PCCI method, in which the stiffness of the suspension are regarded as random parameters, while the added mass and its position are considered as interval parameters. The numerical results show that the PCCI method can provide accurate numerical results for both types of the evaluation indexes. Furthermore, the PCCI method only takes 180.4s, but the Monte-Carlo-Scanning test takes 9915.3s. In addition, the PCCI method is a kind of non-intrusive method, so it can be used to solve black-box type problems.

Reference

Alefeld, G. and G. Mayer (2000). Interval analysis: theory and applications. *Journal of Computational and Applied Mathmatics* **121**: 421-464.

Berz, M. and K. Makino (1999). Efficient Control of the Dependency Problem Based on Taylor Model Methods. *Reliable Computing* **5**: 3-12.

Blanchard, E., C. Sandu, et al. (2009). *Comparison between a polynomial-chaos-based Bayesian approach and a polynomial-chaos-based EKF approach for parameter estimation with application to vehicle dynamics* the ASME 2009 International Design Engineering Technical Conferences & Computers and Inforamtion in Engineering Conference, San Diego, California.

Buras, A. J., M. Jamin, et al. (1996). A 1996 Analysis of the CP. Violating Ratio ε'/ε. *Physics Letters B* **389**(4): 749-756.

Du, X., A. Sudjianto, et al. (2005). Reliability-Based Design With the Mixture of Random and Interval Variables. *Journal of Mechanical Design* **127**(6): 1068.

Ishibuchi, H. and H. Tanaka (1990). Multiobjective programming in optimization of the interval objective function. *European Journal of Operational Research* **48**: 219-225.

Isukapalli, S. S. (1999). Uncertainty analysis of transport-transformation models. New Brunswick, New Jersey, The State University of New Jersey. **Doctor of Philosophy**.

Jackson, K. R. and N. S. Nedialkov (2002). Some recent advances in validated methods for IVPs for ODEs. *Applied Mathematics and Computation* **42**: 269-284.

Lin, Y. and M. A. Stadtherr (2007). Validated solutions of initial value problems for parametric ODEs. *Applied Numerical Mathematics* **57**(10): 1145-1162.

Moore, R. E. (1966). I*nterval analysis*. Englewood Cliffs, New Jersey, Prentice-Hall.

Qiu, Z. and X. Wang (2003). Comparison of dynamic response of structures with uncertain-but-bounded parameters using non-probabilistic interval analysis method and probabilistic approach. *International Journal of Solids and Structures* **40**(20): 5423-5439.

Wu, J., Z. Luo, et al. (2013). Interval uncertain method for multibody mechanical systems using Chebyshev inclusion functions. I*nternational Journal for Numerical Methods in Engineering* **95**(7): 608-630.

Wu, J., Y. Zhang, et al. (2013). A Chebyshev interval method for nonlinear dynamic systems under uncertainty. *Applied Mathematical Modelling* **37**(6): 4578-4591.

Wu, J., Y. Zhao, et al. (2005). An improved interval analysis method for uncertain structures. *Structural Engineering and Mechanics* **20**(6): 713-726.

Xiu, D. and G. E. Karniadakis (2003). Modeling uncertainty in flow simulations via generalized polynomial chaos. *Journal of Computational Physics* **187**(1): 137-167.

Mechanical Characterization of CFRP Contour Core Panels

*A.K. Haldar[1], W.J.Cantwell[2],and Z Guan[1], R.A Alia[1]

[1]School of Engineering, University of Liverpool, Brownlow Hill, Liverpool L69 3GH, U.K.
[2] Department of Aerospace Engineering(ARIC), Khalifa University of Science, Technology and Research (KUSTAR),PO Box 127788,AbuDhabi,UAE

* a.k.haldar@liv.ac.uk

A series of experimental investigations and numerical analyses is presented into the compression response in contour core sandwich panels based on carbon fibre reinforced plastic (CFRP). The contoured-cores were fabricated using a hot press moulding technique and then bonded to face sheets based on the same material, to produce a range of lightweight sandwich panels. The results are compared with the numerical predictions offered by a finite element analysis (FEA).The predictions of the FEA generally show reasonably good agreement with the experimental measurements.

Keywords: CFRP, Contoured Sandwich Panel, Finite element Analysis

1. Introduction

Sandwich panels are continuously being improved by developing new structural geometries with minimum weight and compact volume for automobile, aeroplane, marine and construction industries. Sandwich panels with fiber reinforced plastic skins and cellular core , have been shown to offer superior stiffness ,strength and energy absorption properties compared to their monolithic counterparts

A potential new class of energy absorbing aluminium egg box structure was introduced to understand the collapse mechanism. Experiments suggested that egg-boxes deform by either the rotation of a stationary plastic hinge or by a travelling plastic knuckle, depending upon the in-plane kinematic constraints imposed upon the egg-box (Zupan M, 2002). Egg-box shaped energy absorbing structures made of fabric composites were fabricated to find out the compressive characteristics and energy absorption capacity. The energy absorption per unit mass of composite egg-box panels made of different types of material and stacking sequences was calculated and compared with (Seong Hwan Yoo 2008). Compressive tests on foam-filled composite egg-box panels were carried out to assess their performance as energy absorbers. Material type, number of plies and stacking angle were varied. Collapse trace of the core was used to estimate energy absorption capacity. It was found that the foam-filled composite egg-box sandwich panels had a good energy absorption capacity with a stable collapse response resembling the ideal energy absorber (S.H. Yoo, 2010).

In this paper, the compressive properties of contoured core sandwich panels based on carbon fiber reinforced materials are investigated. Particular focus is placed on identifying the influence of the number of unit cells and the thickness of the cell walls in determining the compression behavior of the panel. The experimental results are compared with the numerical predictions offered by a finite element analysis (FEA).

2. Experimental procedure

2.1 Geometry and fabrication

The geometry of the CFRP contoured core panel are effectively the same as those of composite egg-box panel used in (Chunga J.G 2007).An aluminium mould was used to produce the shaped structure used in ref (Chunga J.G 2007). The mould was manufactured by using a numerically-controlled milling machine. Prepreg carbon/epoxy sheets were used to fabricate the composite cores.

2.2. Sandwich panels

A hot press was used to produce all of the contoured sheets for the sandwich panels. In order to manufacture the composite contour cores, the composite prepreg was placed between the upper and lower moulds, and then cured according to the manufacturer's recommended processing cycle. Once the hot press had cooled to a temperature below 60 C, the sheet was removed from the mould and visually inspected for defects. The contoured sheet was then bonded to the upper and lower skins using a two part epoxy adhesive (Araldite 420 A/B). The contoured sandwich panel was then heated in an oven to a temperature of 120°C for approximately 1 h in order to cure the adhesive. Two sizes of rectangular test specimen were prepared. Initially, the effect of increasing the number of cells in the core was investigated for a constant cell wall thickness. Here, (1x1) and (2x2) unit cells were investigated in order to understand the effect of varying unit cell in contoured sandwich panels. Following this, the effect of varying the cell wall thickness 0.50mm, 1.00mm and 1.50mm were investigated for a constant size of (2x2) unit cells specimen.

3. Finite element analysis

The responses of the contoured-core sandwich panels under compression loading were modelled using the ABAQUS/Standard finite element software package. A quarter model was proposed to save the model running time in abaqus, whilst applying the appropriate boundary conditions. In Fig. 1, three-dimensional shell elements (S4R) were used to model the contoured core part. The core and the platten were connected using a contact interaction formulation. The nodes along the upper and lower edges were fully constrained, except in y-direction at the upper edge. Displacements were applied uniformly to the nodes at the apex of the unit cell to simulate compression of the core. Hashin's damage model assuming, an initial linear elastic behavior followed by evolution of damage, was used to predict the behavior of the composite. The models were used to simulate initial failure and to predict the compression strength of the panels. Table 1 Properties of the materials used to produce the contoured-core sandwich panels for Finite element analysis.
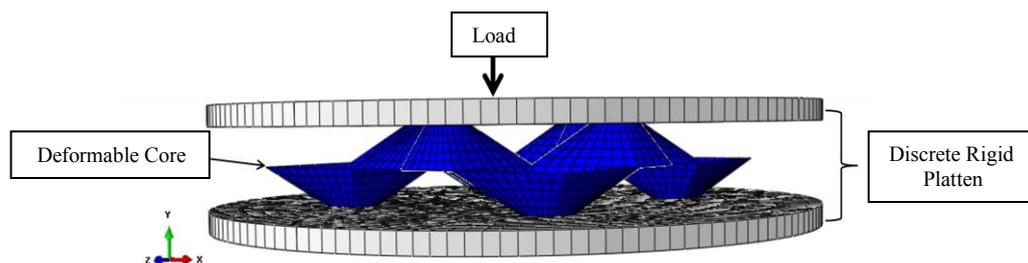
Figure 1. Loading conditions used in the contoured-core model

Table 1. Properties of the materials used to produce the corrugated-core sandwich panels for CFRP.

Symbol	Values (CFRP)	Property
E_{11}	48 GPa	Young's modulus in longitudinal direction
E_{22}	48 GPa	Young's modulus in transverse direction
E_{33}	1 GPa	Young's modulus in thickness direction
G_{12}	9 GPa	In-plane shear modulus
G_{13}, G_{23}	9 GPa	Through-thickness shear modulus
v_{12}	0.1	In-plane Poisson's ratio
v_{13}, v_{23}	0.1	Through-thickness Poisson's ratio
X_t	550	Longitudinal tensile strength
X_c	150	Longitudinal compressive strength
Y_t	550	Transverse tensile strength
Y_c	150	Transverse compressive strength
S_t	120	Transverse shear strength
S_L	120	Longitudinal shear strength

4. Results and discussion

4.1. The effect of varying the number of unit cells

The effect of varying the number of unit cells on the compression strength of the CFRP sandwich panels is shown in Fig.2.A comparison of the finite element predictions and the experimental measurements indicates that the influence of the cell number is accurately predicted.
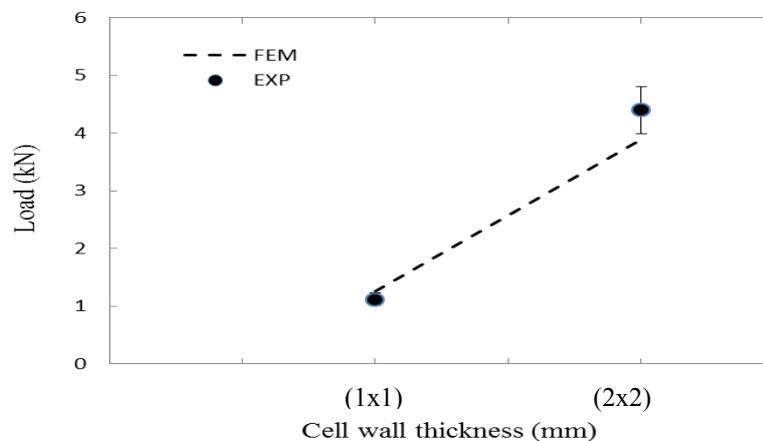
Figure 2. Comparison between measured and FE predictions for compression strength for CFRP as a function of number of unit cells

4.2. The effect of cell wall thickness

The final part of this study investigated the influence of varying the thickness of the cell wall on the compressive properties of the contoured panels. Fig.3 shows the variation of strength of the CFRP contoured panels as a function of wall thickness. As expected, the strength increases highly with wall thickness. The compression strength of the CFRP contoured panel with a wall thickness of 1.50 mm is impressive, with a value in excess of 15.98 MPa being recorded. A comparison of the finite element predictions and the experimental measurements indicates the good agreement.
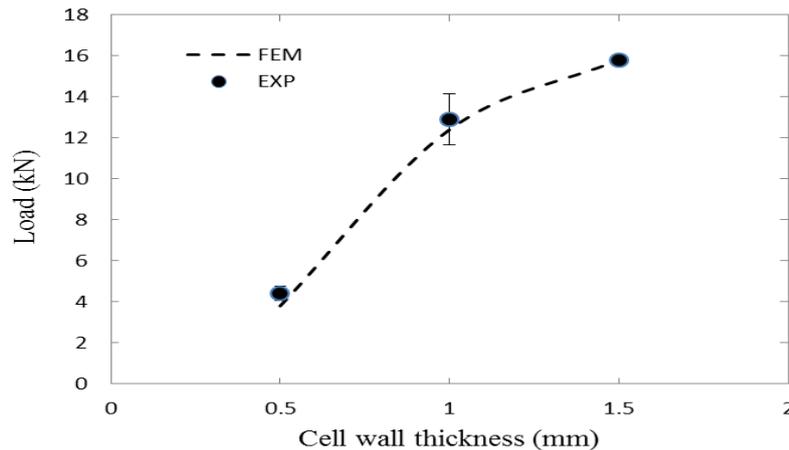
Figure 3. Comparison between measured and FE predictions for compression strength for CFRP as a function of cell wall thickness

5. Conclusion

Contoured-cores, manufactured using an aluminium profiled mould, have been used to produce a range of lightweight sandwich structures. The compressive behavior of carbon fiber reinforced composite material has been investigated both experimentally and numerically. The predictions offered by the numerical models were found to be in reasonably good agreement with the experimental data.

Reference

Chunga J.G., Changa S.H., Sutcliffe M.P.F.. (2007) Deformation and energy absorption of composite egg-box panels, *Composites Science and Technology 67,* 2342–2349.

Seong Hwan Yoo, Seung Hwan Chang.(2008) An experimental study on energy absorbing structures made of fabric composites, *composite structure86,211-219.*

Yoo S.H., S.H. Chang, M.P.F. Sutcliffe .(2010) Compressive characteristics of foam-filled composite egg-box sandwich panels as energy absorbing structures, *composite part-A 41,*427-434

Zupan M, Fleck NA, Ashby MF. (2002) The collapse and energy absorption of egg-box panels. In: Ghosh A, Snders T,Claar D, editors. Processing and properties of lightweight cellular metals and structures. Pennsylvania: TMS.

A 3D Knowledge-Based On complicated Mould Design system

*A.K. Haldar[1], J. Mathew[1], W.J.Cantwell[2], and Z Guan[1]

[1]School of Engineering, University of Liverpool, Brownlow Hill, Liverpool L69 3GH, U.K.
[2] Department of Aerospace Engineering(ARIC), Khalifa University of Science, Technology and Research (KUSTAR),PO Box 127788,AbuDhabi,UAE

* a.k.haldar@liv.ac.uk

This paper presents the basic structure of an interactive knowledge to design and manufacture the complicated mould profile with the help of CAD/CAM system. The basis of this system arises from an analysis of the mould design process for to design the complicated sandwich structures for aerospace industries. This system covers both the mould design process and mould knowledge management. CAD mould design integrates the intelligent design process and knowledge management with many developed interactive tools in a commercial solid modelling software environment. Manufacturing process of this mould including G-code generation with CAM software and Computer Numerical Control machining.

Keywords: CAD/CAM, G-code, Mould Design

1. Introduction

The core parts in Sandwich panels are continuously being improved by developing new structural geometries with minimum weight and compact volume for automobile, aeroplane, marine and construction industries to offer superior mechanical properties. The general method of 3D finishing of a free-form surface is to use a ball endmill to trace along the part surface by maintaining an acceptable tolerance (Chang et al. 1991). Gouging is the main problem in 3D finishing. When a ball endmill cutter is used, the cutter radius must be smaller than the smallest radius of concave curvature to avoid gouging (Lee and Chang 1991). Choi and Jun (1989) introduced an algorithm which avoids gouging by comparing each cutter contact (CC) point with adjacent CC points which locate within the projection of the ball endmill on the XYplane. Another approach is to use a polygon surface to verify gouging (Kuragano et al. 1988). When there is a self-intersection in the polygon surface, the portion bounded by the self-intersection lines is trimmed. The existing methods rely on discrete point data approximation, which does not guarantee the avoidance of gouging. Consequently, a robust procedure to extract machining constraints directly from a free-form surface description is desirable.

This paper introduces fundamental CAD/CAM concepts of a sophisticated geometric structure which can be used for designing sandwich panels for various engineering applications. In ProEngineer (3DCAD) a new part model accurately captures a design from a concept through solid feature-based modelling and enables us to graphically view the product before it is manufactured. This system integrates the initial mould design with both knowledge base and interactive commercial CAD/CAM.

2. Surface design and mould machining

The contoured mould was designed in commercial CAD/CAM software and manufactured to a high precision using a computer-controlled numerical milling machine (CNC) which is composed of following steps:

1. A part model design by scoping the design parameters of the structure.
2. Create a part model by following the required design parameters.
3. Transfer the part model to triangulate surfaces in ProToolmaker(CAM) to create the CNC programme in G codes.
4. Set the raw materials in Haas CNC milling machine and manufacture the desired mould through roughing and finishing operations.

2.1 A part model design by scoping the design parameters of the structure:

The proposed mould design is the combination of two different sectional profiles which is presented in fig 1.And the detail design of the cell is mentioned in fig 2.
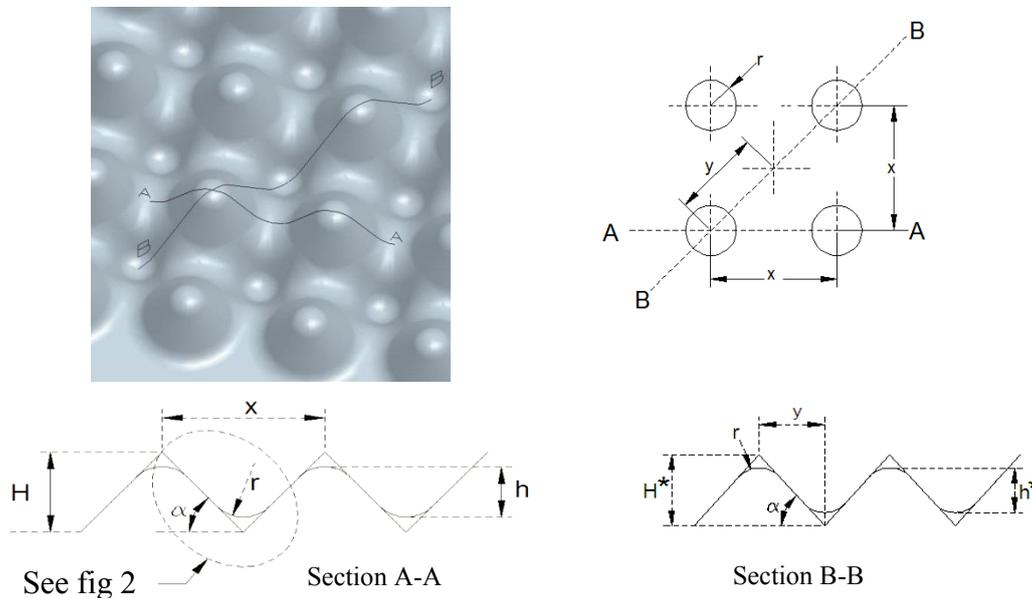
Figure1: Plan view and section Geometry of contoured profile

section A-A,
From Fig 1, r refers to the radius of the curve, which is 4 mm for the proposed design, α is the angle, H is the theoretical height and h is the actual height of the profile.

As shown in fig.2, the radius of the curvature can be calculated as:

$$r = q + s \tag{1}$$

Where
$$q = r\cos\alpha$$

By substitution the value of q in equation (1)
$$s = r(1 - \cos\alpha) \tag{2}$$

The actual height h can be found as:
$$h = H - 2t \tag{3}$$

Where

$$t = p \tan \alpha - s$$

$$p = r \sin \alpha$$

By substituting the value of t in equation (3), the relation between theoretical and actual heights can be described as:

$$h = H + 2r\left(1 - \frac{1}{\cos \alpha}\right) \tag{4}$$

Finally, the actual height can be defined as:

$$h = x\frac{\tan \alpha}{2} + 2r\left(1 - \frac{1}{\cos \alpha}\right) \tag{5}$$
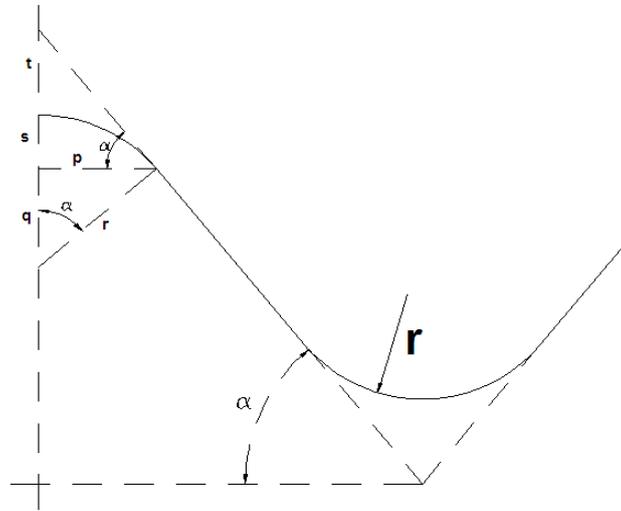
Figure 2: detail design of the cell

For section B-B

The radius and slop of the cell is same as in section A-A. But theoretical and actual heights are H^* and h^* respectively.

An important measure of the geometry can be calculated as:

$$y = \frac{x}{\sqrt{2}} \tag{6}$$

The theoretical height H^* can be defined as:

$$H^* = y \tan \alpha \tag{7}$$

The actual height h^* in section B-B, can be found as:

$$h^* = H^* - 2t \tag{8}$$

By substituting the value of t and H^* in equation (7), the actual height can be defined as:

$$h^* = x\frac{\tan \alpha}{\sqrt{2}} + 2r\left(1 - \frac{1}{\cos \alpha}\right) \tag{9}$$

322

2.2 Create a part model by following the required design parameters:

1. To avoid irregularity on a smooth surface, the 3D CAD drafting process was carefully designed. The basic geometry is simple even though due to the complexity to pattern, a virtual design of the contoured shapes was started from a plain flat 3d shape in ProEngineer software as shown in step1	
2. A basic convex geometry of radius 4mm was drawn with an angle of 50° and 20 mm width, as per the design data in A-A plane.	
3. The geometry of step 2 was revolved about the vertical axis drawn in the sketching plane to get the convex shape.	
4. In this step, the revolved shape was patterned/duplicated. There were 10 members duplicated in both, X and Y direction each at a distance of x=20 mm.	
5.In this step, the concave geometry was drawn as per the design data in B-B plane. For this drawing, the edges of the existing entity were used by selecting it with Use Edge tool.	
6. The geometry of step 5 was Revolved about the vertical axis drawn in the same plane and material removed from the model to create the concave shape.	
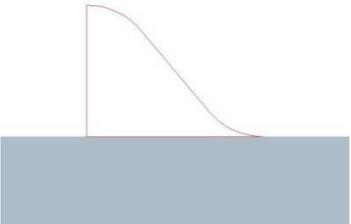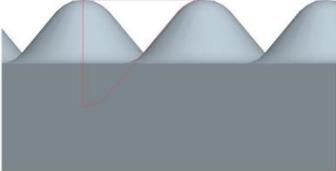
7. In this step, the revolved shape was patterned/duplicated. There were 9 members duplicated in both, X and Y direction each at a distance of x=20 mm.	
8. Round tool was used to remove the material by creating smooth transitions between existing geometry. The rounds were created by selecting edges are constructed tangent to the surfaces adjacent to the selected edges.	
9. The Shell feature was selected to hollows out the inside of a solid model, by leaving a shell of a specified wall thickness. The entire plane surface was Selected to remove from the model to create final contoured ply of 0.2mm.	
10. Finally, the rectangular sketch was extruded to the contoured ply, drawn in step 9 to create the top and bottom parts of the mould as mentioned in the step10.	
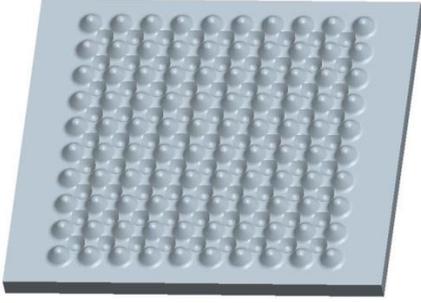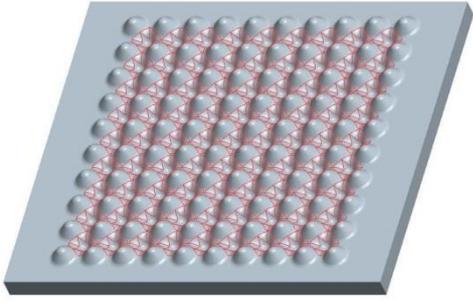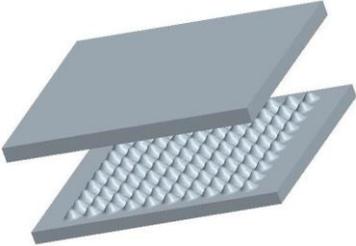
Table1: Steps to create a part model by using ProEngineer software

2.3 Transfer the part model to triangulate surfaces in ProToolmaker(CAM) to create the CNC programme in G codes:

In this section, the 3D part model was converted to IGES format and opened in ProToolmaker in order to create the manufacturing process. ProToolmaker is virtual manufacturing software which could simulate and modify the manufacturing process in a digital environment. When the file had loaded in ProToolmaker, a graphics window was opened and started to triangulate the surfaces for viewing. Triangulation converts the geometric surfaces into triangles. These triangles were used for both, to display on the screen and the machining process. Initially a roughing program was created to remove the bulk waste material from the workpiece with a toroidal carbide cutter, had 6mm diameter and 1mm corner radius was used in CNC machining. Cutting speed and feed used for this operation were 130m/min and .2mm/rev respectively. A Depth cut of 1mm was used in each passes in z direction. In this stage final finishing was left in as shown in fig 3.
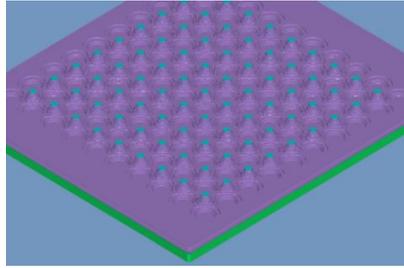
Figure 3: Material left for finishing is 0.5mm in x, y and z direction.
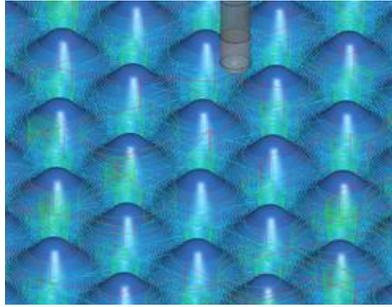
Figure 4: Milling cutter on smooth concave profile.

To get the final smooth profile as shown in fig 4, a carbide ball end milling cutter with 2 cutting edge was used with the cutting speed of 150m/min, feed per tooth 0.05mm and Depth cut of 0.1mm.The small concave radius of 4 mm was produced on final profile by using ball end mill cutter of 3mm radius.

2.4 Set the raw materials in Haas CNC milling machine and manufacture the desired mould through roughing and finishing operations:

Three axis Haas CNC milling machine was used to manufacture the mould. A milling machine is a machine tool that removes metal as the work is fed against a rotating multipoint cutter. In NC system, operating instructions are given to the machine as G-codes. The work piece was clamped to the machine bed which can move horizontally in X and Y axis. And the milling tool can move in Z axis which is perpendicular to the horizontal plane.

The toroid end mill has two cutting edges with a radius of 1mm, each cutting edge almost overlapping at the centre line of the cutter. This design eliminates the unfavorable cutting action that takes place at the tool centre of a conventional full radius end mill, where the cutting speed goes down to zero. Cutting speed and feed used for this operation were 130m/min and .2mm/rev respectively. A depth cut of 1 mm was used for the roughing operation.

A ball nose endmill has a semisphere at the tool end are ideal for machining three dimensional contoured shapes or work pieces with complex surfaces. The stepover value (along with the tool size) will determine whether the model has a smooth finish, or tooling marks are visible. Models with a smaller stepover take longer to cut. Stepover is the distance the tool moves over between subsequent passes. A ball nose end mill cutter with 6mm diameter and 0.1mm stepover were used to finish the mould. The used Cutting speed and feed for this operation were 150m/min and .05mm/rev respectively.

Figure 5: Milling cutter and work piece in Haas CNC milling machine

3. Conclusions

This paper present the method of complex shaped mould design by using computer aided manufacturing technique. The proposed technique explained the automation of complex surface machining for better machining quality. The design process of the framework as a sequence of different steps develops the design pattern to shape the final contoured mould.

References

CHANG, T. C., WYSK, R. A., and WANG, H. P. (1991) *Computer-Aided Manufacturing*, (Englewood CliOEs, NJ: Prentice Hall).

LEE, Y. S., and CHANG, T. C. (1991) CASCAMĐ an automated system for sculptured surface cavity machining. *Computers in Industry*, 16(4), 321-342.

CHOI, B. K., and JUN, C. S. (1989) Ball-end cutter interference avoidance in NC machining of sculptured surfaces. *Computer-Aided Design*, 21(6), 371-378.

KURAGANO, T., SASAKI, N., and KIKUCHI, A. (1988) The FRESDAM system for designing and manufacturing freeform objects. *USA-Japan Cross Bridge. Flexible Automation* edited by R. Martin 2, 931-938.

Thermomechanical analysis of porous SMP plate

R. Tao[1], *Q. S. Yang[1]

Beijing University of Technology, China
Department of Engineering Mechanics, Beijing University of Technology
Beijing 100124, PR China

*Corresponding author: qsyang@bjut.edu.cn

Abstract

Shape memory polymers (SMPs) are a kind of intelligent material with response to the external temperature stimulus. It can recover from a deformed state (temporary shape) to its initial state (permanent shape). In this work, the thermo-mechanical behavior and shape memory effect of a SMP-based porous plate is modeled by the finite element method using a 3D constitutive equation. The whole process of shape memory of porous SMP plate includes loading at high temperature, decreasing temperature with constant load, unloading at low temperature and recovering the initial shape by increasing temperature. The results demonstrate the thermo-mechanical deformation of SMP structure.

Keywords: Smart material and structure, Shape memory polymers, Constitutive model, Porous plate, Shape memory effect

1 Introduction

As an intelligent macromolecule material, applications of shape memory polymers (SMPs) have evoked great interest since the 1980s. Because of their light weight, good durability, large deformation and shape recovery, SMPs provide several advantages over shape memory alloys and ceramics, such as low density, high shape-fixed strain, easy operation, tailorable critical transition temperature and part of them bio-compatible. Therefore, it has great potential application in the textile, biomedical materials, defence and Military and so on [Behl, Zotzmann and Lendlein (2010); Leng, Lan, Liu and Du (2011); Hu, Meng, Li and Ibekwe (2012); Mather, Luo and Rousseau (2009)]. Since the first shape-memory polynorbornene successful development in the world, in the past two decades various SMP materials have been developed out [Lendlein A and Kelch S (2002); Takahashi et al. (1996); Yang et al. (2003)].

It has the ability to retain the deformed shape when subjected to external heat, moisture, electrical and magnetic stimulation factors can be restored to the original shape, which has a memory function to the initial shape. Thus, the storage and release of strain [Meng and Hu (2009); Nelson (2008); Lendlein and Kelch (2002)] can be achieved. Thermally isotropic SMP is the most basic and common of such materials. Achievement of storage and release of strain is due to the occurrence of glass transition of SMPs, as the temperature changes [Takahashi, Hayashi and Hayashi (1998)]. The thermomechanical cycle process in the SMPs involves the following four steps: loading at high temperature, cooling under constant load, unloading at low temperature and heating under free load. The elastic, viscoelastic and thermal deformation, and shape memory effect are displayed in this process. The above mentioned high and low temperatures denote temperatures above the end and below the beginning of glass transition of SMPs. To obtain relatively high mechanical properties, several researches have worked on SMP composites reinforced by variety of fibers and particles [Yang, Huang, Le, Leng and Mai (2012)]. These composite materials with shape memory that use of continuous fibers as reinforcement and use of thermosetting SMP as matrix have a very large potential application. In recent years, they have been widely appreciated in the space expanded structure [Keller and Lake (2003); Gall and Lake (2003)]. In the study of the constitutive models of thermotropic SMPs, which the representative study are as follows. e.g. [Tobushi, Hashimoto, Hayashi and Yamada (1997)] developed a thermomechanical constitutive model by modifying a standard linear viscoelastic model [Tobushi et al. (1997)]. The model involved a slip element due to internal friction and took account of thermal expansion. Therefore, a SMP linear constitutive model of thermodynamic properties is established. Considered the large deformation characteristics of

327

SMP, Tobushi et al. further proposed a nonlinear constitutive model based on the above work [Tobushi, Okumura, Hayashi and Ito (2001)]. These models are macroscopic and experienced but thermodynamic internal mechanism of SMP materials has not been considered. Liu et al. proposed a three-dimensional, small strain, linear elastic, and rate-independent SMP thermodynamics constitutive equations [Liu et al. (1997)]. The model can roughly predict the change trend and recovery of shape memory of SMP materials on different constraints. But this model does not consider the impact of the SMP viscoelastic. Zhou, Liu and Leng formulated a 3D thermomechanical constitutive equation of SMPs [Zhou, Liu and Leng (2009)]. However, applications of these constitutive equations are limited due to the lack of experimental data for material parameters. Another type of constitutive model of SMPs is the micromechanics-based method, such as the phase transition and the mixture theory. Liu, Gall, Dunn, Greenberg and Diani (2006) proposed a 3D linear elastic constitutive model for small deformation that considers the molecular mechanism of the shape memory. Chen and Lagoudas established a thermomechanical constitutive model for large 3D deformation of SMPs. A review of advances of constitutive relations of SMPs was given by Zhang and Yang very recently.

However, these constitutive equations of SMPs are so complicated and contain so many material parameters that it is difficult to apply them in practical engineering. The lack of experimental data for material parameters restricts the finite element implementation of the constitutive models. It is necessary to develop applicable constitutive equations with physical definitions and the finite element procedure for complicated deformation of SMPs. Actually, the available 3D finite element program and numerical investigations are very limited in existing literature.

In this paper, by considering the elastic, viscoelastic and thermal deformation of isotropic SMPs, we propose a three-dimensional form of a thermomechanical constitutive equation for isotropic thermal actuated SMPs, with defined physical significance. A finite element procedure based on the present constitutive model is implemented by using user material subroutine (UMAT) of ABAQUS, and some numerical examples are provided to illustrate the 3D deformation and shape memory effect of SMPs.

2 Thermomechanical constitutive model of SMPs

SMP thermodynamic constitutive equation plays an important role to the commercial aerospace industry adopting SMPs into their structures. In recent years, the research of the constitutive equation has been a hotspot and has made a lot of achievements. However, due to the complexity of the mechanism of SMP itself, research is continuing.

Tobushi et al. proposed a one-dimensional linear constitutive model for SMPs of polyurethane series [Tobushi et al. (1996)]. Shi et al. gave a three-dimensional constitutive equation of SMP in a rate form [Shi et al. (2013)].

$$\dot{\varepsilon}_{ij} + \frac{5\nu - 1}{3(1 - 2\nu)}\delta_{ij}\dot{\varepsilon}_{kk} \qquad (1)$$

$$= \begin{cases} \dfrac{(1+\nu)\dot{\sigma}_{ij}}{E} + \dfrac{\sigma_{ij}}{\mu} - \dfrac{\varepsilon_{ij}}{\lambda} - \dfrac{1}{3}\left(\dfrac{E}{\mu(1-2\nu)} - \dfrac{1}{\lambda}\right)\delta_{ij}\varepsilon_{kk} + \alpha\dot{T}\delta_{ij}, \\[4pt] \qquad\qquad\qquad \text{as } \bar{\varepsilon}^c(t) < \varepsilon^l(T) \\[8pt] \dfrac{(1+\nu)\dot{\sigma}_{ij}}{E} + \dfrac{\sigma_{ij}}{\bar{\mu}} - \dfrac{\varepsilon_{ij} - \bar{\varepsilon}^s(T,C)\delta_{ij}}{\bar{\lambda}} - \dfrac{1}{3}\left(\dfrac{E}{\bar{\mu}(1-2\nu)} - \dfrac{1}{\bar{\lambda}}\right)\delta_{ij}\varepsilon_{kk} + \alpha\dot{T}\delta_{ij}, \\[4pt] \qquad\qquad\qquad \text{as } \bar{\varepsilon}^c(t) \geq \varepsilon^l(T), \dot{\bar{\varepsilon}}^c(t) > 0 \\[8pt] \dfrac{(1+\nu)\dot{\sigma}_{ij}}{E} + \dfrac{\sigma_{ij}}{\mu} - \dfrac{\varepsilon_{ij} - \varepsilon^s(t_1,T)\delta_{ij}}{\lambda} - \dfrac{1}{3}\left(\dfrac{E}{\mu(1-2\nu)} - \dfrac{1}{\lambda}\right)\delta_{ij}\varepsilon_{kk} + \alpha\dot{T}\delta_{ij}, \\[4pt] \qquad\qquad\qquad \text{as } \dot{\bar{\varepsilon}}^c(t) \leq 0 \end{cases}$$

where E is Young's modulus and α is coefficient of thermal expansion. μ and λ are viscosity coefficient and retardation time, respectively, depending on the temperature. $\varepsilon^s(t,T)$ is creep residual strain, unrecovered part of the creep strain, while $\varepsilon - \varepsilon^s$ is retardation strain. It is noted that as the temperature is above the glass transition region, i.e. $T > T_h = T_g + T_w$, where T_g is the glass transition temperature and T_w is temperature amplitude of the glass transition region, the creep strain can be recovered completely, which means the $\varepsilon^s(t,T)$ does not appear. Within the glass transition temperature region, i. e. $T_h = T_g + T_w > T > T_l = T_g - T_w$, there is a critical value of creep strain at which part of the creep strain becomes irrecoverable while for the case below glass transition temperature, the creep residual strain is a constant. The constitutive equations, with the temperature-dependent parameters, can reflect the thermo-mechanical behavior of different types of SMP materials.

Within the range of glass transition of SMP, material parameters are strongly temperature-dependent, which can be expressed

$$X = X_g \exp\left[K\left(\frac{T_g}{T} - 1\right)\right] \qquad (2)$$

where T_g is the glass transition temperature, X denotes one of the material parameters E、μ、λ、C and ε_l. K is a coefficient corresponding to parameter X.

3 Introduction of model
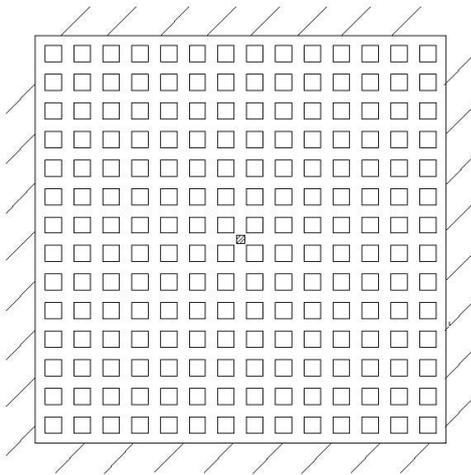
Figure 1. The simplified model of porous SMP plate
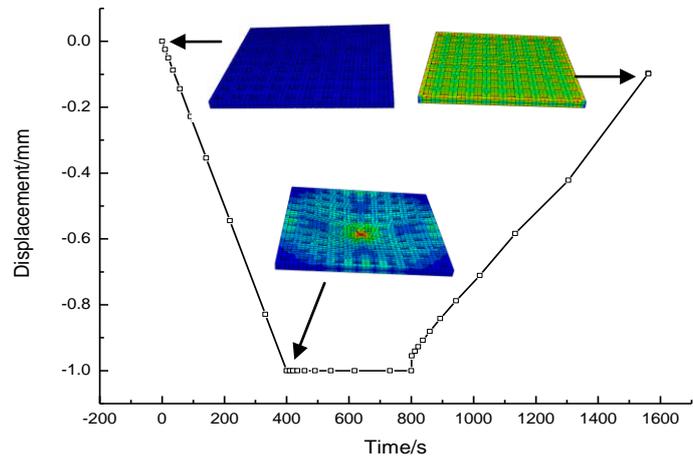
Figure 2. Displacement–time diagram showing the thermomechanical behavior of SMP

In order to realize analysis of thermo-mechanical behavior of the porous SMP plate, a simplified model is established. As shown in Fig.1, the model size is 100*100*5mm, the through holes of 5 * 5mm square uniformly distributed on the planar plate. Boundary conditions are set as follows: four sides in the edge are fixed and a specific displacement is imposed in the central region of the surface of the plate.

4 Numerical results and discussion

4.1 Experiment model and its shape-memory process

In this work, the UMAT program of SMP three - dimensional constitutive equation was developed, with the help of user material subroutine interface provided by the ABAQUS platform. Based on the correctness of the program, a series of numerical simulation of porous SMP plate were carried out. The entire process of numerical simulation has four steps. In the first step, a loading rate 1.5mm/min at 343k is applied displacement into 1mm. In the second step, temperature is reduced from 343K to 313K at the rate of 4.5k/min and the model remains loaded 1mm until the cooling process completed. In the third step, the model keeps the temperature 313k state and relieved the displacement, completing the unloading process at the low constant temperature. In the fourth step, temperature is rises from 313K to 343K at the rate of 4k/min without external load. The recovery process is completed with the heating. The time-displacement curve of whole thermodynamic cycle and deformation process of structure were shown in Fig. 2. Despite the presence of a small residual stress, the final residual strain is very small, so it can be considered to achieve a reasonable recovery effect.

4.2 Effect on the recovery of the different applied displacement
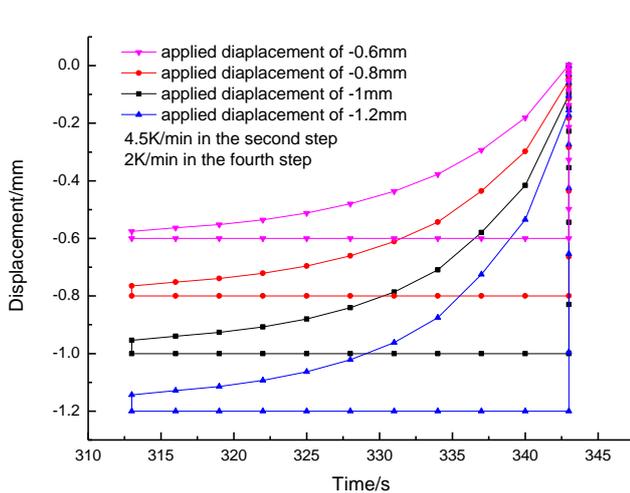
Figure 3. Temperature-displacement curves in different displacement load
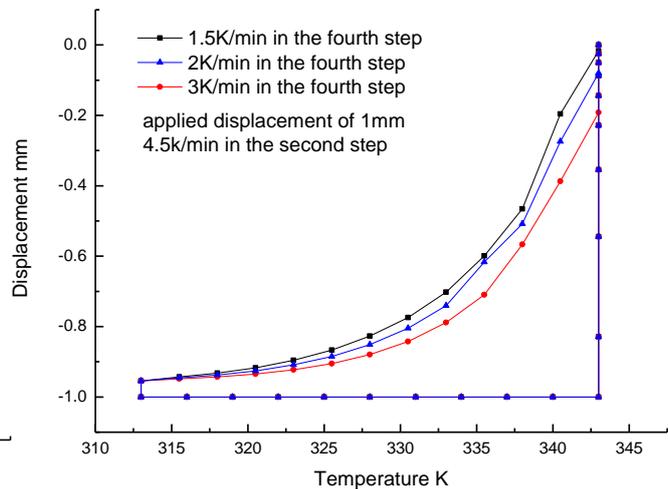
Figure 4. Temperature-displacement curves in different heating rate

Due to the viscoelastic properties and the temperature dependence of the SMP material, different applied displacement has a direct impact on the final curve. The effects that different applied displacement loads on the central region of porous SMP plate are studied. The four throughout the thermodynamic cycle are contrasted. In four experiments, the heating rate and cooling rate maintain constantly. The different displacement is applied on the model. First, all displacement is 0 before being applied displacement in four groups. Again, the loaded specimen is maintained while reducing the temperature of specimen. Then, part of the strain is restored during unloading. Finally, the results of recovery are different because of different applied displacement. As shown in Fig. 3 different applied displacement has great influence on the final residual strain. Residual strain becomes smaller with displacement load reduced from1.2mm to 0.6mm.,

4.3 Effect on the recovery of the different heating rate

As the temperature-dependent material properties of SMP, different heating rate for the restoration of the loaded displacement has impact. Keeping constant applied displacement, the heating rate impact the recovery of residual strain. As shown in Fig. 4, the characteristic of the recovery process varies from fast to slow and the turning point follows the glass transition temperature. It is the main

reason viscosity coefficient and delay time and other parameters are reduced at high temperature. With heating rate reduced from 3k/min to 1.5k/min, residual strain is decreasing. Whether full recovery of residual strain or not is directly affected by different heating rates. When the heating rate is less than 1.5k/min, the residual strain nearly completely recovered. It is mainly reason that the SMP is a viscoelastic material.

4.4 Effect on the recovery of the different thickness

In the structural design of the porous plate, it is inevitable to choose a reasonable thickness and the hole opening ratio based on recovery rate of residual strain and mechanical performance's requirements therefore study of influence of this parameter is important. As shown in Fig. 5, the thickness of plate were set to 2mm, 3mm, 4mm, 5mm, 6mm when the hole opening ratio , heating and cooling rate are constant. The greater the thickness of the porous SMP plate, the smaller the strain in thickness direction. The strain needed to restore is reduced in this case. Therefore, different thicknesses of porous SMP plate have a significant impact for the final residual strain in the stage of heating and recovery.

4.5 Effects of structural parameters on the recovery process

The structure should select the appropriate the hole opening rate in order to save costs. Whether different hole opening ratio will affect result of recovery or not. Thickness of model is 3mm and heating, and cooling rates held constant. The applied displacement of 5mm is imposed on the center region of porous SMP plate. The hole opening ratio of the porous plate were set to 0, 7.84%, 17.64%, 31.36%, and 49%. With the increase of the hole opening ratio, the mass of the porous SMP plate is reduced. As shown in Fig. 6, the effect of different hole opening ratio for results of recovery is small.
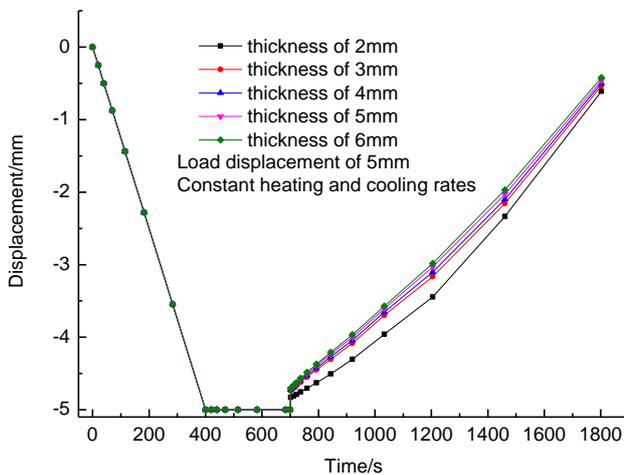
Figure 5. Time-displacement curves in different thickness
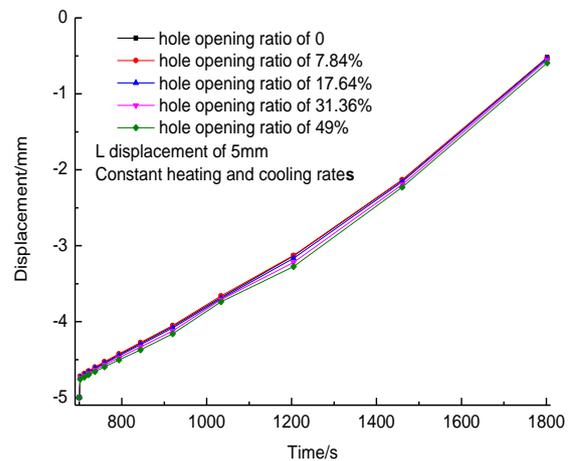
Figure 6. Time-displacement curves in different hole opening ratio

5 Conclusions

Based on the correctness of the SMP three-dimensional constitutive equation, a series of numerical simulation were carried out. The results show that the present 3D thermo-mechanical constitutive model can be used effectively to describe the complicated mechanical behavior of SMP. Experiments show that property of SMP material has strong temperature dependence. So the reasonable control of temperature is essential for the control of the mechanical behavior of SMP.

Whether full recovery of residual strain or not is directly affected on the thickness of the porous plates. However, the effect of hole opening ratio is small. The porous SMP plate has excellent thermo-mechanical property and the structure can be designed according to the needed.

References

Behl, M., Zotzmann, J. and Lendlein, A. (2010) Shape-Memory Polymers and Shape-Changing Polymers, *Adv. Polym. Sci***226**, 1-40.

Liu, B. F., Dui, G. S. and Zhu, Y. P. (2011) A Constitutive Model for Porous Shape Memory Alloys Considering the Effect of Hydrostatic Stress. *CMES: Computer Modeling in Engineering & Sciences* **78**, 247-276.

Hu, J. L. Meng, H. P., Li, G. Q. and Ibekwe, S. I. (2012) A review of stimuliresponsive polymers for smart textile applications. *Smart Mater. Struct* **21**, 053001.

Mather, P. T., Luo, X. F. and Rousseau, I. A. (2009) Shape Memory Polymer Research. *Annu. Rev. Mater. Res* **39**, 445-471.

Lendlein A, Kelch S. (2002) Shape-memory polymers. *Angew.Chem.Int.Ed* **41**,2034-2057.

Takahashi T, Hayashi N, Hayashi S. (1996) Structure and properties of shape memory polyurethane block copolymers. *J.Appl.Polym.Sci* **60**,1061-1069.

Yang J H, Chun B C, Chung Y C et al. Comparison of thermal/mechanical properties and shape memory effect of polyurethane block-copolymers with planar or bent shape of hard segment. *Polymer* **44**, 3251-3258.

Yang, B., Huang, W. M.; Li, C. and Chor, J. H. (2005) Effects of moisture on the glass transition temperature of polyurethane shape memory polymer filled with nano-carbon powder. *Eur. Polym. J* **41**, 1123-1128.

Keller P N, Lake M S, Codell D et al. (2003) Development of elastic memory composite stiffeners for a flexible precision reflector. *AIAA Journal*, 2003-1977.

Gall K R, Lake M S.(2003)Development of a shoekless thermally aetuated release nut using elastic memory composite material *.AIAAJournal*, 2003-1582.

Tobushi, H., Hashimoto, T. Hayashi, S. and Yamada, E. (1997) Thermomechanical constitutive modeling in shape memory polymer of polyurethane series. Intel. *Mater. Syst. Struct* **8**, 711-718.

Tobushi, H., Okumura, K.; Hayashi, S. and Ito, N. (2001) Thermomechanical constitutive model of shape memory polymer. *Mech. Mater* **33**, 545-554.

Liu, Y. P., Gall K, Dunn M L et al. (2006) Thermomechanics of shape memory polymers: Uniaxial experiments and constitutive modeling. *Int.J.plastieity* **22(2)**, 279-313.

Zhou, B. Liu, Y. J. and Leng, J. S. (2009) Finite Element Analysis on ThermoMechanical Behavior of Styrene-Based Shape Memory Polymers. *Acta Polym. Sin* **6**, 525-529.

Shi G H, Yang Q S, He X Q, Kim M L.(2013) A three-dimensional constitutive equation and finite element method implementation for shape memory polymers. CMES: Computer Modeling in Engineering & Sciences. **90(5)**, 339-358.

Relations between the Poynting and axial force-twist effects

Dong Wang, M.S. Wu*

School of Mechanical and Aerospace Engineering, Nanyang Technological University, 50 Nanyang Avenue, Singapore 639798

*Corresponding author: mmswu@ntu.edu.sg

Abstract

The relations between the Poynting effect, in which a cylinder elongates or contracts axially under torsion, and the axial force-twist effect, in which the twist of a torsionally loaded cylinder is affected by the axial loading, are investigated using second-order elasticity for an elastic homogeneous cylinder. The explicit expressions for the two effects and their relations are presented. The relations show that under tension: (a) negative Poynting effect implies negative axial force-twist effect, (b) positive axial force-twist effect implies positive Poynting effect, whereas (c) the converse statements are not true. Further results show that (a) the Poisson ratio captures the difference between the two effects, and (b) reduced elastic coefficients, which uniquely characterize the effects, lead to universal relations between the effects and the applied loading. Both effects also exhibit a strong inverse power law dependence on the radius.

Keywords: Axial force-twist effect, Poynting effect, torsion-axial loading, second-order elasticity

Introduction

Soft materials may exhibit complex nonlinear behavior such as the Poynting effect, in which a cylinder elongates or contracts axially under torsion. Poynting (1909) experimentally found that some metals exhibited the positive effect, i.e., they elongated axially under torsion. Recently, Janmey et al. (2007) found that networks of semiflexible biopolymers such as actin, collagen, fibrin and neurofilaments, exhibited the negative Poynting effect.

Wang and Wu (2014) showed that in contrast to the Poynting effect, an axial force-twist effect may also exist. It refers to their theoretical result that the twist of a cylinder under combined torsion and axial loading can be affected by the axial loading. The axial force-twist effect can also be positive or negative. The former means that both the twists produced by the axial loading and torsion are in the same direction, while the latter means that the twists produced by them are in the opposite directions. Though Wang and Wu (2014) presented the solutions for the Poynting and axial force-twist effect, the relations between them were not investigated.

This paper focuses on these relations, from which some fundamental conclusions can be drawn. The dependence of the two effects on the linear and nonlinear elastic constants is also studied. The organization of the paper is as follows. The derivation of the relations is first presented, followed by numerical results, a further discussion, and a set of conclusions.

Relations between the Poynting effect and the axial force-twist effect

Figure 1 shows a cylinder of length L and radius R under combined axial loading P and torsion T. Here P represents either a tensile or compressive stress. The materials are nonlinear elastic, isotropic and homogeneous. The initial coordinates of a particle of the cylinder are chosen as (r, θ, z). The strain energy density of Murnaghan (1951) is adopted, i.e.:

$$W = \frac{\lambda + 2\mu}{2} J_1^2 - 2\mu J_2 + \frac{l + 2m}{3} J_1^3 - 2mJ_1J_2 + nJ_3,$$ (1)

where λ and μ are the second-order and l, m, n the third-order elastic constants, respectively, and J_1, J_2, and J_3 are the strain invariants of the Lagrangian strain **E**. The detailed solutions of the stress and displacement fields are given in Wang and Wu (2014). For the purpose of deriving the relations between the effects, the results on the axial and circumferential displacements from the earlier paper are given below.

The axial displacement u_z under pure torsion loading can be written as:

$$u_z = Dz,$$ (2)

where D is the Poynting effect coefficient given by:

Figure 1. A homogeneous elastic cylinder with radius R and length L under combined torsion T and axial loading P.

334

$$D = -C_D \frac{T^2}{4\pi^2 R^6}, \qquad C_D = \frac{n\lambda + 4\mu m + 4\lambda\mu + 8\mu^2}{\mu^3(3\lambda + 2\mu)}. \qquad (3)$$

Note that a change in the direction of T does not change the sign of D. The parameter C_D is a reduced coefficient of the four elastic constants. It uniquely characterizes the quadratic relation between the Poynting effect and T. If a modified Poynting effect coefficient $\bar{D} = D / C_D$ is defined, then a universal relation between $\bar{D}$ and T can be obtained:

$$\bar{D} = -\frac{T^2}{4\pi^2 R^6}. \qquad (4)$$

Furthermore, the circumferential displacement under combined axial loading P and torsion T is:

$$u_\theta = u_\theta^L + u_\theta^{NL}, \qquad (5)$$

where u_θ^L represents the linear twist due to torsion T:

$$u_\theta^L = \frac{2T}{\pi R^4 \mu} rz, \qquad (6)$$

and u_θ^{NL} represents the nonlinear twist associated with the axial force-twist effect:

$$u_\theta^{NL} = -\frac{PT(n\lambda + 4\mu m + 6\lambda\mu + 8\mu^2)}{2\pi R^4 \mu^3(3\lambda + 2\mu)} rz. \qquad (7)$$

The axial force-twist effect coefficient can be defined as:

$$H = \frac{u_\theta^{NL}}{u_\theta^L} = -C_H \frac{P}{4}, \qquad C_H = \frac{n\lambda + 4\mu m + 6\lambda\mu + 8\mu^2}{\mu^2(3\lambda + 2\mu)}. \qquad (8)$$

C_H is a reduced coefficient which characterizes the relation between H and P. It is similar in form to C_D. By defining the modified axial force-twist effect coefficient $\bar{H} = H / C_H$, a universal linear relation between $\bar{H}$ and P can be obtained:

$$\bar{H} = -\frac{P}{4}. \tag{9}$$

It can be seen from Eq. (7) that the axial force-twist effect only exists under combined axial loading and torsion, i.e., $P \neq 0$ and $T \neq 0$. Eq. (8) implies that if H is positive, u_θ^{NL} has the same direction as u_θ^L and the axial force twist effect is positive; otherwise, it's negative. Because H depends on P and not T, two further observations can be made from Eq. (8):

(1) Change of the direction of T does not change the sign of H.
(2) Change of the sign of P changes the sign of H.

Eqs. (3) and (8) show that materials with different elastic constants can have the same Poynting effect or the axial force-twist effect, provided the reduced coefficients of these materials are the same. Another observation of Eq. (3) is that for a particular μ, if m and n are chosen in a way that makes $(n + 4\mu)/3 = (4m + 8\mu)/2$, or $8\mu + 6m - n = 0$, then λ has no influence on the Poynting effect. A similar conclusion can be made for H on the basis of Eq. (8). If m and n are chosen such that $(n + 6\mu)/3 = (4m + 8\mu)/2$, or $6\mu + 6m - n = 0$, then λ has no influence on the axial force-twist effect.

The relation between H and D in dimensionless form can be obtained easily from Eqs. (3) and (8):

$$\frac{H}{P/4\mu} = \frac{D}{T^2/4\pi^2 R^6 \mu^2} - \frac{2\lambda}{3\lambda + 2\mu}. \tag{10}$$

Since $v = \lambda/(2\lambda + 2\mu)$, the above equation can be rewritten as:

$$\frac{H}{P/4\mu} = \frac{D}{T^2/4\pi^2 R^6 \mu^2} - \frac{2v}{1+v}. \tag{11}$$

The term on the left-hand side represents the axial force-twist effect coefficient normalized by the axial loading, while the first term on the right-hand side represents the Poynting effect normalized by the torsion. An explicit relationship between the axial force-twist effect and the Poynting effect is thus established.

Since v is positive generally, several conclusions can be drawn from Eq. (11), assuming that the axial loading P is tensile:

(a) If $D < 0$, then necessarily $H < 0$,

(b) If $H > 0$, then necessarily $D > 0$.

(c) If H and D have different signs, then necessarily $H < 0$ and $D > 0$.

It should be emphasized that the converses of (a) and (b) are not true, i.e., $H < 0$ does not necessarily imply $D < 0$, and $D > 0$ does not necessarily imply $H > 0$. A further observation is that the sign of H will change if the sign of P changes. Thus for the case of compressive axial loading, the above three conclusions should be changed to:

(d) If $D < 0$, then necessarily $H > 0$,

(e) If $H < 0$, then necessarily $D > 0$.

(f) If H and D have the same sign, then necessarily $H > 0$ and $D > 0$.

The Poisson ratio plays a key role since the difference between the normalized H and the normalized D is the term $2\nu/(1+\nu)$. This difference reaches its maximum when $\nu = 0.5$, i.e., the material is incompressible.

Finally, the size dependence of the Poynting effect can be judged from Eq. (3) to be inversely proportional to the sixth power of the cylinder radius. For the axial force-twist effect, Eq. (7) shows that the maximum circumferential displacement ($r = R$) is inversely proportional to the third power of the cylinder radius. Hence, the Poynting effect is relatively more important than the axial force-twist effect for small cylinders, and the reverse holds for large cylinders.

Numerical results

This section focuses on the influence of the elastic constants on the Poynting effect and the axial force-twist effect. The elastic constants of the soft materials were adapted from Wang and Wu (2013, 2014) for poly(acrylic acid) (PAA) gels and capillary muscles, respectively, and Catheline et al. (2003) for an agar-gelatin. The geometry of the cylinder is fixed as $R = 0.002$ m and $L = 0.01$ m. The applied axial loading and torsion may vary for different figures.

Fig. 2 plots the $H = 0$ and $D = 0$ contours in the $\mu-\nu$ space, for $m = -2420$ kPa and $n = -2350$ kPa. The axial loading P is chosen as positive. It can be seen that the $\mu-\nu$ space is partitioned into three regions: Region I with $H > 0$ and $D > 0$, Region II with $H < 0$ and $D < 0$ and Region III with $H < 0$ and $D > 0$.

Several interesting phenomena can be observed, in agreement with the conclusions (a) to (c) stated above. First, negative Poynting effect ($D < 0$) implies negative axial force-twist effect ($H < 0$) as shown in Region II. However, the converse is not true, i.e., negative axial force-twist effect ($H < 0$) does not imply negative Poynting effect ($D < 0$) necessarily, as shown in the small Region III. Secondly, positive axial force-twist effect implies positive Poynting effect (i.e., $H > 0$ means $D > 0$, as shown

Figure 2. Contours of the Poynting effect coefficient $D = 0$ (dashed line) and axial force-twist effect coefficient $H = 0$ (solid line) in $\mu-\nu$ space for a homogeneous elastic cylinder. The contours partition the space into three regions.

in Region I). However, the converse is not true. Positive Poynting effect does not imply positive axial force-twist effect (i.e., $D > 0$ does not necessarily imply $H > 0$ as shown in region III). Moreover, when the two effects differ in sign, the Poynting effect must be positive and the axial force-twist effect must be negative (as shown in Region III). Region III, where the two effects have different signs, is generally small, suggesting that only careful choices in the material parameters can lead to different signs for the two effects.

Fig. 3 plots H and D against the Poisson ratio ν. The material parameters are based on those of polymers with $\mu = 10.3$ kPa, $m = -24.2$ kPa and $n = -23.5$ kPa. The loadings are $P = 10$ kPa and $T = 300$ kPa·m^3. It can be seen that when ν increases, both H and D decrease from positive to negative monotonically. Thus, the Poisson ratio can be an important parameter in controlling the two effects. Secondly, the magnitudes of H and D are of the order of 10^{-1}, suggesting that the nonlinear effects can be significant. Note that ν_1 and ν_2 are the particular Poisson ratios which make $H = 0$ and $D = 0$, respectively. This figure further shows that (a) if $H > 0$, then $D > 0$, as shown when $\nu < \nu_1$, (b) if $D < 0$, then $H < 0$, as shown when $\nu > \nu_2$, and (c) if H and D have different signs, then $H < 0$ and $D > 0$, as shown when $\nu_1 < \nu < \nu_2$.

Fig. 4 shows how the linear elastic constants μ and ν affect the Poynting effect and the axial force-twist effect. The parameters are $m = -360$ kPa, $n = 20$ kPa, $P = 10$ kPa and $T = 1000$ kPa·m^3. It can be seen that there exists a μ_1 for which H is independent of ν. Similarly, there exists a μ_2 for which D is independent of ν. As mentioned above,

Figure 3. Dependence of H and D on the Poisson ratio ν, with $\mu = 10.3$ kPa, $m = -24.2$ kPa, and $n = -23.5$ kPa.

Figure 4. Dependence of H and D on the shear modulus μ for different Poisson ratios $\nu = 0.1$, 0.3, 0.4 and 0.49, with $m = -360$ kPa and $n = 20$ kPa.

μ_1 and μ_2 can be determined from the equations $(n + 6\mu_1)/3 = (4m + 8\mu_1)/2$ and

$(n + 4\mu_2)/3 = (4m + 8\mu_2)/2$, respectively, yielding $\mu_1 = 363.3$ kPa and $\mu_2 = 272.5$

kPa. A further observation is that the negative H and D values appear to have upper bounds, while the positive values are unbounded. More generally, however, D or H may either have a positive or negative bound, depending on the values of m and n.

Fig. 5 shows how the nonlinear elastic constant m can significantly influence both

effects. Here H and D are plotted against μ for $m = \pm 2 \times 10^6$, $\pm 10^6$ and 0 kPa. The

other elastic parameters are $\lambda = 60$ kPa and $n = -23.5$ kPa. The loadings are $P = 0.01$

kPa and $T = 10$ kPa·m^3. For this set of parameters, increasing μ will decrease the magnitudes of the coefficients. Secondly, both effects are positive for negative m and

negative for positive m. Changing the sign of m will change the sign of both H and D. Thirdly, decreasing the magnitude of m will also decrease the magnitudes of H and D. The magnitudes can reach the order of 10^{-2} to 10^{-1} when μ is small; thus the nonlinear behavior can be significant when the material is very soft with a small μ.

Fig. 6 plots H and D versus μ for the same sets of m, with $\lambda = 35700$ kPa and $n = -23500$ kPa. The loadings are $P = 0.01$ kPa and $T = 10$ kPa·m^3. The nonlinear effects are different from those shown in Fig. 5. For m positive, both H and D decrease to a negative maximum and subsequently decrease slowly to zero with increasing μ. However, for m negative, they decrease monotonically to zero with μ.

Figure 5. Dependence of H and D on the shear modulus μ for $m = -2\times10^6$, -10^6, 0, 10^6, 2×10^6 kPa, with $\lambda = 60$ kPa and $n = -23.5$ kPa. The loadings $P = 0.01$ kPa and $T = 10$ kPa·m^3.

Figure 6. Dependence of H and D on the shear modulus μ for $m = -2\times10^6$, -10^6, 0, 10^6, 2×10^6 kPa, with $\lambda = 35700$ kPa and $n = -23500$ kPa. The loadings $P = 0.01$ kPa and $T = 10$ kPa·m^3.

Discussion

Many biological materials, from soft to hard, are subjected to complex loading in their physiological environment. A few examples are described here. Arterial walls associated with human brain aneurysms were subjected to combined extension, torsion and inflation in finite element studies, in order to mimic the real physiological conditions (Tóth et al., 2005). The behavior of lumbar spinal units under torsion, compression and flexion/extension were also experimentally studied (Haberl et al., 2004). It is also well-known that articular cartilage is subjected to combined compression and shear during normal activities (Mansour, 2003). Fatigue tests were conducted on cylindrical bovine cortical bone specimens under axial, torsional and combined axial-torsional loadings (Vashishth et al., 2001). Finite extension and torsion were applied on capillary muscles in order to characterize their behavior under physiological conditions (Criscione et al., 1999).

Because of the prevalence of combined loadings, the Poynting effect and the axial force-twist effect may be highly relevant. In particular, large stresses may be generated by both effects if the specimen is confined in one way or another, i.e., the additional axial and rotational displacements are restrained. These large stresses can, for instance, alter the overall force balance and the cytoskeleton structure of cells, or the movement of a human red blood cell through narrow capillaries. The diameter of a human red blood cell is 7.0-8.5 μm, while that of narrow capillaries is smaller than 3 μm (Bao and Suresh, 2003).

The effects can also be utilized in the design of devices such as actuators and sensors. One can imagine a bio-inspired polymer actuator based on the axial force-twist effect, i.e., a torsionally loaded cylinder may generate an additional output twist, if subjected to an input axial force. By carefully selecting the elastic parameters of the materials and the structural dimensions, the amount of twist can be increased significantly and the desired output can be achieved.

Conclusions

Explicitly expressions for the Poynting effect, the axial force-twist effect and their relation are presented in this paper. The dependence of the relation on elastic constants is investigated.

The results show that under a tensile stress P, (a) negative Poynting effect implies negative axial force-twist effect, (b) positive axial force-twist effect implies positive Poynting effect, and (c) if the two effects differ in sign, the Poynting effect must be positive and the axial force-twist effect negative. The loadings P and T are such that (d) changing the direction of T will not change the sign of both effects, and (e) changing the direction of P will change the direction of the axial force-twist effect. Moreover, the Poynting and axial force-twist effects exhibit a very significant size

dependence, respectively of the inverse sixth and third power of the cylinder radius.

Reduced elastic coefficients characterize universal relations between the effects and the applied loadings. The elastic constants μ, ν and m have significant influence on the magnitude and direction of the Poynting and axial force-twist effects. For certain combinations of elastic constants, changing the sign of m can directly change the sign of the two effects. The two effects may have a positive or negative bound, depending on the elastic constants. From the perspective of material design, the elastic constants are thus of vital importance.

References

Bao, G. and Suresh, S. (2003) Cell and molecular mechanics of biological materials. *Nature Materials* **2**, 715-725.

Catheline, S., Gennisson, J.L. and Fink, M. (2003) Measurement of elastic nonlinearity of soft solid with transient elastography. *Journal of the Acoustical Society of America* **114**, 3087-3091.

Criscione, J.C., Lorenzen-Schmidt, I., Humphrey, J.D. and Hunter, W.C. (1999) Mechanical contribution of endocardium during finite extension and torsion experiments on papillary muscles. *Annals of Biomedical Engineering* **27**, 123-130.

Haberl, H., Cripton, P.A., Orr, T.E., Beutler, T., Frei, H., Lanksch, W.R. and Nolte, L.P. (2004) Kinematic response of lumbar functional spinal units to axial torsion with and without superimposed compression and flexion/extension. *European Spine Journal* **13**, 560-566.

Janmey, P.A., Mccormick, M.E., Rammensee, S., Leight, J.L., Georges, P.C. and Mackintosh, F.C. (2007) Negative normal stress in semiflexible biopolymer gels. *Nature Materials* **6**, 48-51.

Mansour, J.M. (2003) Biomechanics of Cartilage, in *Applied Kinesiology*, Lippincott Williams and Wilkins, Philadelphia, pp. 66-79.

Murnaghan, F.D. (1951) Finite Deformation of an Elastic Solid. John Wiley, New York.

Poynting, J.H. (1909) On pressure perpendicular to the shear planes in finite pure shears, and on the lengthening of loaded wires when twisted. *Proceedings of the Royal Society of London Series A-Containing Papers of a Mathematical and Physical Character* **82**, 546-559.

Tóth, B.K., Raffai, G. and Bojtar, I. (2005) Analysis of the mechanical parameters of human brain aneurysm. *Acta of Bioengineering and Biomechanics* **7**, 3-23.

Wang, D. and Wu, M.S. (2013) Second-order elasticity of soft multilayer capsules: Universal relations and parametric studies. *International Journal of Engineering Science* **73**, 17-32.

Wang, D. and Wu, M.S. (2014) Poynting and axial force-twist effects in nonlinear elastic mono- and bi-layered cylinders: Torsion, axial and combined loadings. *International Journal of Solids and Structures* **51**, 1003-1019.

Vashishth, D., Tanner, K.E. and Bonfield, W. (2001) Fatigue of cortical bone under combined axial-torsional loading. *Journal of Orthopaedic Research* **19**, 414-420.

Computational Otolaryngology – Modeling and Simulations of Nasal Airflows

*Heow Pueh Lee[1], Jianhua Zhu[1], Shu Jin Lee[2] and De YunWang[3]

[1]Department of Mechanical Engineering, National University of Singapore, Singapore
[2] Division of Surgery, Yong Loo Lin School of Medicine, National University of Singapore, Singapore
[3] Department of Otolaryngology, National University of Singapore, Singapore

*Corresponding author: mpeleehp@NUS.EDU.SG

Abstract

There is a growing interest in the use of computational fluid dynamics (CFD) modeling and simulations with detailed nasal airway models constructed from CT or MRI images for surgical planning and assessment of outcome of nasal surgery. As compared to rhinomanometry and acoustic rhinometry, which provide quantitative information only of nasal airflow, resistance, and cross sectional areas, CFD enables additional informations of airflow passing through the nasal cavity that help visualize the physiologic impact of alterations in intranasal structures. Therefore, it becomes possible to quantitatively measure, and visually appreciate, the airflow pattern (laminar or turbulent), velocity, pressure, wall shear stress, particle deposition, and temperature changes at different flow rates, in different parts of the nasal cavity. The effects of both existing anatomical factors, as well as post-operative changes, can be assessed. With recent improvements in CFD technology and computing power, there is a promising future for CFD to become a useful tool in planning, predicting, and evaluating outcomes of nasal surgery. This presentation presents some of the recent studies performed in our research group related to clinical applications or assessment of nasal obstructions, Turbinectomy, Functional Endoscopic Sinus Surgery, Cleft palates and Rhinoplasty. I will also highlight our recent works in terms of modeling and simulations of nasal airflows due to the long duration wearing of N95 respirators.

Keywords: Computational Fluid Dynamics, Nasal Airways, Sinuses, Nasal Obstructions, Functional Endoscopic Sinus Surgery, Turbinectomy, N95 respirators

Introduction

The nose is the guardian angel of the respiratory tract. It has several important physiological functions which include air-conditioning, filtrating the inspired air, and smell. It also plays an important defence function, as the nose is the first place where foreign pathogens and allergens contact the host. To serve these important functions, a functional or patent nasal passage is essentially needed. A better understanding of how the nose functions is important and related to the treatment of respiratory related medical conditions such as snoring, Obstructive Sleep Apnea (OSA), and the contraction of diseases such as SARS and Bird Flu. Nasal obstruction is also a common complaint which is difficult to quantify clinically. The etiologic factors for nasal obstruction include anatomic variations of the nose and various local and systemic diseases. Hence, objective assessment of the nasal airway will aid diagnosis, treatment, research and medico-legal documentation.

During the last 10 to 20 years, attempts have been made to quantify nasal functionality or patency. This is an exciting time in the field of rhinology with regard to basic research and clinical practices. The technological advancements and greater insight into understanding of normal nasal functions

and their relationship with other organs has spurred research in this field. Functional nasal airway measurement is important not only for research in nasal physiology, but is a useful and important diagnostic tool for patients with nasal or its related disorders. The sensation of nasal obstruction, nasal resistance and minimal cross sectional area are three distinct entities measuring nasal patency, which are closely related to each others. In a combination of these techniques, it represents the state of the art for a functional and quantitative study of the nasal airway. However, to date, it is still not possible to quantitatively measure the changes in nasal physiologic functions, which are caused by various type and degree of anatomical or pathophysiological changes of the nasal cavity and patency.

The relevance of objective assessment of nasal resistance and patency has been documented in many rhinological situations: (a) In routine rhinological practice, it is particularly useful in the differential diagnosis and management of common nasal diseases, e.g., allergic or non-allergic rhinitis and nasal septum deformity. Objective testing is useful in appropriate decision-making. Routine employment of objective quantitative assessment will result in improved diagnosis and the medical management. (b) Measurement of nasal airway patency is very much appreciated in rhinological research, since it provides a quantitative piece of information on the changes of nasal mucosal response to intranasal application with allergens and any kind of physical and chemical agents. (c) It provides objective data on the nature of nasal airway that can be used as the medicolegal documentation.

More recently, the advantages of computational fluid dynamics (CFD) enable researchers to obtain detailed flow patterns in the human upper airway by reconstructing models from computed tomography (CT) and Magnetic Resonance Imaging (MRI) images, which has become a new reliable trend of nasal airway exploration. However, there are very few reported comparisons of computational fluid mechanics simulation results with clinical measurements of airways conditions, let alone using the common techniques of rhinomanometry and acoustic rhinometry. It is generally agreeable that experimental investigations in nasal airway models give only limited compatibility with actual physiological conditions of the nose although the simplified experimental model allows for detailed examinations of various complicating effects. The lack of correlation studies between the engineering simulations and clinical practices, especially clinical examinations on actual nasal airway is a major reason as to why the engineering simulations have not found its way into the clinical examinations and also the lack of CFD related publications in clinical otolaryngology journals.

The present research currently being carried out in the Department of Mechanical Engineering, Department of Otolaryngology and Department of Surgery within the National University of Singapore is therefore an attempt to bridge this gap by correlating the engineering simulations to actual physiological functions of the nasal airways. The present study will also attempt to establish the relations between the engineering simulation results and the measurements from rhinomanometry and acoustic rhinometry. These two methods provide complementary and important objective information concerning the nasal airway

Rhinomanometry is well established as a useful clinical method for objective assessment of nasal patency. Nasal resistance to airflow is calculated from measurements of nasal airflow and transnasal pressure. Sstandardization of rhinomanometry was established in 1983 and accepted worldwide. The nasal resistance is calculated from the measurement of the nasal airflow at a fixed transnasal pressure point and is expressed in Pa/cm3/s. Three types of rhinomanometry can be used: active anterior rhinomanometry (AAR), active posterior rhinomanometry (APR), and passive anterior rhinomanometry (PAR). AAR uses a facemask, one nostril is sealed off with adhesive tape; a hard

plastic tube passed through the tape measuring the nasopharyngeal pressure. It is a dynamic test that studies nasal ventilation; showing the nature of the air stream and a difference in the shape of the inspiratory and expiratory limbs at the individual nasal cavity. This method is well standardized and it is the most common and accurate method for clinical use. The major disadvantage of this method is that it cannot be performed in the presence of a septal perforation or a complete unilateral nasal blockage. For PAR, a fixed airflow of 250cm3/s is blown through a nozzle into one nostril. The pressure induced by the nasal resistance to this airflow at a given level of the nozzle is measured. This is an easy and fast procedure that can be used even in infants. Nasal patency measurement on one side is still possible when the other side is completely blocked. It has been proven to be a qualitative as well as a quantitative method for the objective evaluation of the degree of nasal obstruction in patients with allergic rhinitis following the nasal allergen challenge. However, the sensitivity and accuracy of this measurement is somewhat lesser than AAR.

In contrast, acoustic rhinometry does not measure airflow parameters but explores the geometry of the nasal cavity. The principle of acoustic rhinometry is that an audible sound (150-10,000 Hz), propagated in a tube, is reflected by local changes in acoustic impedance. This method provides estimates of cross-sectional endonasal areas and of the endonasal volume, and helps to define objectively the structural and mucosal components of the nasal passage. Since its introduction, there has been an explosion of research using this tool. Due to the rapid acquisition of data which can be completed in a minute, it has become a valuable clinical and research tool. Patient tolerance is excellent even in children.

The proposed research is multi-disciplinary and multi-physics in nature and will not be possible without close collaborative efforts among engineers, clinicians and biomedical researchers. By examining the engineering analysis results with the respective clinical observations and measurements performed using rhinomanometry and acoustic rhinometry for actual nasal models of various medical conditions, the engineering analysis can be developed into a virtual objective tool to be used in clinical otolaryngology and to help in a more accurate diagnosis and documentation of nasal conditions.

The main objective of this research is therefore to bridge the gap between the engineering analysis and simulations, and the clinical practices for assessing objectively the physiological conditions of nasal airways. The specific tasks are as follows:

- To establish computational models of human nose based on CT and MRI imaging for various nasal conditions.
- To investigate the relationship between the geometrical configuration of the nasal cavity and the simulated results for airflow and thermodynamic behaviour in healthy conditions and obstructive nasal diseases.
- To correlate the simulation results to the nasal conditions assessed by rhinomanometry, acoustic rhinometry and clinical examinations.
- To establish an in vitro nasal airway model with actual human nasal functional and physiological characteristics.
- To develop a virtual objective tool to be used in clinical otolaryngology and to help in a more accurate diagnosis and documentation of nasal conditions.

Methodology

The research performed involves both numerical and experimental studies. The main tasks for the numerical studies are modeling and simulations. Three dimensional surface models were first created by image segmentation of CT or MRI scans provided by the collaborators from clinicians. This was usually done using several commercially available softwares such as MIMICS 12.1 (The Materilize Group, Leuven, Belgium), Hypermesh 9.0 (Altair Engineering, Bangalore, India), and TGrid 4.0 (ANSYS, Inc., Canonsburg, PA, USA). Smoothing of the highly corrugated surfaces due to digitization was performed to facilitate computational meshing of the three-dimensional model. Smoothing of the boundary surface in such relatively larger 3-D nasal cavity will not affect the main flow pattern inside, but will help to decrease computational effort and increase computational efficiency. Computational fluid dynamics studies were carried out using Fluents for pure flow simulations or Adina for fluid structural interaction such as the interaction of soft plate and airflow for the analysis of obstructive sleep apnea. These tasks were performed at the Applied Mechanics Laboratory, National University of Singapore.

The human subject studies, conducted in two phases were conducted at the Investigator trial unit, National University Hospital. A total of 100 volunteers (47 for the first phase and 53 for the second phase) was recruited for this study. The first phase was conducted over a period of five days and the second phase was conducted for another three days. The typical tasks involved the completion of study questionnaires, the performance of acoustic rhinometry, rhinomanometry, mucociliary clearance evaluation using saccharin tests, and smell test using the University of Pennsylvania Smell Identification test kits,

Several nasal cavity models have been created based on CT or MRI of patients. The initial phase of the research focused on the analysis of nasal blockade or Inferior Turbinate Hypertrophy on the aerodynamic pattern and physiological functions of the turbulent airflow. Subsequent studies extended the flow simulations to particle deposition related to drug delivery as well as the thermal effect. The study was then extended to the other geometric effects such as septal deviation and septal perforation as well as the effect of various surgical procedures such as inferior turbinectomy and towards the later part of the research, on the effect of Functional Endoscopic Sinus Surgery (FESS) as well as nasal fractures. All the numerical simulations were examined and co-related to clinical observations and therefore most of the findings were in fact published in clinical journals related to otolaryngology such as the Laryngoscope, Rhinology, Journal of Aerosol Medicine and Pulmonary Drug Delivery, American Journal of Rhinology and Allergy, Respiratory Physiology and Neurobiology. As highlighted by one of the reviewers, there has been "a boom in the number of publications describing the flow patterns in the nasal cavity of various nasal pathologies, especially in clinical journals".

Results and Discussion

A major portion of the project is to examine the effect of various pathological conditions on the nasal functions. Two pathological conditions related to septal deviation and inferior turbinate hypertrophy are the main focus of the study.

In the nose model with septal deviation, major changes in the pattern of inspiratory airflow (e.g., flow partitioning and nasal resistance, velocity and pressure distributions, intensity and location of turbulence), wall shear stress and increasing of total negative pressure through the nasal cavity were demonstrated qualitatively and quantitatively. For the healthy nose, the main airflow occurs in the middle of the airway (between the inferior and middle turbinates around the septum), with the peak

velocity in this area. On the contrary, in the model with a septal deviation, the main airflow is found passing through the floor (left side) and superior part (right side) of the nasal cavity. In the healthy nose, the area with the highest intensity of turbulent flow was found in the functional nasal valve region, but it became less apparent or even disappeared in the septal deviation one. For healthy case, the pressure decreases smoothly along the airway from nostril to the nasopharyngeal region. However, for deviated nose, greater pressure gradient or abrupt pressure jump is found posterior to the site of deviation (more in the right side than the left). For the deviated nose, the air flow mainly goes through the upper passageway in the cavity; two noticeable vortex areas are detected. Vortex areas mainly locate near the nasal valve region and dorsal regions near the superior cavity. With a large inspiratory flow rate of 34.8 L/min, maximum transient velocities of 5.69 m/s and 7.67 m/s are detected inside the healthy and septum deviation noses respectively. A higher shear stress distribution is found in the floor of the left nasal cavity than that in the right side. This CFD study provides detailed information of the aerodynamic effects of nasal septal deviation on nasal airflow patterns and their associated physiological functions.

In the healthy nose, the main respiratory air stream occurs mainly in the middle of the airway, accompanied by a diffused pattern of turbulent flow on the surface of the nasal mucosa. The peak value of turbulent flow is found in the functional nasal valve region. However, this aerodynamic flow pattern has partially or completely changed in the models with enlarged inferior turbinate. With an inhalation flow rate of 34.8 L/min, a maximum velocity of 5.69 m/s, 7.39 m/s and 11.01 m/s, are detected respectively in the healthy, moderately and severely obstructed noses. Both total negative pressure and maximum shear stress has increased by more than three and two times, respectively, in severely blocked noses compared to the healthy one. Data of this study provide quantitative and quantitative information of the impact of inferior turbinate hypertrophy on the aerodynamic pattern and physiological functions of nasal airflow. By including the model of turbulent airflow, the results of this experimental study will be more meaningful and useful in predicting the aerodynamic effects of surgical correction of the inferior turbinate hypertrophy.

Nasal patency is an essential condition that has a major impact on particle deposition. For the healthy one, due to its complete existence of the MCA, the particle number escaping the cavity is the largest one; for moderately blocked nose, due to its relatively larger penetration via MCA (two thirds left) than the severely blocked one (one third left), the particle number escaping the cavity is larger than the severely blocked one. The particle percentage escaping the nasal cavity decreased to less than a half and one tenth for the moderately and severely blocked noses. Decreasing of flow rate and particle diameter increased the escaping ratio; however, zero escaping percentage was detected with the absence of airflow and the effect was less noticeable when the particle diameter was too small. The existence of inspiratory flow and head tilt angle helped to change the particle escaping ratio for the healthy nose; however, such changes were not significant for the moderately and severely blocked noses. Thus it is noted that the patients with nasal obstruction need to ensure the presence of a middle inspiratory flow rate, when using the nasal spray device for higher escaping ratio. To insert the spray deeper with particles beyond the MCA region or a decongestive treatment to increase the MCA may be necessary in patients with moderately or severely blocked noses as the effects of the changes of nose flow rate, initial particle velocity, particle diameter and head tilt angles for drug delivery are limited or even negligible.

Another major portion of the research is the evaluation of surgical procedure on nasal airflow and particle deposition. The aim for one of these studies was to evaluate the effects of inferior turbinate surgery on nasal airway heating capacity using computational fluid dynamics (CFD) simulations. Heat transfer simulations were performed for a normal nasal cavity and others with severely enlarged inferior turbinates, before and following three simulated surgical procedures: (1) resection

of the lower third free edge of the inferior turbinate, (2) excision of the head of the inferior turbinate and (3) radical inferior turbinate resection. The models were run with three different environmental temperatures. The changes of airflow pattern with the reduction of inferior turbinate were found to affect heat transfer greatly. However, the distribution of wall heat flux showed that the main location for heat exchange was still the anterior region. Under the cold environment, the nasal cavities with the head of inferior turbinate reduction were capable of heating the inspired air to 98.40% of that of the healthy one; however, for the case with lower third of inferior turbinate excised, the temperature was 11.65% lower and for the case with radical inferior turbinate resection, 18.27% lower temperature compared to the healthy nasal cavity. The healthy nasal cavity is therefore deemed to be able to warm up or cool down the inspiratory airflow under different environmental temperature conditions; for the nasal cavities with turbinate surgeries, partial inferior turbinate reduction can still sustain such heating capacity. However, too much or total turbinate resection may impair the normal function of temperature adjustment by nasal mucosa.

In another piece of work, we examined the effect of nasal surgery on particle deposition and drug delivery. Intranasal medications are commonly used in treating nasal diseases. However, technical details of the correct usage of these medications for nasal cavity with obstruction are unclear. A 3-dimensional model of nasal cavity was constructed from MRI scans of a healthy human subject. Nasal cavities corresponding to healthy, moderate and severe nasal obstruction were simulated by enlarging the inferior turbinate geometrically, which was documented by approximately one third reduction of the minimum cross sectional area for the moderate and two thirds for the severe obstruction. The discrete phase model based on steady state computational fluid dynamics was used to study the gas-particle flow. The results were presented with drug particle (from 7×10^{-5} to 10^{-7} m) deposition distribution along the lateral walls inside these three nasal cavities and comparisons of the particle ratio escaping from the cavity were also presented and discussed. Nasal patency is an essential condition that had the most impact on particle deposition of the factors studied; the particle percentage escaping the nasal cavity decreased to less than a half and one tenth for the moderately and severely blocked noses. Decreasing of flow rate and particle diameter increased the escaping ratio; however, zero escaping percentage was detected with the absence of airflow and the effect was less noticeable when the particle diameter was very small. The existence of inspiratory flow and head tilt angle helped to improve the particle escaping ratio for the healthy nose; however, such changes were not significant for the moderately and severely blocked noses. Therefore, when using an intranasal medication, it is advisable to have a moderate inspiratory airflow rate and small size particles in order to improve particle escaping ratio. Various head positions suggested by clinicians do not seem to improve drug escaping ratio significantly for the nasal cavities with inferior turbinate hypertrophy.

Another surgical procedure that has been examined is Functional Endoscopic Sinus Surgery (FESS). For the nasal cavities with FESS, the ostia may be enlarged and some ethmoid cells may also be removed. Therefore more airflow may be directed into the paranasal sinus regions, upper ethmoid and sphenoid sinus regions and thus affects their physiological functions. Moreover, the airflow inside was inappropriate to be assumed quasi-steady due to increased local Womersley and Strouhal numbers and increased local inertial effects.

Thus the study focused on the evaluation of the effects of a particular FESS case on unsteady nasal aerodynamic flow patterns and to attempt to answer whether there were existences of circulations inside the sinus regions. Different from previous steady state flow modeling, the study focused on unsteady airflow characteristics to investigate inertia effects. A 3-dimensional model of nasal cavity was constructed from CT scans of a patient with FESS interventions on right side nasal cavity. CFD simulations were then carried out for unsteady aerodynamic flow modeling inside the nasal cavity

as well as the sinuses. Comparisons of the local velocity magnitude and streamline distributions inside the left and right nasal cavity and maxillary sinus regions were carried out. Due to the FESS procedures in the right nasal cavity, existences and distributions of local circulations (vortexes) were found to be significantly different for a same nasal airflow rate but at different acceleration, deceleration or quiet phases in the maxillary sinus region on the FESS side. Due to inertia effects, local internal airflow with circulation existences was continuous throughout the whole respiration cycle. With a larger peak inspiration flow rate, the airflow intensity inside the enlarged maxillary sinus increased significantly.

The current model was the first attempt to investigate such abnormal aerodynamic behavior around the FESS and sinus regions with transient airflow conditions. Although the current model is based on one particular individual and may not be representative for FESS procedures, it still provides qualitative and quantitative information for better knowledge of the internal geometrical changes with surgical effects on aerodynamic differences. Existences of airflow circulations or re-circulations inside the main nasal cavity volume were found for healthy nasal models. It was usually believed that the airflow going into the sinus region was negligible for a healthy nose, and possibly a small amount of airflow into the sinuses through ostia helped to maintain stable mass (nitric oxide) and humidity exchanges for the nasal cavity after the current manner of FESS, relatively larger amount of airflow were circulated into these regions and affected the local aerodynamic flow patterns. From engineering perspective, the open airway passage from the paranasal sinus regions to the nasal cavity behaves as a hollow cavity, which results in relatively smaller local pressure distribution with flow instability. Moreover, it was usually believed that the sinus regions were parts of the possible locations for productions of nitric oxide (NO). A relatively stable concentration of NO is needed to maintain a sterile local environment for a nasal breathing and respiratory systems to ensure normal functions related to human being host defense and local immunologic reactions. An uncinate process probably has a protective role in preventing deposition of bacteria and allergens in the sinuses during the inspirational phase. Thus for the current FESS case with the removal of uncinate process, the existences of continuous airflow and circulations inside the sinus regions may affect local nitric oxide re-distributions and its stability. Moreover with such circulations, bacterial infection rates may increase with possible depositions of bacteria and allergens in the sinus after FESS.

The first phase of the human subject study demonstrated the concordance of all the various clinical tests used in the assessment of nasal physiology in allergic and healthy volunteers. Impaired nasal function due to temperature change was found to be more exaggerated in allergic rhinitis. Dynamic interval functional assessment at (e.g. rhinomanometry, nasal spirometry) can better accentuate and demonstrate nasal hypersensitivity as compared to traditional single-point static testing. The preliminary findings of the second phase of human subject study showed that the airway is likely to be blocked more with smaller nasal valve minimal cross sectional area why lying down compared to body upright position. This will cause a shift of the frequency spectrum of the breathing sound. Detailed analysis is being carried out.

The team is currently embarking on the experimental study of "Objective Assessment of the Effects of long-duration wearing of N95 and surgical facemasks on upper airway functions" funded by the Ministry of Health involving 100 healthcare workers from the local hospitals and clinics.

A list of the publications arising from the research is appended in the list of references.

Conclusion

The existing CFD technology enables detailed study and objective measurements of the physical characteristics of the airflow within the anatomically exact numerical human nose models, which are constructed based on CT scan or MRI images. Although the predicted results from CFD study are derived from complex calculations which may not represent real life conditions, it does provide clinically useful and additional information in additional to typical clinical assessment such as rhinomanometry and acoustic rhinometry of the breathing airflow. Information that could be obtained from modeling and simulations include airflow pattern (laminar or turbulent), velocity, pressure, wall shears stress, and temperature changes under different flow rates in different part of nasal cavity, as well as the effect of various anatomical factors (e.g., septal deviation or perforation, inferior turbinate hypertrophy and turbinectomy, nasal and maxillary bone fracture, opening of rhinosinuses after surgery, and etc.). With the improvement in CFD technology and research, there is a promising future of this technique to become a useful tool in predicting outcomes and planning of nasal surgery.

Acknowledgement

References

Chen, X.B., Lee, H.P., Chong, F.H.V. and Wang, D.Y. (2009) Assessment of Septal Deviation Effects on Nasal Air Flow-A Computational Fluid Dynamics Model. Laryngoscope **119**, 1730-1736.

Chen, X.B., Lee, H.P., Chong, F.H.V. and Wang, D.Y. (2010) Impact of inferior turbinate hypertrophy on the aerodynamic pattern and physiological functions of the turbulent airflow - a CFD simulation model. Rhinology, **48**, 163-168.

Chen, X.B., Lee, H.P., Chong, F.H.V. and Wang, D.Y. (2010) Numerical Simulation of the Effects of Inferior Turbinate Surgery on Nasal Airway Heating Capacity. American Journal of Rhinology & Allergy, **24**, 118-112.

Chen, X.B., Lee, H.P., Chong, F.H.V. and Wang, D.Y. (2010) A Computational Fluid Dynamics Model for Drug Delivery in a Nasal Cavity with Inferior Turbinate Hypertrophy. Journal of Aerosol Medicine and Pulmonary Drug Delivery **23**, 329-338.

Chen, X.B., Leong, S., Lee, H.P., Chong, F.H.V. and Wang, D.Y. (2010) Aerodynamic Effects of Inferior Turbinate Surgery on Nasal Airflow. Rhinology **48**, 1-6.

Chen, X.B., Lee, H.P., Chong, F.H.V. Wang, D.Y. (2011) Assessments of nasal bone fracture effects on nasal airflow: A computational fluid dynamics study. American Journal of Rhinology & Allergy **25**, 39-43.

Chen, X.B., Lee, H.P., Wang D.Y. and Chong, F.H.V. (2011) Aerodynamic Characteristics inside the Rhinosinosal Cavity after Functional Endoscopic Sinus Surgery". American Journal of Rhinology & Allergy **25**, 388-392.

Chen, X.B., Lee, H.P., Chong, F.H.V. and Wang, D.Y. (2012) Drug Delivery in a Nasal Cavity with Functional Endoscopic Sinus Surgery : A Computational Fluid Dynamics Study". Journal of Laryngology and Otology **126**, 487-494.

Lee, H.P., Poh, H.J., Chong, F.H.V. and Wang, D.Y. (2009) Changes of airflow pattern in inferior turbinate hypertrophy – a computational fluid dynamics model. American Journal of Rhinology & Allergy **23**, 153-158.

Lee, H.P., Revanth, R.G., Chong, F.H.V. and Wang, D.Y. (2009) Effects of septal perforation on Nasal airflow: computer simulation study". Journal of Laryngology and Otology **24**, 1-7.

Lee, H.P. and Wang, D.Y. (2011) Objective Assessment of Increase in Breathing Resistance of N95 Respirators on Human Subjects. Annals of Occupational Hygiene **55**, 917-921.

Lee, H.P., Revanth, R.G., Chong, F.H.V. and Wang, D.Y. (2013) Comparison Between Effects of Various Partial Inferior Turbinectomy options on Nasal Airflow – a Computer Simulation Study". Computer Methods in Biomechanics and Biomedical Engineering **16**, 112-118.

Lee, S.J, Liong, Y.K.J., Tse, K.M. and Lee, H.P. (2010) Biomechanics of the Deformity of Septal L-struts". Laryngoscope **120**, 1508-1515.

Lee, W.L., Pahala, G.J., Tan, Z.J., Le, D.V., Lee H.P. and Khoo, B.C. (2011) Muco-ciliary transport: effect of mucus viscosity, cilia beat frequency and cilia density. Computers & Fluids, **49**, 214-221.

Leong, S., Chen, X.B., Lee, H.P. and Wang, D.Y. (2010) A review of the implications of computational fluid dynamic studies on nasal airflow and physiology. Rhinology **48**, 139-145.

Liong, Y.K.J., Lee, S.J. and Lee H.P. (2010) Deformation of nasal septum during nasal trauma". Laryngoscope **120**, 1931-1939.

Pahala, G.J., Tan, Z., Le, D.V., Lee, H.P. and Khoo, B.C. (2012) Three-dimensional numerical simulation of human pulmonary cilia in the periciliary liquid layer by the immersed boundary method". Computers & Fluids **67**, 130-137.

Pahala, G.J., Tan, Z., Le, D.V., Lee, H.P. and Khoo, B.C. (2014) A Numerical Study of Muco-iliary Transport under the Condition of Diseased Cilia". Computer Methods in Biomechanics and Biomedical Engineering DOI:10.1080/10255842.2013.864285.

Revanth, R.G., Lee, H.P., Chong, F.H.V. and Wang, D.Y. (2009) Indicators for the correct usage of intranasal medications – A computational fluid dynamics study. Laryngoscope **119**, 1975-1982.

Wang, D.Y., Lee H.P. and Gordon, B.R. (2012) Impacts of fluid dynamics simulation in study of nasal airflow physiology and pathophysiology in realistic human three dimensional nose models. Clinical and Experimental Otorhinolaryngology **5**, 181-187.

Zhu, J., Lim, K.M., Thong, K.T.M., Wang, D.Y. and Lee H.P. (2014) Assessment of airflow ventilation in human nasal cavity and maxillary sinus before and after targeted sinonasal surgery: a numerical case study. Respiratory Physiology & Neurobiology **194**, 29-36.

Zhu, J., Lee, H.P., Lim, K.M., Lee, S.J., Teo, L.S.L. and Wang, D.Y. (2013) Inspirational airflow patterns in deviated noses: a numerical study". Computer Methods in Biomechanics and Biomedical Engineering **16,** 1298-1306.

Zhu, J., Lee, H.P., Lim, K.M., Gordon, B.R. and Wang, D.Y. (2012) Effect of Accessory Ostia on Maxillary Sinus Ventilation: a Computational Fluid Dynamics (CFD) Study. Respiratory Physiology & Neurobiology **183**, 91-99.

Zhu, J., Lee, H.P., Lim, K.M., Lee, S.J., Teo, L.S.L. and Wang, D.Y. (2012) Passive movement of human soft palate during respiration: a simulation of 3D fluid/structure interaction. Journal of Biomechanics **11**, 1992-2000.

Zhu, J., Lee, H.P., Lim, K.M., Lee, S.J. and Wang, D.Y. (2011) Evaluation and comparison of nasal airway flow patterns among three subjects from Caucasian, Chinese and Indian ethnic groups using computational fluid dynamics simulation. Respiratory Physiology & Neurobiology **175**, 62-69.

Zhu, J., Lim, K.M., Gordon, B.R., Wang, D.Y. and Lee, H.P. (2014) Effects of Anterior Ethmoidectomy, With and Without Antrotomy and Uncinectomy on Nasal and Maxillary Sinus. Airflows: a CFD Study. J. Med. Biol. Eng. doi: 10.5405/jmbe.1710.

A Computer Based Objective Grading System for Facial Paralysis

*Heow Pueh Lee[1], Saurabh Garg[1] and Shu Jin Lee[2]

[1]Department of Mechanical Engineering, National University of Singapore, Singapore
[2] Division of Surgery, Yong Loo Lin School of Medicine,
National University Health System, Singapore

*Corresponding author: mpeleehp@NUS.EDU.SG

Abstract

A computer assisted objective grading system based on the asymmetry analysis of patient three-dimensional surface models obtained from a 3D facial imaging system is developed for facial paralysis diagnosis. To overcome the subjectivity of traditional diagnosis methods, facial asymmetry grading is carried out according to fine registration result of the original and mirror facial meshes which do not rely on any landmarks. As a higher order property of 3D surfaces, the Principal, Gaussian and Mean Curvatures, Shape Index and Curvedness are introduced for grading the asymmetry of the faces as descriptions of the surface local features besides examining the local differences in distance between the original and mirrored surfaces. The present results show that the proposed approach is able to objectively assess the asymmetrical of human faces as well as to detect the improvement of the patients quite well.

Keywords: Computational Methods, Facial expression, 3D curvatures, Bell's Palsy

Introduction

The face is three dimensionally complex consisting of many tissues namely, bone, soft tissue including fat and muscle, and overlying skin. The soft tissue of the face is unique to the whole body as the underlying muscles are densely enveloped in a continuous layer making up the superficial musculoaponeurotic system or SMAS, which has direct attachments to the facial skin allowing for displacement of the facial soft tissues for facial animation. This ability of animation makes us distinct from other primates and animals as we can produce a myriad of facial expressions. Another unique feature is the presence of two opposite sides which in the ideal situation are mirror images against a vertical midline that bears the nose. This kind of facial structure is the exception rather than the rule as a spectrum of asymmetry exists. The extremes of asymmetry can be due to facial trauma, facial nerve injury and also in birth defects. The physical and psychological suffering is immense as it is both a socially crippling disorder and also a possible functional disorder that can lead to problems with respect to vision, breathing, articulation, mastication, vocalization and many others. Common problems in childhood are cleft lip and palate seen in 1:700 live births and in adulthood facial trauma that is often in association with injury to the soft tissue and underlying bone. The loss of the soft tissue attachments between the bone, muscle and skin leads to descent of the affected site's soft tissues; together with loss of bone support and soft tissue scarring results in asymmetry and deformity. The descent of soft tissues besides bony and skin changes is seen with the commencement of facial aging from the later 20's and continues relentlessly to produce the stigmata of the aged face.

To date, the study on the soft tissue of the face has largely been limited to the two aspects, namely the bony changes via CT scan imaging and skin changes via external limited views of the other

lying skin such as with 2D photography. This has allowed for a large body of work on bone changes in congenital and acquired disease and a myriad of surgical options for skin closure in congenital and acquired disease such as post tumor resection. Bone work and skin closure alone are not enough for the restoration of facial symmetry. Bone restoration creates a stable bony platform and skin closure very often creates an intact skin pocket. The contour of the soft tissue contents which are three dimensional may be different from the normal side leading to postoperative asymmetry. We believe that the next step to improving facial surgery and analysis as a whole is the understanding of the whole 3D contour of the facial soft tissues allowing for mirroring and comparison to the contralateral side. This will objectively quantify areas of excess and deficit, descent and contour change which can then be addressed appropriately for a full reconstruction.

Appraisal or rating of facial deformity is usually required to assess the severity of the initial facial injury as well as the degrees of improvement due to the surgical or medical treatments. The appraisal or grading procedures can be broadly divided into qualitative and quantitative methods. Many of the existing methods are qualitative in nature and the grading often suffers from the lack of consistency among different clinicians as well as the difficulty in evaluating marginal changes during the follow-up treatments. The quantitative method analyses the extent of abnormal facial morphology and the degree of disproportion through various facial measurements (for example, Whittle, 2004, Farkas et al., 1993, Sandy, 2003) based on numerical data obtained from photographs or digital models obtained using 3D imaging equipment.

With the proliferation of the use of computer systems and digital photography, a computer based system for the objective assessment of facial deformity or asymmetry based on 3D digital models is a viable solution for the clinicians having to assess the severity of the initial facial injury as well as the degrees of improvement due to the followed up medical treatments. The proposed solution will provide tangible improvements in health services delivery, patient care and health outcomes with a more repeatable and objective assessment of the patients' conditions.

On the mathematical methods for facial analysis in terms of asymmetry or facial deformity, many methods have been proposed for characterizing the parameters of a beautiful face and analyzing the differences between male and female faces, and between faces of people with different racial backgrounds. These methods can be grouped into two main categories: anthropometric studies using either 2D or 3D data, and studies using geometric algorithms using 3D data. Wang et al. (2009) measured nasal length, nasal tip-projection, dorsal height, radix height, nasolabial angle, and nasofrontal angle on 2D photos. These measurements were used to compare the nose shapes of healthy young Koreans with those of rhinoplasty patients. Choe et al. (2006) used 22 standard measurements from 2D photos to compare the differences in nose shapes between Korean American women and North American white women. Szychta et al. (2011) analyzed 3D face models to compare the aesthetic results of post traumatic rhinoplasty. The analysis is based on two indices of the nasal proportions and four angles of the region computed from 18 anthropological points. Dong et al. (2010) used nine linear measurements, three angular measurements and seven proportions derived from a set of 17 landmark points on 3D face models to characterize Chinese nose and find the differences between males and females.

The main objective of the present research is to develop a computer based system based on 3D images taken from patients using 3D imaging technique for objective and quantitative assessment of facial and nasal deformity. The proposed computer based system will be developed and tested for three clinical conditions and applications – facial deformity due to stroke and cleft lip and palate, congenital nasal deformity or nasal deformity due to facial or nasal trauma, and objective assessment of the mid face infra-orbital hollow and upper and lower cheek fullness, lower face

lower lip fullness, lip wrinkles (at rest and dynamic) and the oral commissure for facilitating facial reconstructive surgery. As a preliminary report, we will focus on the analysis of the facial symmetry of healthy human subject as a benchmark for future research.

Methodology

For the proposed study, a total of 100 healthy human subject volunteers have been recruited for the initial study from NUS students and staff. The approval from the relevant Institutional Review Board has been obtained for this study. Another 100 volunteers will be selected from the patients visiting the clinics of Otolaryngology and plastic and reconstructive surgery at the National University Hospital. These human subjects and patients will form the database for the analysis. The 100 patients are expected to be recruited from patients with cleft lip and palate, facial paralysis and patients with facial deformity due to facial trauma.

For the first phase of this study, 100 healthy (without any facial deformities) participants were recruited. Participants involved in the study were at least 18 years old and students from the National University of Singapore. Among the scans, 55 were males and 45 were females. Most of the scans are Chinese with 3 Malays, 1 Indian and 1 Eurasian. Participants were required to remove their spectacles or clip up any hair that is covering their face to aid the measurement taking process that will be performed later on in the study. Before the facial scanning was carried out, 3 photographs were required to be taken from each participant, the frontal, profile and the oblique views (Figure 1).

Frontal View   Profile view   Oblique View

Figure 1 Frontal, Profile and Oblique View of Participant

The facial scanning was done using the Artec3D EVA white light scanner (Figure 2) acquired at the beginning of the proejct. It makes use of the white light technology which will not present any safety risk to people and does not produce the kind of magnetic radiation that can damage medical instruments. There is also no need for any special markers. Each facial scan lasts for about 20 seconds. The scans were taken in a quiet environment with good lighting to minimise the noise disturbances and to optimise the scans taken. The scans were taken at an optimal distance in order for them to be more accurate.

Figure 2: Artec3D Eva white light scanner

The facial scans were initially proposed to be taken with the participants' eyes open. However, due to the white light which may result in discomfort to the eyes and also due to the continuous blinking of the eyes by the participants, it might lead to inaccuracy and inconsistency of the scan. Hence, it was decided that the scans should be done with participants having their eyes closed. Two scans were also taken per participant and the scan of a higher quality and accuracy was chosen for data analysis.

The scanned data had to be processed before they can be used. Fine alignment procedure was done by applying fine serial registration on the scan for a quality model. Then the model had to go through global data registration where the global registration algorithm converted all one-frame surfaces into a single coordinate system using information on the mutual position of each pair of surfaces. Upon successful global registration, all the processed data could be fused into a polygonal 3D model. After which, tools in the software were used to correct any surface defects either due to scanning or registration errors. Lastly, texture will be applied to the scans.

Before proceeding to take the facial measurements from the 3D scans, the scans had to be orientated in the right direction with respect to the axis. All facial scans were normalized to natural head posture with the origin set at the subnasale point, the x-axis pointing left, from right to left eye; y-axis pointing vertically upwards, from chin to forehead; z-axis pointing outwards, in the nose direction.

Landmarks were then manually placed on the scans using Artec Studio 9 software as shown in Figures 3 to 7.
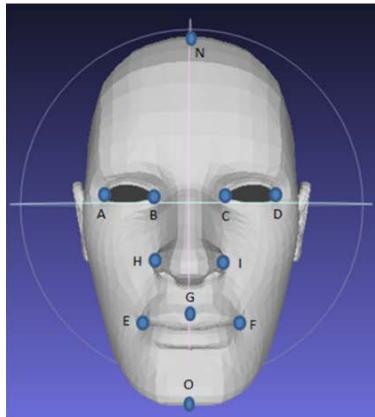
Figure 3
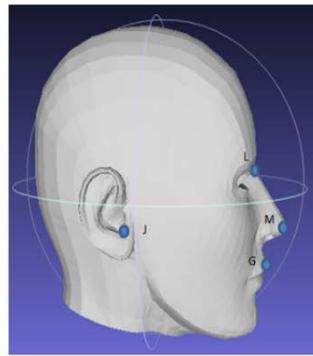
Figure 4
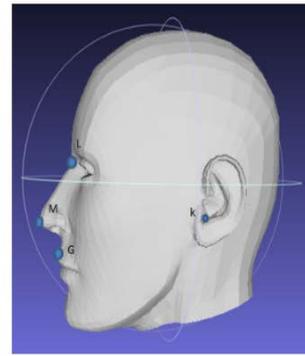
Figure 5
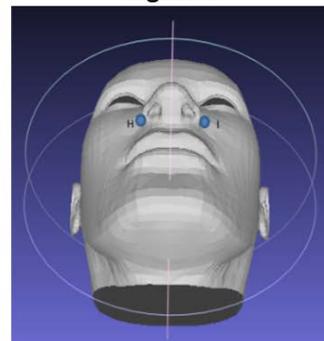
Figure 6
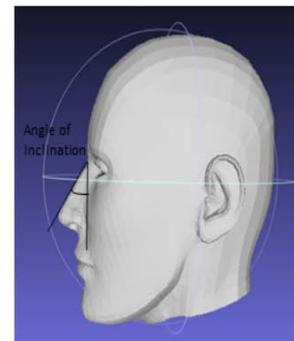
Figure 7

Figures 3-7 Landmarks on facial scan

In total, 16 landmarks were manually identified and placed on each scan. The x, y and z coordinates of each landmark were recorded (48 coordinates in total).

Brief Description of the landmarks are as follows:

Ectocanthion	: The point at which the outer ends of the upper and lower eyelids meet
Endocanthion	: The point at which the inner ends of the upper and lower eyelid meet
Cheilion	: The lateral most point at the angle of the lips
Labiale Superius	: Midpoint of the vermilion line of the upper lip
Alare	: The most lateral point on the ala of the nose
Intertragic notch and earlobe	: The small groove between the bump of cartilage between the ear and temple
Nasion	: The craniometrical point at the bridge of the nose where the frontal and nasal bones of the skull meet
Pronasale	: The most prominent point on the nasal tip
Trichion	: Point where the hairline meets the midpoint of the forehead
Gnathion	: Lowest median landmark on the lower border of the mandible

The coordinates were used to get the measurements for various dimensions of the facial features which includes the length of the right and left eyes, width and height of the face and others. There were 45 scans for females and 55 scans for males. The various dimensions measured for both the left and the right side of the face include the length of the eye, distance between the Intertragic Notch and Cheilion, distance between the intertragic notch and nasion, and lastly the distance between intertragic notch and ectocanthion.

Results and Discussion

The measurements for the various dimensions for the left and right side of the face did not tally. This suggests that all the participants do not have a perfectly symmetrical face.

	Average (mm)	Std. Deviation
Distance between the eyes	37.65	3.05
Length (right eye)	27.24	2.02
Length (left eye)	27.43	1.77
Face Height	184.16	7.96
Face Width	140.09	5.35
Height/Width of face	1.32	0.07
Length of Lips	44.94	3.38
Angle of inclination of nose	13.34	3.96

Table 1 Average and standard deviation for females

T-tests were performed on different groups such as distance between the length of the right and left eyes, distance between the Intertragic notch (Left and Right) and Nasion, Distance between Left and Right Intertragic notch and left and right Cheilion respectively and lastly distance between right and left Intertragic notch with right and left Ectocanthion respectively. An alpha value of 0.05 was used. From the t-tests, the P values are all more than 0.05, showing no statistical significance in all of the measurements.

	Average (mm)	Std. Deviation
Distance between the eyes	40.19	3.43
Length (right eye)	27.86	1.78
Length (left eye)	28.08	1.78
Face Height	192.75	10.16
Face Width	150.12	6.42
Height/Width of face	1.29	0.07
Length of Lips	47.84	4.18
Angle of inclination of nose	12.76	3.74

Table 2 Average and standard deviation for males

Comparison of the difference in dimensions between females and males were tested using the t-test. Dimensions included difference between the length of the right and left eyes, Difference in the distance between the Intertragic notch (Left and Right) and Nasion, difference in distance between

Left and Right Intertragic notch with Left and Right Cheilion respectively, difference in distance between Right and left Intertragic notch with right and left Ectocanthion respectively, difference in the inclination of nose, difference in the face height and lastly the difference in face width. The alpha value of 0.05 was used. From the t-tests, the P values are all more than 0.05 except for 2 which is the face height and the width of the face. The T-test for face height between the males and females obtained a value almost equal to zero, showing that the results are statistically significant. The mean of the face height for males is 192.75mm while the mean face height for females is 184.16mm as shown in Table 3.

t-Test: Two-Sample Assuming Unequal Variances		
	Face Height (Females)	Face Height (Males)
Mean	184.16	192.75
Variance	63.38	103.29
Observations	42	51
Hypothesized Mean Difference	0	
Df	91	
t Stat	-4.57	
P(T<=t) one-tail	7.61E-06	
t Critical one-tail	1.66	
P(T<=t) two-tail	1.52-05	
t Critical two-tail	1.99154	

Table 3 T-test for face height between females and males

The T-test for face width has also proven to be statistically significant with the p value being close to zero. The face width mean for males is 150.12mm while for females is 140.09mm. This is shown in table 4.

t-Test: Two-Sample Assuming Unequal Variances		
	Face Width (Females)	Face Width (Males)
Mean	140.0859524	150.1156863
Variance	28.63356614	41.26434902
Observations	42	51
Hypothesized Mean Difference	0	
df	91	
t Stat	-8.214317689	
P(T<=t) one-tail	6.88149E-13	
t Critical one-tail	1.661771155	
P(T<=t) two-tail	1.3763E-12	
t Critical two-tail	1.986377154	

Table 4: T-test for face width between females and males

Hence, this shows that the male has a greater face height as compared to females. The software developed was used to test a typical healthy subject as well as a patient with facial deformity. The software will be used in the coming study for patients.
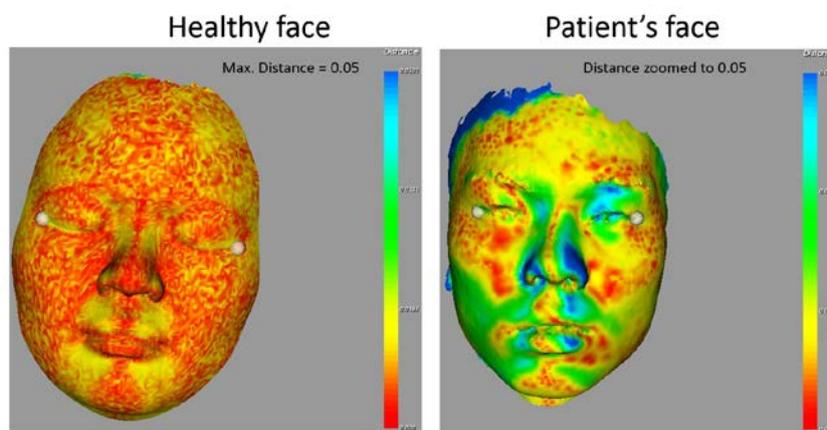
Distance between closest points in source and mirror faces

Figure 8 Sample output from the software developed

The software will be used in the coming study for patients.

Conclusions

From the results, it can be seen that there are no one with perfectly symmetrical faces even though the difference measured is not large. It was also observed that the face height and face width of males are generally larger than those of females. There is no other statistical significance between the other dimensions of facial measurements. 3D analysis using the Surface Analyzer indicated that RMS values of patients with facial deformities were generally larger than those of healthy participants. Distance Maps and Gaussian Maps also showed that faces of patients had higher asymmetry as compared to healthy participants.

It should be noted that when surgeons perform reconstructive surgery for patients with facial deformities, they should not do the surgery solely based on the symmetry of the face but rather also seek the opinions of the patients involved. In the era of modern orthodontics when the soft-tissue paradigm and patient perception often dictate the success of treatment outcomes, it is not acceptable for the orthodontist to simply identify problems and proceed with treatment. There is also a need to understand individual differences in the perception of face symmetry by surgeons, patients and those involved in the treatment process to get optimal results (Jackson, et al., 2013)

Acknowledgement

References

Choe, K.S., Yalamanchili, H.R., Litner, J.A., Sclafani, A.P., Quatela, V.C. (2006) The Korean American woman's nose: an in-depth nasal photogrammatic analysis. Arch Facial Plast Surg. **8(5)**, 319-323.

Dong, Y., Zhao, Y., Bai, S., Wu, G. and Wang, B. (2010) Three-dimensional anthropometric analysis of the Chinese nose. J Plast Reconstr Aesthet Surg. **63(11)**, 1832-1839.

Farkas, L.G., Hajnis, K. and Posnick, J. (1993) Anthropometric and anthroposcopic findings of the nasal and facial region in cleft patients before and after primary lip and palate repair. Cleft Palate Craniofac J **30**, 1-12.

Jackson, T. H., Mitroff, S. R., Clark, K., Proffit, W. R., Lee, J. Y., & Nguyen, T. T. (2013). Face symmetry assessment abilities: Clinical implications for diagnosing asymmetry. American Journal of Orthodontics and Dentofacial Orthopedics, 663-671.

Sandy, J.N. (2003) An aesthetic index for evaluation of cleft repair. Eur J Orthod **25**, 243-249.

Szychta, P., Rykała, J. and Kruk-Jeromin, J. (2011) Individual and ethnic aspects of preoperative planning for posttraumatic rhinoplasty. Eur J Plast Surg. **34(4),** 245–249.

Wang, J.H., Jang, Y.J., Park, S.K. and Lee, B.J. (2009) Measurement of aesthetic proportions in the profile view of Koreans. Ann Plast Surg. **62(2)**, 109-113.

Whittle, J. (2004) Preoperative anthropometric analysis of the cleft children's face: a comparison between groups. Int J Surg **2**, 91-95.

Global vs. Local Instabilities of Pure Bcc Iron

*K. Yashiro¹ and T. Yamane²

¹Graduate School of Engineering, Kobe University, JAPAN.
²Former Student of Graduate School of Engineering, Kobe University, JAPAN.
*Corresponding author: yashiro@mech.kobe-u.ac.jp

Abstract

As a series study that discusses the deformation criteria based on the atomic elastic stiffness (AES), we performed various molecular dynamics simulations of perfect bulk, thin plate/wire, and tilt/twist grain boundaries (GBs) of bcc Fe. Contrary to our previous results of fcc Ni and Al, bcc Fe can reach the global instability, where all the atoms has negative AES and the system can't follow up the stress increase anymore, prior to the local deformation or nucleation of lattice defects in the tensile simulation of bulk, plate and wires. We also discuss about the negative AESs of the tilt/twist GBs against various misorientation angles for comparison with the GB energy in the no-load equilibrium. Finally we perform tensile simulations on the (322) Σ11 and (111) S3 tilt GBs ([110] tilt axis) and revealed the change in the negative AES atoms and deformation morphology.

Keywords: Local lattice instability, Atomic elastic stiffness, Surface, Grain boundary, Bcc Fe

Introduction

We considered the onset of local structural change as "local lattice instability" and attempted to find the criteria based on the atomic elastic stiffness (AES), $B_{ij}^\alpha = \Delta \sigma_i^\alpha / \Delta \varepsilon_j$, or the deformation resistance at each atom point [Yashiro and Tomita (2001)]. Here, the elastic stiffness coefficients are originally defined as the stress-strain gradient in the nonlinear elasticity [Wallace (1972)]. In the case of brittle material Si, we can observe the local to global unstable phenomena, i.e. the unstable stress drop of the whole system is triggered by the emergence of the $\det B_{ij}^\alpha < 0$ atoms [Yashiro and Fujihara (2011)]. On the other hand, we can't conclude which is the first, the onset of precursor deformation or the emergence of $\det B_{ij}^\alpha < 0$ atoms in many cases of Ni and Al [Yashiro et al. (2006), Nishimura et al. (2010)]. In the present paper, we show the relationships between the AES and the system response of bcc-Fe, derived from our investigations on the AES in bulk, at the surfaces and grain boundaries. In order to eliminate the effect of thermal vibration in this discussion, all the simulations is implemented at extremely low temperature, T=0.1K. FS potential is adopted for bcc-Fe system.

Unstable behavior of perfect bulk under [001] tension

Figure 1 shows the change in the average, standard deviation of $\det B_{ij}^\alpha$, and the stress-strain curve of the perfect bulk Fe under [001] tension. The bulk is represented by a cubic periodic cell of 20x20x20 unit lattices. The change in the number of $\det B_{ij}^\alpha < 0$ atoms is also indicated in the lower figure. The tensile simulation is implemented by usual strain control in a quasi-static manner; all the atom position is scaled according to the small strain increment, $\Delta\varepsilon_{zz}$=1.0×10⁻⁶, at each time step of 1fs. Here, the cell lengths of L_{xx} and L_{yy} are also scaled to cancel the normal stress by the Poisson's contraction. The system shows the unstable stress drop at the strain of 0.1252 due to the emergence of internal disorder or the structural relaxation as seen in the error bars of standard deviation of $\det B_{ij}^\alpha$; however, all the atoms already become $\det B_{ij}^\alpha < 0$ at the smaller strain of ε_{zz}=0.0982, i.e. the crystal already reached the global instability. The stress begins to decrease

after that point, resulting in the stress-peak of smooth arch, without any local change in the lattice structure. Thus the crystal would show unstable elongation under stress control tension since it couldn't follow up the stress increase anymore. We can thus conclude that the global instability is preceding to the local structural change in the bulk bcc-Fe under [001] tension. Here, we can't track what happen after the global instability since we have no answer so far for the boundary condition about the external loading due to the time-scale gap.

Unstable behavior of thin plate and wire

Then we have investigated the effect of the initial structural inhomogeneity by surfaces, preforming tensile simulations on thin plates and wires eliminating the periodic boundary condition in the normal directions against the tensile axis. Figure 2 shows the results of the (100) and (110) surface plates under the [001] tension. The right diagram of the (110) plate shows same tendency as the perfect bulk above mentioned; the stress-strain curve shows smooth
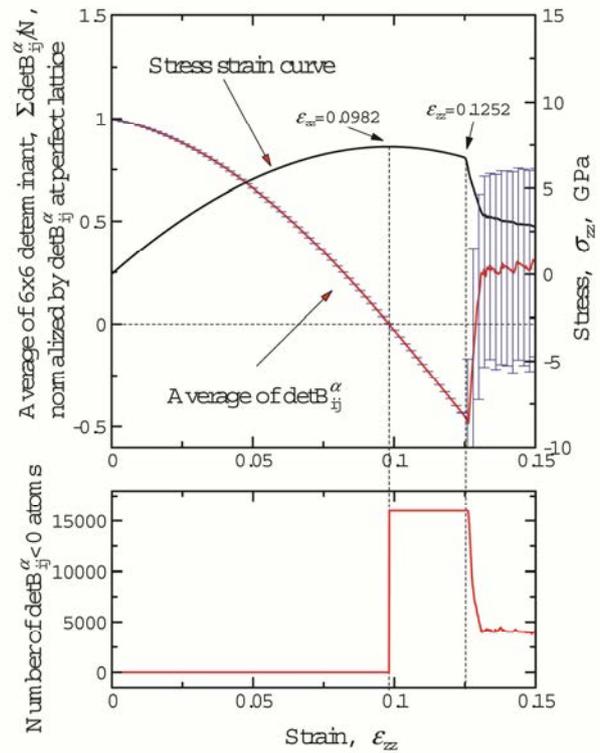
Figure 1. Change in the average of $\det B_{ij}^{\alpha}$, number of $\det B_{ij}^{\alpha} < 0$ atoms, and the stress-strain curve of bulk Fe under [001] tension.
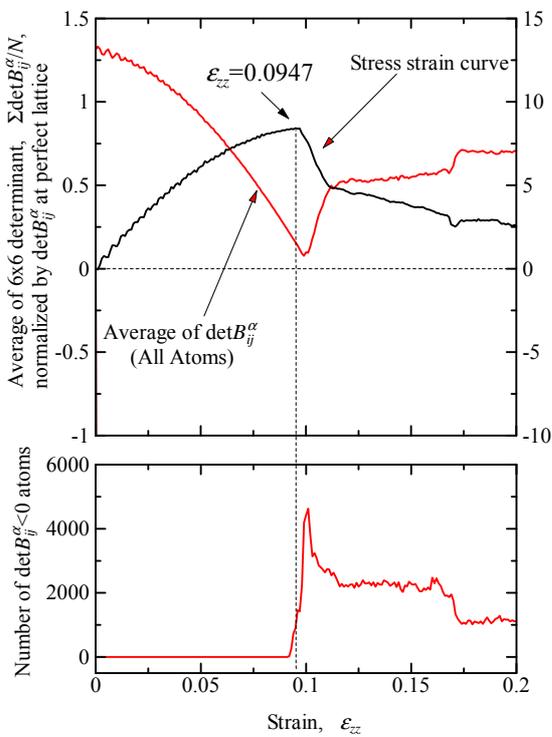
(a) (100) surface plate
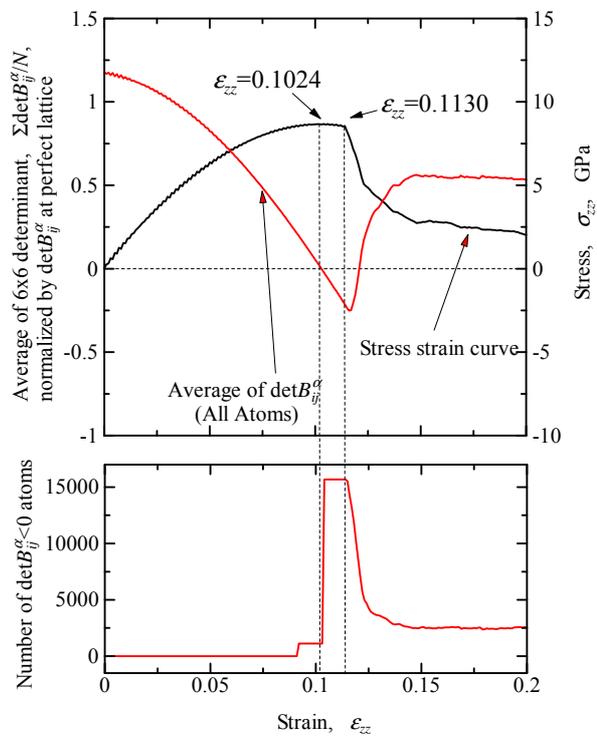
(b) (110) surface plate

Figure 2. Change in the average of $\det B_{ij}^{\alpha}$, number of $\det B_{ij}^{\alpha} < 0$ atoms, and the stress-strain curve of thin plate Fe under [001] tension.

arch and all the atoms becomes $\det B_{ij}^{\alpha} < 0$ at the peak, although the unstable stress drop by the internal structural change is observed at the later strain of $\varepsilon_{zz}=0.1130$. Thus we can deduce that the global instability is prior to the local deformation from the (110) surface. On the other hand, the unstable stress drop initiates at the strain of 0.0947 in the case of the (100) plate, despite of the positive value of the average of $\det B_{ij}^{\alpha}$ as often observed in fcc Ni and Al; i.e. the local deformation precedes to the global instability. The bottom of the left diagram reveals that $\det B_{ij}^{\alpha} < 0$ atoms emerge just before the stress drop and explode, then we can observe the onset of the local structural change and stress drop. Thus we can conclude that the local instability appears prior to the global one, in the case of (100) surface under the [001] tension. Both the wire models of (100)-(010) and (110)-($\bar{1}$10) surfaces also shows same tendency, that is, the local structural change occurs from the surfaces before the global instability point. On the other hand, in the case of the [$\bar{1}\bar{1}$2] tension, both the thin plate of (111) surface and the wire of (111)-($\bar{1}$10) surfaces shows the stress drop after the global instability or the smooth stress peak; thus we can deduce that the surfaces are not potential candidate of the local deformation, under the [112] tension of bcc Fe.

Atomic elastic stiffness of symmetrical tilt/twist grain boundaries

Figure 3 illustrates the number of $\det B_{ij}^{\alpha} < 0$ atoms and the grain boundary (GB) energy per unit area against the misorientation angle of the [110] tilt and the [001] twist GBs. Here infinite laminate structures of GBs are assumed under the periodic boundary conditions. The size of the simulation cells is different for each GBs due to the coincidence site lattice (CSL) so that we don't discuss about the averages of $\det B_{ij}^{\alpha}$. We may recognize same analogy between negative atoms and GB energy, e.g. there is no $\det B_{ij}^{\alpha} < 0$ atoms on the (112) Σ3 tilt and Σ3 twist GBs, which are well-known stable GBs as can be seen in the energy cusp. However, the number of $\det B_{ij}^{\alpha} < 0$ atoms shows more complicated change, contrast to the rather monotonic change of GB energy. For example, the (332) Σ11 tilt GB has many negative $\det B_{ij}^{\alpha}$ atoms despite of the energy cusp, while the (111) Σ3 tilt GB has few $\det B_{ij}^{\alpha} < 0$ atoms although there is no remarkable cusp in the GB energy curve.

(a) [110] tilt G.B.　　　　　　(b) [001] twist G.B.
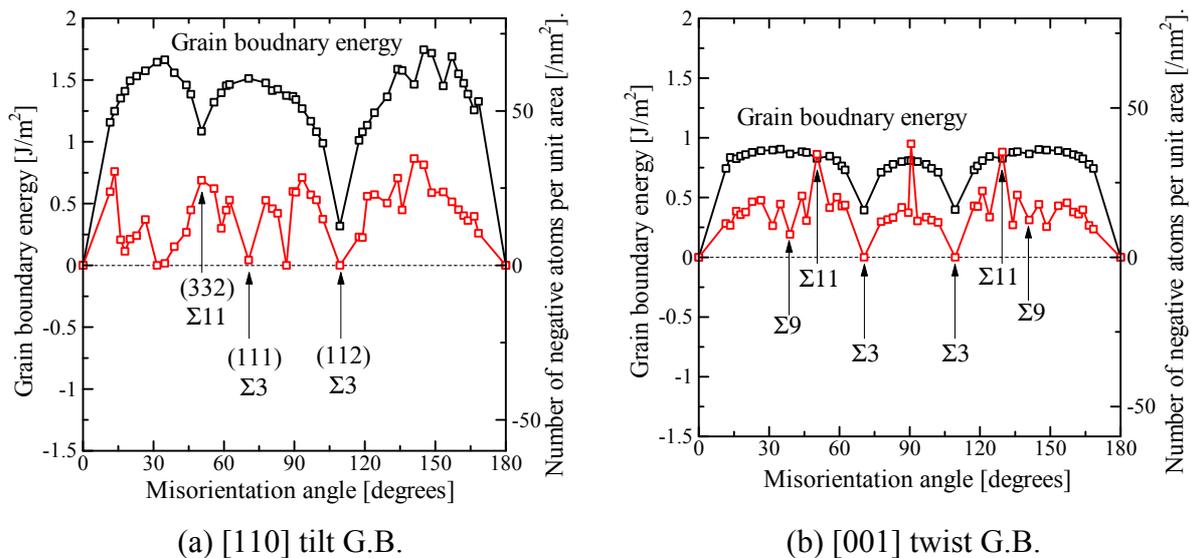
Figure 3. Relationships between grain boundary energy and misorientation angle, and the ratio of negative $\det B_{ij}^{\alpha}$ atoms.

(a) (332) Σ11 tilt G.B.　　　　　　　　　　　　(b) (111) Σ3 tilt G.B.
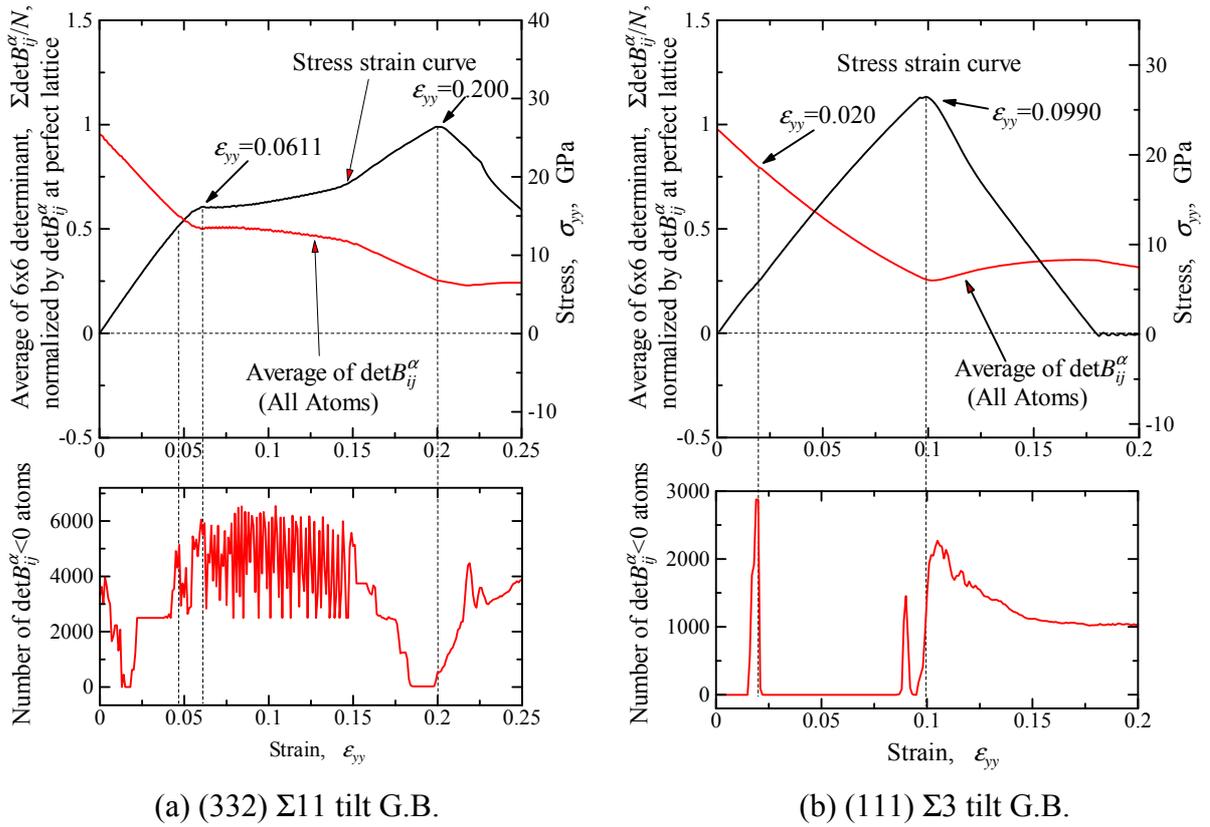
Figure 4. Stress-strain curve, the change in the average of $\det B_{ij}^{\alpha}$ and the number of $\det B_{ii}^{\alpha}<0$ atoms under tension (infinite laminate structure of tilt GBs).

We then performed tensile simulations on the (332) Σ11 and the (111) Σ3 tilt GBs, in the direction normal to the GB plane (y-axis). Here we didn't control the Poisson's contraction or the lateral strain ε' is fixed at zero during the simulations. The simulation results are summarized in Fig.4. As already indicated in Fig.3, the (332) Σ11 tilt GB has many negative atoms at the initial equilibrium; however, the negative atoms decrease and finally vanish by the external load during ε_{yy}=0~0.015, although there is no remarkable change in the stress-strain curve. Then atoms on the layers just above and below the GB planes become negative at the strain of 0.022, as recognized with the similar snapshot at ε_{yy}=0.0474 in Fig.5. The stress-strain curve shows the remarkable blunting of stress increase at ε_{yy}=0.0611. Figure 5 shows the internal change during the blunting-plateau response by the sign of $\det B_{ij}^{\alpha}$ and central symmetry parameter of Atomeye [Li (2003)]. The bcc-bct phase transition occurs from the GB and propagates in the grains, finally the bct phase covers whole the grains. We can find the negative AES atoms at the forward edge of the expanding phase, or at the migrating boundary. The strain hardening at the later stage is caused by this phase transition. The crystal shows the highest peak at ε_{yy}=0.200 where a cleavage cracking emerges at the new boundary seen in the snapshot at ε_{yy}=0.151 in Fig.5. On the other hand, the stable (111) Σ3 tilt GB doesn't show GB migration but cleavage cracking at the first and highest peak of ε_{yy}=0.099, after showing pulse-like pop-up of negative AES atoms at ε_{yy}=0.020 and ε_{yy}=0.090. The negative AES atoms can be seen only at the GB as shown in Fig.6. From the GB energy at the initial equilibrium, the (111) Σ3 tilt GB is considered more stable than the (322) Σ11 one. In fact, the magnitude relation of the stress and strain of the first peak obeys this prediction; however, the later has deformability and shows better ductility than the former.
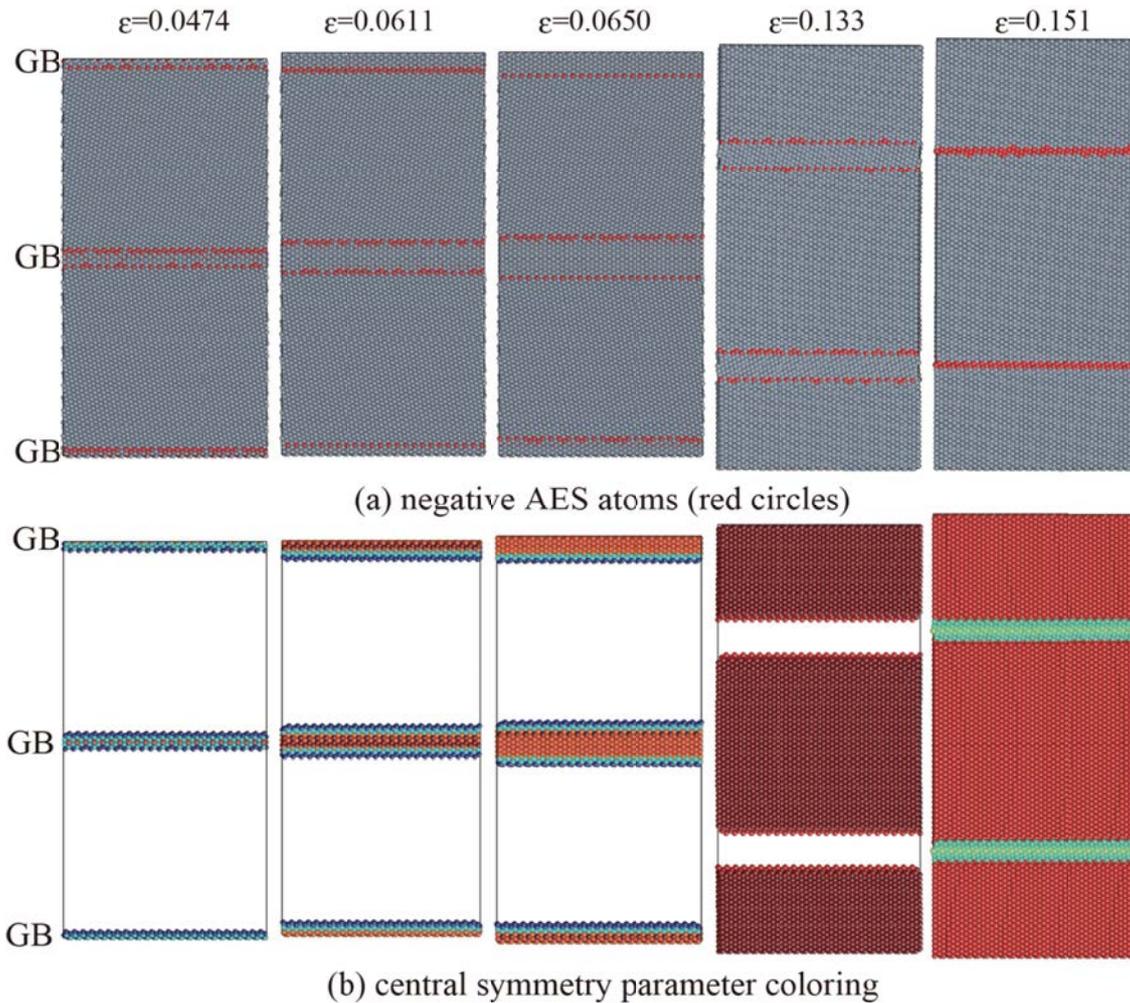
Figure 5. Snapshots of (332) Σ11 symmetric tilt grain boundary under tension

Conclusions

In the present paper, we performed various primitive simulations on bcc Fe to reveal the characteristics of the atomic elastic stiffness (AES) in the perfect bulk, thin plate/wires, tilt/twist grain boundaries (GBs). It is remarkable difference from our previous results of fcc Ni/Al and diamond lattice of Si, that bcc Fe can reach the global instability prior to the local deformation or nucleation of lattice defects in the tensile simulation of bulk, plate and wires. Here, the global instability means all the atoms has negative AES and the system can't follow up the stress increase anymore against external loading. We also investigated the negative AES per unit area against various misorientation angles of tilt/twist grain boundaries, comparing with the GB energy diagram. We also performed tensile simulations on the (322) Σ11 and (111) Σ3 tilt GBs ([110] tilt axis) and revealed that the later actually shows higher elastic limit than the former as predicted by the GB energy before loading; however, the former shows better ductility due to the phase transformation from GBs after the elastic limit, while the later immediately fractured at the GB by cleavage cracking. It is still difficult to connect the stress-strain response to the change in the negative AES, however, it should be noted that we have observed the negative AES at the front edge of the propagating boundary of the phase transformation.

References

Li J. (2003) AtomEye: an Efficient Atomistic Configuration Viewer, *Modelling Simul. Mater. Sci. Eng.* Vo.11, pp.173-177.

Nishimura, M., Yashiro, K. and M. Arai (2010) Local Lattice Instability Analysis on Amorphous Metals: Stability Switching between Stable and Unstable, *Journal of Solid Mechanics and Materials Engineering*, Vol.4, No.11, pp.1550-1562.

Wallace, D. C. (1972) *Thermodynamics of Crystals*, Newyork.

Yashiro, K. and Tomita, Y (2001) Local Lattice Instability at a Dislocation Nucleation and Motion, *Journal de Physique* IV, 11, Pr5-3－Pr5-10, EDP Sciences.

Yashiro, K. Nishimura, M. and Tomita, Y. (2006) Deformation Analysis of Amorphous Metals Based on Atomic Elastic Stiffness Coefficients, *Modelling and Simulation in Materials Science and Engineering*, Vol.14, pp.597-605.

Yashiro, K. and M. Fujihara (2012) Molecular dynamics study on atomic elastic stiffness in Si under tension: homogenization by external loading and its limit, *Modelling and Simulation in Materials Science and Engineering*, Vol. 20, No. 4, 045002.
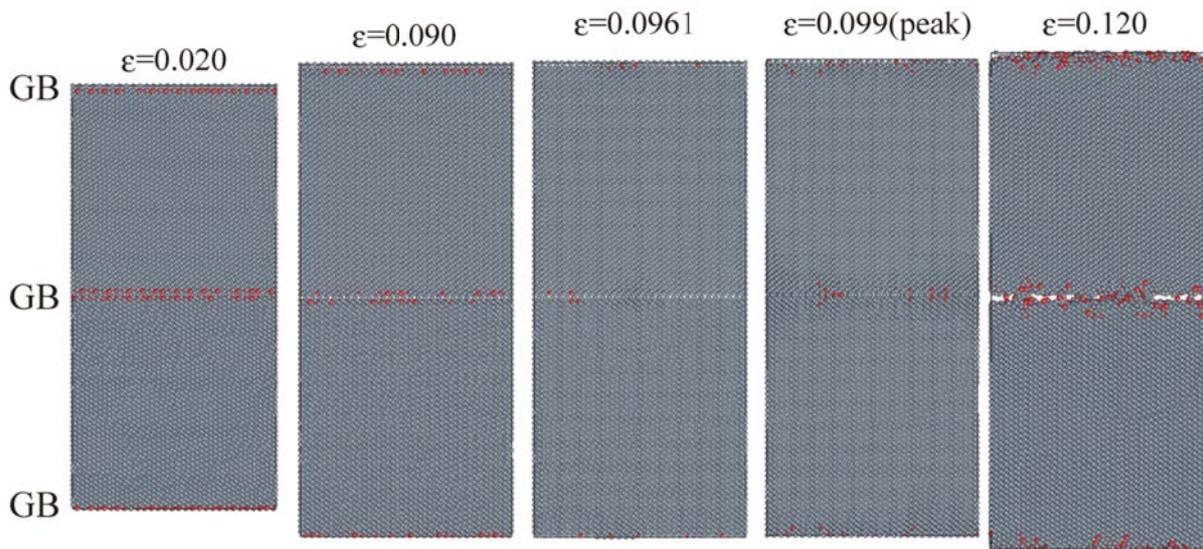
Figure 6. Snapshots of (111) Σ3 symmetric tilt grain boundary under tension. Red circles indicate negative AES atoms.

Two computational approaches for the simulation of fluid problems in rotating spherical shells

*F. Garcia[1,2], E. Dormy[1], J. Sánchez[2] and M. Net[2]

[1]MAG, LRA, Département de Physique, Ecole Normale Supérieure, 24 rue Lhomond, 75005 Paris, France.
[2]Departament de Física Aplicada, Universitat Politècnica de Catalunya, Jordi Girona Salgado s/n. Campus Nord. Mòdul B4, 08034 Barcelona, Spain.

*Corresponding author: ferran@fa.upc.edu

Abstract

Many geophysical and astrophysical phenomena such as magnetic fields generation, or the differential rotation observed in the atmospheres of the major planets are studied by means of numerical simulations of the Navier-Stokes equations in rotating spherical shells. Two different computational codes, spatially discretized using spherical harmonics in the angular variables, are presented. The first code, PARODY, solves the magneto-hydrodynamic anelastic convective equations with finite a difference discretization in the radial direction. This allows the parallelization on distributed memory computers to run massive numerical simulations of second order in time. It is mainly designed to perform direct numerical simulations. The second code, SPHO, solves the fully spectral Boussinesq convective equations, and its variationals, parallelized on shared memory architectures and it uses optimized linear algebra libraries. High-order time integration methods are implemented to allow the use of dynamical systems tools for the study of complex dynamics.

Keywords: Hydrodynamics, Spherical shells, Parallelism, Direct Numerical Simulation, Dynamical Systems

Introduction

The Due to the increase of computing power in the last decades, many geophysical and astrophysical phenomena, such as magnetic fields generation, or the differential rotation observed in the atmosphere of the major planets, are studied by means of three-dimensional numerical simulations of the magneto-hydrodynamic or thermal convection equations in rotating spherical geometries. The introductory sections of [Dormy et al. 2004; Net et al. 2008], among others, provide good reviews of the state of the art on this subject. The difficulties related to the experimental studies, such as the radial gravity which can only be reproduced by means of either an electrostatic radial field or approximated by the centrifugal force, enhance the importance of the numerical approach in these fields. However, non-stationary tridimensional waves arise at the onset of convection due to the boundary curvature, and thus finding a solution requires very high resolutions. Frequently, as in [Pino et al. 2000], and [Plaut and Busse 2005], a two-dimensional annular geometry is used to approximate the real problem. For this reason the development and improvement of the numerical techniques provides the basis for such research.

Several numerical codes to simulate these type of problems were developed independently by different research groups and benchmarked in [Christensen et al. 2001]. A common feature of these codes is that the velocity and magnetic fields are expressed in terms of poloidal and toroidal scalar potentials following the formulation of [Chandrasekhar 1981]. For the spatial discretization of the equations on the sphere, many of these codes use pseudo-spectral methods based on spherical harmonics basis functions in the angular variables, which provide highly accurate solutions with relatively few grid points [Canuto et al. 1988]. These methods are based on transformations from the spectral to the physical space [Orszag 1970]. The calculation of the quadratic terms, appearing in the truncated equations, is performed in the physical space. The main differences between the

codes arise in the discretization along the radial direction, in the implementation of the boundary conditions and in the time-stepping procedures. There exist however other approaches such as that of [Kageyama and Sato 1995] that use finite differences or that of [Matsui and Okuda 2004] that use a finite-element-method in all directions.

Most of the current tridimensional studies consist of direct numerical simulations of periodic, quasi-periodic, and even turbulent flows to study the variation of the time-averaged physical properties in the parameter space and to obtain scaling laws [Christensen and Aubert 2006; Oruba and Dormy 2014]. These numerical simulations are performed with second order time integration semi-implicit schemes which only treat the diffusive terms implicitly. For a deeper understanding of the origin of the laminar flows and their dependence on parameters, pseudoarclength continuation methods [Sánchez et al. 2004; Sánchez et al. 2010], and the linear stability analysis of the time dependent solutions [Net et al. 2008; Garcia et al. 2008] have been successfully applied thanks to the use of high-order time integration methods which provide accurate enough solutions. On the other hand, high-order time integration can also be useful for evolving turbulent flows efficiently [Garcia et al. 2014a].

In this paper two different computational parallel codes, spatially discretized using spherical harmonics in the angular variables, are presented and their applicability for studying geophysical and astrophysical problems is discussed. Also, their parallel performance on the high performance computing center MesoPSL (http://www.mesopsl.fr) is analyzed and possible improvements of the codes are suggested.

The first code, PARODY, solves the magneto-hydrodynamic anelastic convective equations, although in this paper we only comment the Boussinesq implementation, with a finite difference discretization in the radial direction. This allows the parallelization on distributed memory computers to perform massive numerical simulations of second order in time. It is mainly designed to perform direct numerical simulations and it has been widely used by many researchers, see for instance [Dormy et al. 1998; Raynaud and Dormy 2013; Schrinner et al. 2012; Schrinner et al. 2014].

The second code, SPHO, solves the fully spectral Boussinesq convective equations, and its variationals, parallelized on shared memory architectures and it uses optimized linear algebra libraries. High-order time integration methods [Garcia et al. 2014a; Garcia et al. 2010; Garcia et al. 2014b] are implemented to allow the use of dynamical systems tools, such as that of [Sánchez et al. 2004; Sánchez et al. 2010; Sánchez and Net 2013], for the study of complex dynamics.

The model and the equations

The thermal convection and magnetic field generation of a spherical electrically conducting fluid shell differentially heated, rotating about an axis of symmetry with constant angular velocity $\mathbf{\Omega} = \Omega \mathbf{k}$, and subject to radial gravity $\mathbf{g} = -\gamma \mathbf{r}$, where γ is a constant, and $\mathbf{r}$ the position vector, is implemented in the code PARODY. The mass, momentum, energy and induction equations are written by using an usual formulation and non-dimensional units (see [Christensen et al. 2001; Dormy 1997; Dormy et al. 1998; Schrinner et al. 2012] for details). The units are the gap width, $d = r_o - r_i$ for the distance, ΔT (the difference of temperature between the innner and outer boundaries) for the temperature, d^2/ν for the time, and $(\rho \mu \eta \Omega)^{1/2}$ for the magnetic field, ν being the kinematic viscosity, μ the magnetic permeability, η the magnetic diffusivity and r_i and r_o the inner and outer radii, respectively. With these units the equations governing the dynamics of the fluid in the rotating frame of reference are

$$(\partial_t \mathbf{v} + (\mathbf{v} \cdot \nabla)\mathbf{v} - \nabla^2 \mathbf{v})E = -2\mathbf{\Omega} \times \mathbf{v} - \nabla p + (\mathbf{r}/r_o)RaT + P_m^{-1}(\nabla \times \mathbf{B}) \times \mathbf{B}, \qquad (1)$$

$$\nabla \cdot \mathbf{v} = 0, \qquad (2)$$

$$\partial_t T + \mathbf{v} \cdot \nabla T = Pr^{-1}\nabla^2 T, \qquad (3)$$

$$\partial_t \boldsymbol{B} = \nabla \times (\boldsymbol{v} \times \boldsymbol{B}) + P_m^{-1} \nabla^2 \boldsymbol{B}, \tag{4}$$

$$\nabla \cdot \boldsymbol{B} = 0. \tag{5}$$

The non-dimensional parameters are the modified Rayleigh number Ra, the Prandtl number Pr, the magnetic Prandtl number P_m, the Ekman number E, and the radius ratio χ, They are defined by

$$Ra = \frac{g_o \alpha \Delta T d}{\nu \Omega}, \quad E = \frac{\nu}{\Omega d^2}, \quad Pr = \frac{\nu}{\kappa}, \quad P_m = \frac{\nu}{\eta}, \quad \chi = \frac{r_i}{r_o},$$

where α is the thermal expansion coeficient, κ is the thermal diffusivity, and g_o the gravity at the outer radius.

The boundary conditions for the velocity field can be either no slip or stress free at both boundaries, or mixed boundary conditions with no slip at the inner and stress free at the outer sphere. For the magnetic field, a conducting or insulating inner core can be imposed [Dormy 1997; Schrinner et al. 2012], although only the insulating case will be considered in this paper. The temperature is fixed at both boundaries.

The solenoidal velocity field, $\boldsymbol{v}$, is expressed in terms of toroidal, u_t, and poloidal, u_p, scalar potentials $\boldsymbol{v} = \nabla \times (u_t \boldsymbol{r}) + \nabla \times \nabla \times (u_p \boldsymbol{r})$. With the same expression for the magnetic field and by applying the operators $\boldsymbol{r} \cdot \nabla \times$ and $\boldsymbol{r} \cdot \nabla \times \nabla \times$ to the Navier-Stokes equation (Eq. (1)), and $\boldsymbol{r} \cdot$ and $\boldsymbol{r} \cdot \nabla \times$ to the induction equation (Eq. (4)), the equations for the potentials can be deduced. Finally, the functions $X = (u_t, u_p, b_t, b_p, T)$ are expanded in spherical harmonic series up to degree L in the angular variables, namely

$$X(t, r, \theta, \phi) = \sum_{m=-L}^{L} \sum_{l=|m|}^{L} X_l^m(r, t) Y_l^m(\theta, \phi) \tag{6}$$

with $X_l^{-m} = \bar{X}_l^m$, and $[u_t]_0^0 = [u_p]_0^0 = [b_t]_0^0 = [b_p]_0^0 = 0$, to uniquely determine the four scalar potentials, and $Y_l^m(\theta, \phi) = P_l^m(\cos(\theta)) e^{im\phi}$, P_l^m being the normalized associated Legendre functions of degree l and order m. Since $X_l^{-m} = \bar{X}_l^m$, only the $m \geq 0$ amplitudes are retained. With the latter expansion, the equations can be written in terms of their complex coefficients $X_l^m = X_l^m(t, r)$ which are functions of time and radius. The coefficients of the nonlinear terms of Eqs. (1-5) are obtained following [Dormy 1997].

A similar model is implemented in the code SPHO without the induction equation (Eq. (4)). The energy equation (Eq. (3)) is written in terms of the temperature perturbation $\Theta = T - T_c$ from the conductive state $\boldsymbol{v} = 0, T_c = T_c(r)$. The unit for the temperature is $\nu^2 / \gamma \alpha d^4$. The main difference between the codes arise in the radial discretization of the amplitudes $X_l^m(t, r)$, in the time-stepping techniques, and in the parallel strategy used to solve the equations. All these issues are addressed in the following section.

Parallel implementation

The PARODY code

Finite differences are used on a non-uniform mesh of N_r+1 points, stretched near the boundaries to cope with thin Ekman-Hartmann layers. Although finite differences are local and less accurate with respect to other discretizations such as global collocation methods, they are suitable for a parallel implementation on distributed computers in the way we now describe. The radial grid is partitioned among the processors, p_i, $i=1,\ldots,N_p$, each one having all the spherical harmonic amplitudes at $r_{d_i},\ldots,r_{d_i+n_i}$ consecutive n_i+1 radial points. The radial derivative operators are of second order except in the case of the poloidal scalar velocity which is of fourth order. If centered finite differences are used, to apply the derivative operators each processor p_i has to communicate all the amplitudes at r_{d_i} with p_{i-1}, and all the amplitudes at $r_{d_i+n_i}$ with p_{i+1}. In the case of the poloidal scalar velocity the amplitudes at r_{d_i+1} and at $r_{d_i+n_i-1}$ must also be send to processors p_{i-1} and p_{i+1}, respectively. This parallelization is suitable because the evaluation of the nonlinear terms is the most demanding task and it is performed separately by each processor with the only need of communication for two vectors.

Once the original equations Eqs. (1-5) are discretized a large system of $N=2\left(L^2+2L\right)N_r+\left(3L^2+6L+1\right)\left(N_r-1\right)$ ordinary differential equations must be advanced in time. For time-integration, semi-implicit methods are used, namely, only the diffusive terms are treated implicitly with a Crank-Nicholson scheme, and the rest of the terms which include the non-linear and the Coriolis terms are treated explicitly with an Adams-Bashforth method. Thus the linear systems of equations to be solved at every step can be separated into spherical harmonic components, which can be solved independently, so that only a set of small linear systems must be solved at each time step. These linear systems are pentadiagonal in the case of the poloidal velocity and tridiagonal for the other scalars. More specifically, the pentadiagonal matrix comes from the radial discretization of $\left(\partial_t-\Delta\right)\Delta$, while the tridiagonal matrices come from the radial discretization of $\partial_t-\beta\Delta$, where $\beta=1$ in the case of the toroidal velocity potential, $\beta=1/Pr$ in the case of the temperature equation, and $\beta=1/P_m$ for the equations of the magnetic field potentials.

The linear systems in PARODY are usually [Dormy et al. 1998] solved with the parallel implementation of the LU factorization described in [Lakshmivarahan and Sudarshan 1990]. The main drawback of this solver is that it becomes sequential when decreasing the number of radial points of each processor and increasing the number of processors significantly. In the current parallel LU implementation a minimum of 4 radial points are needed for each processor. An implementation of a parallel Krylov iterative solver [Barrett et al. 1994; Saad 1996] could improve the solution of the linear systems. More precisely the IBiCGStab (Improved Stabilized version of BiConjugate Gradient Squared) method is an alternative form to BiCGStab which only involves a single global reduction operation instead of the usual 3 (or 4) [Yang and Brent 2002]. This solver allows to assign only one radial point at each processor. Although this method is highly parallelizable because it only makes use of matrix products, its performance (number of iterations) depends strongly on the condition number of the matrix, which in our case is mainly influenced by the number of radial points and the time step used in the time integration. Thus several tests, with different N_r and time steps corresponding to different physical regimes, must to be performed to compare the performance of both solvers. Preliminary results addressing this issue will be shown later.

The SPHO code

In contrast to PARODY, this code employs a collocation method on a Gauss-Lobatto mesh of $N_r + 1$ points ($N_r - 1$ being the number of inner points). With this global discretization the radial grid can not be partitioned into several processors of a distributed memory cluster for an efficient parallelization. Thus the parallelization of the code is performed in the angular variables assuming shared memory architectures to avoid communications. The linear discretized operators of the equations for the spherical harmonics amplitudes are decoupled with respect to the order *m*. The same occurs for the Legendre transforms needed for the computation of the nonlinear terms. Then, the triangular grid $\{ X_l^m, \ m = 0, ..., L, \ l = m, ..., L \}$ is partitioned among the processors by assigning a set of amplitudes with consecutive order $m_{d_i}, ..., m_{d_i + n_i}$ at each processor. In this case, the number of orders, $n_i + 1$ assigned to each processor increases as m_{d_i} increase, to maintain a similar number of amplitudes X_l^m. Finally, the fast Fourier transforms and the computations in the physical space needed for evaluating the nonlinear terms are also parallelized by evenly partitioning the colatitude physical grid among the processors.

Once the thermal convection equations have been discretized a large system of ordinary differential equations of size $N = (3L^2 + 6L + 1)(N_r - 1)$ must be integrated in time. Notice the smaller number of equations with respect to the PARODY code. In SPHO the induction equation is not considered. If N_v variational equations are integrated the size of the systems becomes $N_v N + N$.

Two classes of high order (up to five) time integration methods are implemented in SPHO. The first class of methods are the implicit-explicit (or fully implicit) backward differentiation formulas (IMEX-BDF) methods [Garcia et al. 2010; Garcia et al. 2014b]. The IMEX methods treat the nonlinear terms explicitly in order to avoid solving nonlinear equations at each time step. The Coriolis term is treated either semi-implicitly or fully implicitly, giving rise to different algorithms. The use of *matrix-free* Krylov methods (GMRES in our case) for the linear systems facilitates the implementation of a suitable order and time stepsize control. In contrast to PARODY, the matrices of linear systems to be solved in SPHO have dense blocks of dimension $O(N_r)$ (see [Garcia et al. 2010] for details on the structure of these matrices). A second alternative implementation for the time stepping is the so called exponential Rosenbrock methods proposed in [Hochbruck et al. 1998]. A wide range of numerical simulations has shown that such exponential methods are more accurate by at least one order of magnitude than the equivalent order IMEX scheme [Garcia et al. 2014a]. This is especially true when they are employed with large time steps and at small Ekman number.

Performance of the codes in MesoPSL

In this section we investigate the performance of PARODY and SPHO codes on the high performance computer MesoPSL, which consist of an array of 92 nodes with 16 cores and 64 Gb of memory ram each one. More precisely, each node is a bi-processor with 8-cores Intel E5-2670 at 2,60 Ghz and the nodes are interconnected with infiniband QDR.

Parody

Three different dynamo test cases, corresponding to different physical regimes with the same geometry $(\chi = 0.35)$ and Prandtl number $(Pr = 1)$, have been considered for studying the behavior of the iterative solver. The first test case, T_1, corresponds to a laminar dynamo with $P_m = 5$ at relatively high $E = 10^{-3}$

and weakly supercritical modified Rayleigh number $Ra=100$. This is the benchmark case 1 of [Christensen et al. 2001]. The radial resolution is $N_r=160$ and the spherical harmonics truncation parameter is $L=64$. The time integration is performed with a time step of $\Delta t=10^{-4}$. The second case, T_2, corresponds to a chaotic dynamo with $P_m=0.5$ at $E=10^{-4}$ and $Ra=700$. The radial resolution is $N_r=256$, the spherical harmonics truncation parameter is $L=80$ and the time step is $\Delta t=10^{-6}$. Finally in the third case, T_3, the complexity of the dynamo with $P_m=0.25$ is increased because of the lower $E=3\times10^{-5}$ and higher $Ra=2\times10^3$. The radial resolution is $N_r=320$, the spherical harmonics truncation parameter is $L=112$, and the time step is $\Delta t=3\times10^{-7}$.

As commented previously, the iterative history of the IBiCGStab solver depends strongly on the condition number of the matrices A_2 and A_1 coming from the discretization of $(\partial_t-\Delta)\Delta$ and of $\partial_t-\beta\Delta$, respectively. These matrices depend on the time step, but also on the degree l of the spherical harmonic amplitudes. The condition number of both matrices decreases with increasing the degree l, thus we have only computed the condition numbers of the case $l=1$. For an easier implementation of the iterative solver we solve all the linear systems for all X_l^m as a single linear system, i.e, we perform the same number of iterations for each l. Then, as the condition number decreases with l, the residuals for the amplitudes decrease with increasing l. Tables (1) and (2), show the condition numbers of the matrices A_1 and A_2 and their preconditioners, respectively. When decreasing the time step, the matrix A_1 becomes close to a multiple of the identity, and the matrix A_2 always has a larger condition number than A_1. For the latter we have used a diagonal preconditioner to improve the convergence (see table (1)) while minimizing the number of communications. For the former we have used a little bit more complicated preconditioner that we explain below. In both cases left preconditioning is better than right preconditioning.

Table 1. Condition number dependence on the radial resolution and the type of preconditioner with $\Delta t=10^{-4}$. $M_i A_1$, mean left preconditioner where M_i is the Jacobi preconditioner with i iterations. For $i=1$ is the diagonal preconditioner.

N_r	40	80	160	250	350	500
A_1	5.6	25.3	112.8	295.7	595.3	1260.8
$M_1 A_1$	3.6	13.0	50.4	122.9	239.3	489.1
$M_2 A_1$	1.5	3.8	13.1	31.2	60.3	122.8
$M_3 A_1$	1.5	4.4	16.8	41.0	79.8	163.0
$M_4 A_1$	1.1	2.2	6.8	15.9	30.4	61.6

Table 2. Condition number dependence on the radial resolution and the type of preconditioner. $M_i A_2$, mean left preconditioner where $M_i = P_i Q_i$ is the preconditioner. P_i and Q_i are the matrices corresponding to i Jacobi iterations with matrices D and A_1, respectively.

	N_r	40	80	160	250	350	500
	A_2	5.3×10^4	1.2×10^6	2.3×10^7	1.6×10^8	6.3×10^8	2.8×10^9
	$M_1 A_2$	4.3×10^2	4.1×10^3	5.5×10^4	3.2×10^5	1.2×10^6	5.0×10^6
$\Delta t = 10^{-4}$	$M_2 A_2$	8.8×10^1	5.5×10^2	6.2×10^3	3.4×10^4	1.3×10^5	5.3×10^5
	$M_3 A_2$	8.3×10^1	6.0×10^2	6.5×10^3	3.6×10^4	1.4×10^5	5.6×10^5
	$M_4 A_2$	6.4×10^1	3.4×10^2	2.8×10^3	1.5×10^4	5.4×10^4	2.2×10^5

	N_r	40	80	160	250	350	500
	A_2	8.3×10^3	4.5×10^4	2.2×10^5	6.8×10^5	1.7×10^6	5.1×10^6
	$M_1 A_2$	2.3×10^2	9.6×10^2	4.0×10^3	1.1×10^4	2.3×10^4	5.5×10^4
$\Delta t = 10^{-7}$	$M_2 A_2$	5.9×10^1	2.4×10^2	9.8×10^2	2.4×10^3	4.9×10^3	1.1×10^4
	$M_3 A_2$	7.7×10^1	3.2×10^2	1.3×10^3	3.2×10^3	6.3×10^3	1.3×10^4
	$M_4 A_2$	2.9×10^1	1.2×10^2	5.0×10^2	1.3×10^3	2.6×10^3	6.0×10^3

Consider the matrix D coming from the discretization of the laplacian Δ with the appropriate boundary conditions. The preconditioning matrix for A_2 is $M_i = P_i Q_i$ where P_i and Q_i are the matrices corresponding to i Jacobi iterations with matrices D and A_1, respectively (see [Barrett et al. 1994; Saad 1996] for further details on preconditioning techniques). In all the cases we have set $i=2$ which reduces significantly the condition number and for which the preconditioning operation only requires one additional communication. See Table 2 for the dependence of the condition numbers on the type of preconditioning, radial resolution and time step.

In figure (1) the run time for performing one time step when using the LU and IBiCGStab solvers is plotted versus the number of MPI tasks for each of the cases considered. In all the cases the tolerance for the IBiCGStab is set in a way that the mean physical properties (such as volume averaged kinetic energy densities or the Nusselt number) differ by less than 3% with respect that obtained with the LU solver when starting the integration from an initial condition as in [Christensen et al. 2001]. The solution at the previous time instant has been chosen as initial seed for starting the iterations. In the

case T_1 (Fig. 1(a)) due to the relatively large time step $\Delta t = 10^{-4}$ the
matrices are ill conditioned and the IBiCGStab solver requires at least 100 iterations for obtaining a
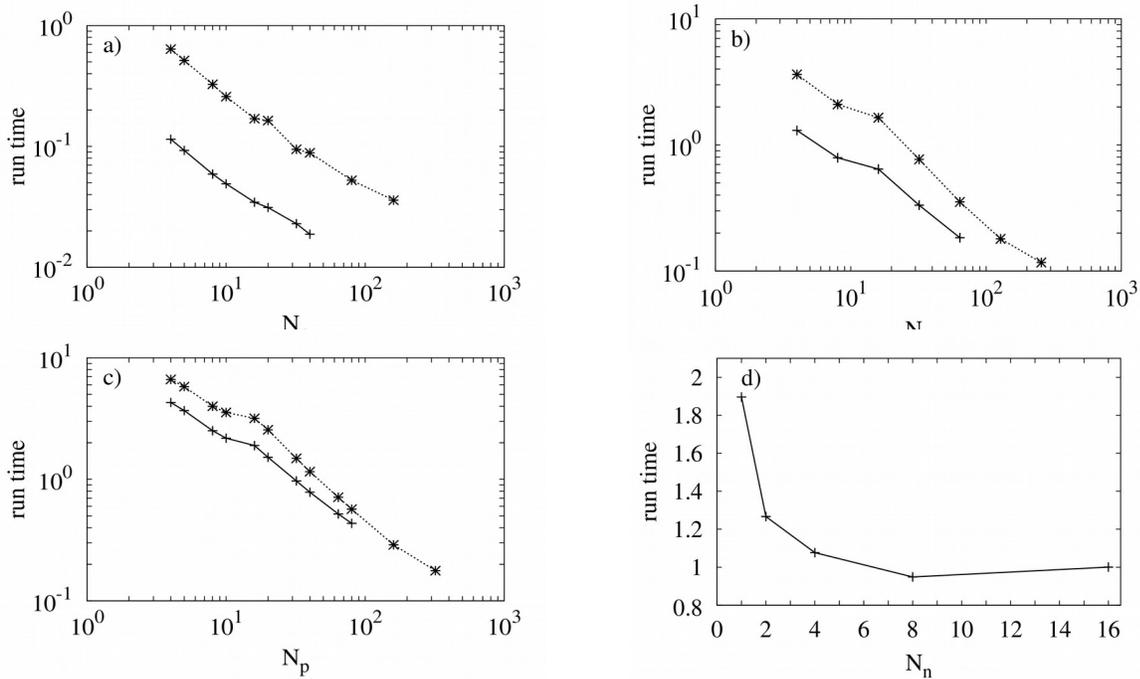
Figure 1. (a) Run time for advancing one time step plotted versus the number of MPI tasks for the test case T_1. (b) and (c) are as (a) but for the test cases T_2 and T_3, respectively. (d) Same as (a) but plotted versus the number of nodes when using 16 MPI task for the test case T_3. The symbols and types of lines indicate: iterative solver (*, dotted line) and direct solver (+, solid line).

residual of order 10^{-6} when solving the linear systems with the matrix A_2. With this number of iterations the iterative solver requires considerably much more computing time than the direct one. For the case T_2 (Fig. 1(b)), the number of iterations is about 50 and thus the difference between the LU and IBiCGStab curves decreases. Finally, for the case T_3 (Fig. 1(c)) only 20 iterations are needed to achieve a residual of order 10^{-4} which has been found enough for obtaining good time-averaged values.

Notice in the slopes of the curves of Figs. 1(b,c) that the IBiCGStab solver has slightly better scalability when using a larger number of processors. In this figure a degradation of the scalability is also evident when using 16 processors because of the architecture of the computer (each node has 16 cores and there is thus competition for memory access). To address this issue in figure 1(d) the run time is plotted versus the number of nodes when using 16 MPI task for the test case T_3. It is clear that is better not to use all the cores of each node to avoid memory access competition, in this way, the computing time can be halved.

SPHO

In this section we describe the performance of the code SPHO parallelized using OpenMP directives and optimized by using basic linear algebra public libraries (GOTO [Goto and Geijn 2008] and ATLAS [Whaley et al. 2000]) and the FFTW3 library for the fast Fourier transforms [Frigo and Johnson 2005]. A test for the integration of the variational equations [Hirsch et al. 2004] is also performed.

In figure 2(a) the run time for advancing one time step obtained with one core divided by the run time obtained by p cores $\left(p \leq 16\right)$ is plotted versus the number of equations for several representative resolutions which are shown in Table 3. The run time for advancing one time step with a fixed time step integration method is basically that for computing the nonlinear terms and for solving the linear systems which are solved by an LU method. Because a direct solver is used, the physical regime plays no role and the performance depends only on the discretization mesh.
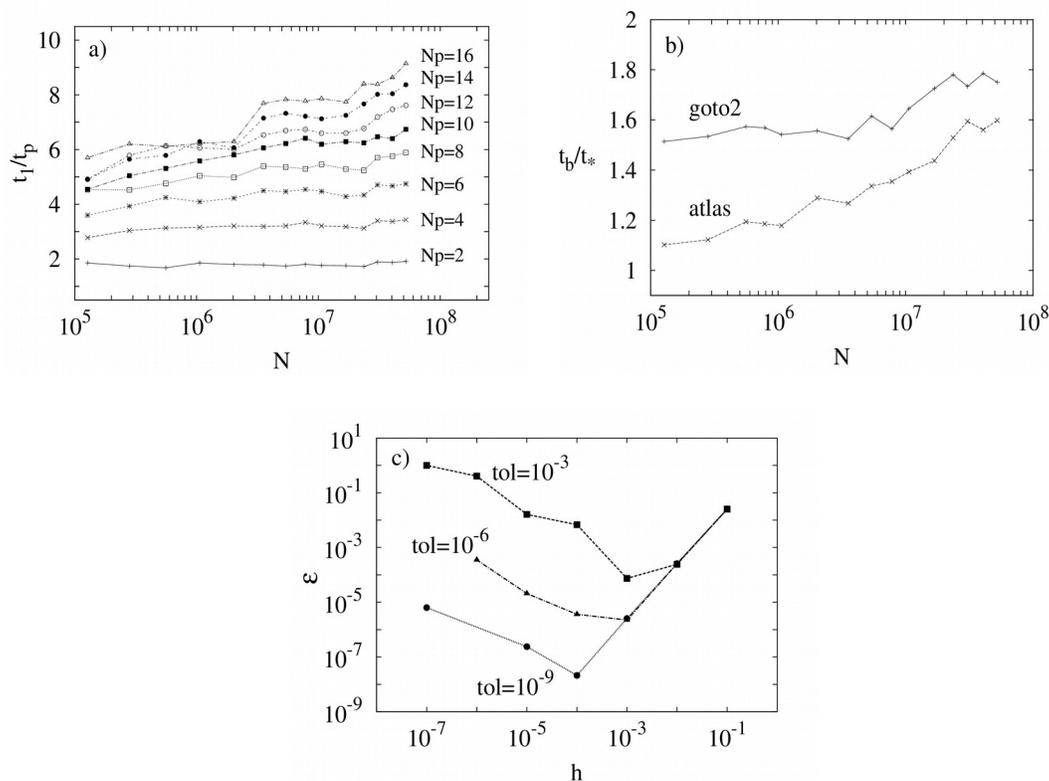
Figure 2. (a) The ratios t_1/t_p, where t_p means the run time obtained with p processors, plotted versus the number of equations. (b) Sequential run time for advancing one time step obtained with the basic BLAS library divided by the sequential run time obtained with the ATLAS and GOTO optimized libraries plotted versus the number of equations. (c) Test for the variational equations: the relative error, ε, plotted versus the centered finite difference approximation step h for three tolerances (labeled on the curves) of the VSVO time integration code.

Table 3. Radial resolution, N_r, and spherical truncation parameter, L, used in figures 2(a,b).

N_r	24	32	38	50	60	72	80	88	94	106	120	130	150	170
L	42	54	70	84	106	128	150	172	194	230	256	280	300	320

When the number of equations is relatively small (up to 2×10^6) the performance degrades when using more than 8 cores because of the access memory competition, however, as the number of equations is increased there is more computational work and the competition for the memory decreases, increasing the performance. We obtain speed ups $S_p=9.1$ for $p=16$ for the high resolution mesh $N_r=170$ and $L=320$, more specifically we obtain $S_p=1.53\,p^{0.64}$. Notice that the slope of the p-curves of Fig. 2(a) increases with increasing p.

As commented before, when using a collocation method to radially discretized the equations, all the radial operators of the original equations are substituted by dense matrices. When the evaluation of an operator is required all similar computations are grouped to call efficient implementations of the matrix-matrix product subroutine DGEMM of BLAS. Also the Legendre transforms needed for the evaluation of the nonlinear terms are implemented with this subroutine. In figure 2(b) the sequential run time for advancing one time step obtained with the basic BLAS library divided by the sequential run time obtained with the ATLAS and GOTO optimized libraries is plotted versus the number of equations for the same resolutions as in Fig. 2(a). Important savings can be obtained with the GOTO library for the larger number of equations where the run time can nearly be halved with respect the basic BLAS library.

Finally, a test for the integration of the variational equations is performed in the following. Assume that the evolution equation for $u\in\mathbb{R}^N$, where u is the vector of all the unknowns of the discretized equations, is

$$\partial_t u=L_0^{-1}\left(Lu+B(u,u)\right), \tag{7}$$

and let $u(t)=\phi_t(u_0)$ be its solution with initial condition $u(0)=u_0$ at $t=0$. In the latter equation, L_0 and L are linear operators including the boundary conditions. The former is invertible, and the latter, for the scheme used, includes the diffusive, the buoyancy, and the Coriolis terms. The operator B, which will be treated explicitly in the IMEX-BDF formulae, will always contain only the nonlinear terms. The variational equations along $u(t)$ are

$$\partial_t u=L_0^{-1}\left(Lu+B(u,u)\right), \tag{8}$$

$$\partial_t v=L_0^{-1}\left(Lv+\frac{1}{2}\left(B(u+v,u+v)-B(u-v,u-v)\right)\right) \tag{9}$$

with $w(t)=(u(t),v(t))\in\mathbb{R}^{2N}$ the solution with initial condition $w(0)=(u_0,v_0)$. The property $D\phi_t(u_0)v_0=v(t,v_0)$ allows us to validate the numerical integration of Eqs. (8-9):

$$v\left(t, v_0\right) = D\, \phi_t\left(u_0\right) v_0 \approx \frac{\phi_t\left(u_0 + h v_0\right) - \phi_t\left(u_0 - h v_0\right)}{2h} = \tilde{v}\left(t, v_0\right). \tag{10}$$

Algorithm

1. Initialise $u_0, v_0 \in \mathbb{R}^N$, the final time at which the errors will be computed, $t > 0$, and the step for the centered formula $h > 0$.
2. Integrate t time units the variational equations Eqs. (8-9) with initial condition $\left(u_0, v_0\right)$.
3. Integrate t time units the original system Eq. (7) with initial conditions $u_0 + h v_0$ and $u_0 - h v_0$ to obtain $\phi_t\left(u_0 + h v_0\right)$ and $\phi_t\left(u_0 - h v_0\right)$, respectively.
4. Compute $\tilde{v}\left(t, v_0\right) = \left(\phi_t\left(u_0 + h v_0\right) - \phi_t\left(u_0 - h v_0\right)\right)/2h$.
5. Check the error

$$\frac{\left\|\tilde{v}\left(t, v_0\right) - v\left(t, v_0\right)\right\|}{\left\|v\left(t, v_0\right)\right\|} \tag{11}$$

Notice that $\varepsilon = \varepsilon\left(t, u_0, v_0, h, tol\right)$, where $tol = \varepsilon^a = \varepsilon^r$ is the tolerance for the Q-implicit VSVO time integration code fully described in [Garcia et al. 2010] (ε^a is the absolute and ε^r the relative error tolerance).

To check the time integration of the variational equations we will consider a case in which the Ekman number is $E = 10^{-4}$, the Prandtl number is $Pr = 0.1$ and the radius ratio is $\chi = 0.35$. More precisely, a modulated travelling wave with azimuthal wave number $m_d = 6$ which is stable at the weakly supercritical Rayleigh number $Ra_e = 2.59929964 \times 10^5$ $\left(Ra_e = \left(\gamma \alpha \Delta T d^4\right)/\left(\kappa \nu\right)\right)$ is considered. This is a quasiperiodic resonant orbit which has two frequencies $f_1 = 60.21680$ and $f_2 = 26.75897$. They satisfy the relation $\left(4 f_1 - 9 f_2\right)/f_2 = O\left(tol\right)$, where tol is the tolerance of the time integration method used to obtain the initial condition u_0.

The initial conditions of Eqs. (8-9) are $v_0 = u_0$ where u_0 is the initial condition of the quasiperiodic orbit and the final time of the time integration is $t \approx 1/f_2$. We compute $\tilde{v}\left(t, v_0\right)$ for several values of the finite difference step h and we integrate Eqs. (8-9) with several time integration tolerances tol. The results are shown in Fig. 2(c), where the relative error ε of Eq. (11) is plotted versus the finite difference tolerance h for three different tolerances $tol = 10^{-3}, 10^{-6}, 10^{-9}$ of the VSVO time integration code. In this figure the error due to the time integration and that due to the truncation can be identified. The latter is exibited for $h > 10^{-2}$ where the curve has an slope 2. The error due to time integration appears for $h < 10^{-2}$.

Discussion

Two different approaches for solving hydrodynamical problems in rotating spherical shells are studied in this paper. In the first approach a finite differences radial discretization is used to allow the parallelization with MPI directives by partitioning the shell in the radial direction into different processors. This is suitable because several types of architectures can be used to run the code. The implementation of the improved version of the BiCGStab Krylov solver could improve the efficiency

of the code in certain physical regimes, which need very small time steps for their integration in time. With this iterative solver a larger number of processors can be used to minimize the computing time for obtaining time-averaged physical properties of chaotic and turbulent dynamo models.

In the second approach the parallelization is performed by partitioning the triangular mesh of spherical harmonics and by using OpenMP directives. The code can only be executed on shared memory architectures. The implementation of the code is performed in such a way to rely on the use of matrix-matrix products with the DGEMM subroutine of the BLAS library. In this case the code is fully spectral, integrates the variational equations, and the time integration schemes are of high order to obtain high accurate solutions which are needed when using dynamical systems tools for a deep study of the physical system.

In certain architectures, such as that of MesoPSL (a cluster of nodes with several cores each one) it is better not to use all the cores to avoid competition for a memory access. Notice that the approach followed in SPHO can be also performed in PARODY by assigning one MPI task at each node and using the cores of it to parallelize the computations on the spherical harmonics mesh using OpenMP directives. The systematic use of the DGEMM subroutine can also improve the code.

Possible slight improvements of the SPHO code with MPI directives will consist on separating independent computations on different nodes. For instance one node could compute the velocity field and another node the vorticity field, which are both needed for the evaluation of the nonlinear terms.

Acknowledgments

References

Barrett, R. and Berry, M. and Chan, T. F. and Demmel, J. and Donato, J. M. and Dongarra, J. and Eijkhout, V. and Pozo, R. and Romine, C. and Van Der Vorst, H. (1994) Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods, *SIAM*.

Canuto, C. and Hussaini, M. Y. and Quarteroni, A. and Zang, T. A. (1988) Spectral Methods in Fluid Dynamics, *Springer*.

Chandrasekhar, S. (1981) Hydrodynamic And Hydromagnetic Stability, *Dover, New York.*

Christensen, U. and Aubert, J. (2006) Scaling properties of convection-driven dynamos in rotating spherical shells and application to planetary magnetic fields, *Geophys. J. Int.* **166**, 97–114.

Christensen, U.R. and Aubert, J. and Cardin, P. and Dormy, E. and Gibbons, S. and Glatzmaier, G.A. and Grote, E.and Honkura, Y. and Jones, C. and Kono, M. and Matsushima, M. and Sakuraba, A. and Takahashi, F. and Tilgner, A. and Wicht, J. and Zhang, K. (2001) A numerical dynamo benchmark, *Phys. Earth Planet. Inter.* **128**, 25–34.

Dormy, E. (1997) Modelisation numerique de la dynamo terrestre, *Ph.D. Thesis, Institut de physique du globe de Paris.*

Dormy, E. and Cardin, P. and Jault, D. (1998) MHD flow in a slightly differentially rotating spherical shell, with conducting inner core, in a dipolar magnetic field, *Earth and Planetary Science Letters* **160**, 15–30.

Dormy, E. and Soward, A. M. and Jones, C. A. and Jault, D. and Cardin, P. (2004) The onset of thermal convection in rotating spherical shells, *J. Fluid Mech.* **501**, 43–70.

Frigo, Matteo and Johnson, Steven G. (2005) The Design and Implementation of FFTW3, *Proceedings of the IEEE* **93**, 216–

231.

Garcia, F. and Bonaventura, L. and Net, M. And Sánchez, J. (2014a) Exponential versus IMEX high-order time integrators for thermal convection in rotating spherical shells, *J. Comput. Phys.* **264**, 41–54.

Garcia, F. and Net, M. and García-Archilla, B. and Sánchez, J. (2010) A comparison of high-order time integrators for thermal convection in rotating spherical shells, *J. Comput. Phys.* **229**, 7997–8010.

Garcia, F. and Net, M. and Sánchez, J. (2014b) A comparison of high-order time integrators for highly supercritical thermal convection in rotating spherical shells, *Proc. of International Conference on Spectral and High   Order Methods for Partial Differential   Equations-ICOSAHOM 2013*, *Lecture Notes in Computational Science and Engineering* **95.**

Garcia, F. and Sánchez, J. and Net, M. (2008) Antisymmetric polar modes of thermal convection in rotating   spherical fluid shells at high Taylor numbers, *Phys. Rev. Lett.* **101**, 194501(1–4).

Goto, Kazushige and van de Geijn, Robert A. (2008) Anatomy of high-performance matrix multiplication, *ACM Trans. Math. Softw.* **34**, 1–25.

Hirsch, M. W. and Smale, S. and Devaney, R. L. (2004) Differential Equations, Dynamical Systems, and Itroducction to Chaos , *Pure and applied mathematics series, Elsevier.*.

Hochbruck, M. and Lubich, C. and Selhofer, H. (1998) Exponential integrators for large systems of differential equations, *SIAM J. Sci. Comput.* **19**, 1552–1574.

Kageyama, A. and Sato, T. (1995) The complexity simulation group, computer simulation of a magnetohydrodynamic dynamo, *Phys. Plasma* **2**, 1421–1431.

Lakshmivarahan, S. and Sudarshan, K. D. (1990) Analysis and design of parallel algorithms: Arithmetic and matrix problems, *Series in supercomputing and parallel processing,   McGraw-Hill.*

Matsui, H. and Okuda, H. (2004) MHD Dynamo simulation using the GeoFEM platform: Comparison with a spectral method, *Pure appl. geophys.* **161**, 2199–2212.

Net, M. and Garcia, F. and Sánchez, J. (2008) On the onset of low-Prandtl-number convection in rotating spherical shells: non-slip boundary conditions, *J. Fluid Mech.* **601**, 317–337.

S. A. Orszag (1970) Transform method for calculation of vector-coupled sums: Application to the spectral form of the vorticity equation, *J. Atmos. Sci.* **27**, 890–895.

Oruba, L. and Dormy, E. (2014) Predictive scaling laws for spherical rotating dynamos, *Geophys. J. Int.*, Accepted.

Pino, D. and Mercader, I. and Net, M. (2000) Thermal and inertial modes of convection in a rapidly rotating   annulus, *Phys. Rev. E* **61**, 1507–1517.

Plaut, E. and Busse, F. H. (2005) Multicellular convection in rotating annuli, *J. Fluid Mech.* **528**, 119–133.

Raynaud, R. and Dormy, E. (2013) Intermittency in spherical Couette dynamos, *Phys. Rev. E* **87**, 1–6.

Saad, Y. (1996) Iterative methods for sparse linear systems, *PWS pub. Company, New York.*

Sánchez, J. and Net, M. (2013) A parallel algorithm for the computation of invariant tori in large-scale dissipative systems, *Physica D* **252**, 22–33.

Sánchez, J. and Net, M. and García-Archilla, B. and Simó, C. (2004) Newton-Krylov continuation of periodic orbits for Navier-Stokes flows, *J. Comput. Phys.* **201**, 13–33.

Sánchez, J. and Net, M. and Simó, C. (2010) Computation of invariant tori by Newton-Krylov methods in large-scale dissipative systems, *Physica D* **239**, 123–133.

Schrinner, M. and Petitdemange, L and Dormy, E. (2012) Dipole collapse and dynamo waves in global direct numerical simulations, *Astrophys. J.* **752**, 1–12.

Schrinner, M and Petitdemange, L. and Raynaud, R. and Dormy, E. (2014) Topology and field strength in spherical anelastic dynamo simulations, *Astronomy & Astrophysics* **564**, 1–13.

Whaley, R. C. and Petitet, A. and Dongarra, J.. (2000) Automated Empirical Optimization of Software and the {ATLAS} Project, *Netlib, Lapack working notes* , UT-CS-00-448.

Yang, L.T. and Brent, R. (2002) The Improved BiCGStab Method for Large and Sparse Unsymmetric Linear Systems on Parallel Distributed Memory Architectures, *Proceedings of the Fifth International Conference on Algorithms and Architectures for Parallel Processing, IEEE.*

Investigating the influence of bending in the structural behavior of tensegrity modules using dynamic relaxation

L. Rhode-Barbarigos[1], *S. Adriaenssens[1]

[1]Form Finding Lab, Dept. of Civil and Environmental Engineering, Princeton University, Princeton NJ, 08544, USA.

*Corresponding author: sadriaen@princeton.edu

Abstract

Tensegrity modules are spatial structures composed of tensile and compression components. Tensile and compression components are assembled together in a self-equilibrated state that provides stability and stiffness to the structure. Modules can be combined to create lightweight structures with good structural efficiency. Furthermore, tensegrity structures are good candidates for adaptive and deployable systems having thus applications in various scientific and engineering fields. Research into tensegrity structures has resulted in reliable techniques for their form-finding and analysis. Although bending is not considered in these techniques, tensegrity structures often sustain bending in their elements due to dead load and imperfections. Therefore, this paper investigates the effect of bending in a tensegrity "simplex" module. Dynamic relaxation is used to analyze the module with strut and strut-beam elements. The study reveals that bending increases stresses in elements and therefore should not be neglected.

Keywords: Tensegrity, Structural Behavior, Bending Elements, Dynamic Relaxation

Introduction

Tensegrity structures are spatial reticulated structures composed of cables and struts in a self-equilibrated pre-stressed state that provides stability and stiffness to the structure. The concept exists for almost 60 years now and has received significant interest in disciplines such as structural engineering [Motro (2005)], aerospace engineering [Skelton and de Oliveira (2009)] and biology [Ingber et al. (2014)]. In biology, tensegrity is used to explain cell mechanics [Ingber (2003)] while in aerospace engineering and structural engineering it is used to design strong yet lightweight modular structures [Skelton and de Oliveira (2009); Adriaenssens and Barnes (2001); Rhode-Barbarigos et al. (2010)]. Tensegrity is also attractive for adaptive applications as actuators and structural elements can be combined [Skelton and de Oliveira (2009); Rhode-Barbarigos et al. (2012a)]. Therefore, tensegrity systems have also been proposed for deployable structures [Sultan and Skelton (2003); Rhode-Barbarigos et al. (2012b)] and robots [Paul et al. (2006); SunSpiral et al. (2013)].

Research into tensegrity systems has resulted in reliable techniques for their form-finding and analysis [Tibert and Pellegrino (2003); Masic et al. (2005)]. In these techniques, compressive elements are modeled as struts with no bending as a pure compression state is desired. However, in reality elements in tensegrity structures are strut-beam elements sustaining bending due to dead load and imperfections such as initial curvature or eccentricity in their joints. Therefore, this paper focuses on the effect of bending in a tensegrity "simplex" module. Dynamic relaxation is used to analyze the module numerically and study the effect of considering strut-beam elements with initial curvature in its structural behavior.

"Simplex" module topology and geometry

The tensegrity structure studied in this paper is the "simplex" module. The "simplex" module is the basic spatial tensegrity system [Motro (2006)]. It is composed of 3 struts and 9 cables (Figure 1). The module topology is given in the appendix. The module has a single state of self-stress and a single infinitesimal mechanism which involves the translation and rotation of the upper triangle [Motro (2005)].
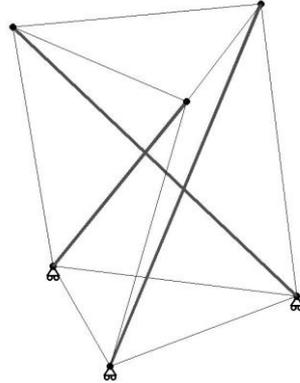
Figure 1. The tensegrity "simplex" module

In this study, strut elements and strut-beam elements are steel hollow tubes with a modulus of elasticity of 210GPa, while tensile elements are stainless steel with a modulus of elasticity of 120GPa. Strut and strut-beam elements have a length of approximately 1.4m, a diameter of 76.1mm and a thickness of 4mm. Cables have a cross-sectional area of 0.2826mm^2 and a tensile strength of 31.8kN. Cable lengths depend on their topology with horizontal cables having a length of 0.866m and vertical cables having a length of 1.032m. Finally, vertical displacements on all nodes on the basis of the module are restrained.

Dynamic Relaxation

In this study, dynamic relaxation is employed for the static analysis of the "simplex" module. Dynamic relaxation is an explicit numerical form-finding and analysis method of tensile structures [Barnes (1999); Adriaenssens and Barnes (2001); Bel Hadh Ali et al. (2011)] that avoids stiffness-matrix calculations [Brew and Brotton (1971)]. Therefore, it is suitable for the analysis of nonlinear structures such as tensegrity modules.

In dynamic relaxation, a structure is modeled as a mesh of elements connected with nodes. A mass is assigned to every node. Loading is also applied to the nodes, while pre-stress is applied through the definition of an initial element length. The method explores the fact that the static solution for a structure subject to loading can be seen as the equilibrium state of a series of damped vibrations. Consequently, the governing equation is:

$$F_{ext} - F_{int} = M\dot{v} + Dv$$

(1)

where F_{ext} and F_{int} are the external and internal forces at each node respectively, M corresponds to the nodal mass and D corresponds to damping. However, mass M and damping D are fictitious parameters optimized for the stability and convergence of the method [Belytschko and Hugues (1983)]. $\dot{v}$ and v are the acceleration and the velocity at each node respectively. In this study, kinetic damping is employed [Cundall (1976)]. Therefore, kinetic energy is monitored and when a peak in kinetic energy is detected, the velocity is reset to zero, the geometry is updated and convergence is

checked. Expressing the acceleration in a finite difference form gives the velocity and the updated geometry for each node:

$$v^{t+\Delta t/2} = v^{t-\Delta t/2} + \frac{F_{ext}-F_{int}}{M}\Delta t \; ; \tag{2}$$

$$x^{t+\Delta t} = x^t + v^{t+\Delta t/2}\Delta t \tag{3}$$

where $v^{t+\Delta t/2}$ and $v^{t-\Delta t/2}$ are the nodal velocities at times $t+\Delta t/2$ and $t-\Delta t/2$ respectively. $x^{t+\Delta t}$ is the nodal position at time $t+\Delta t$ and Δt is the time step applied. The new geometry obtained allows updating the internal forces F_{int} and thus starting over. Convergence is obtained when the term F_{ext} - F_{int} is sufficiently small (equilibrium).

To encounter the effect of bending in the "simplex" module, the bending-element formulation by Adriaenssens and Barnes (2001) is employed for strut-beam elements. Strut-beam elements are thus decomposed on a series of links and bending moments are estimated based on a finite difference modeling of a continuous beam. Bending moments are decomposed into shear forces that are added to the existing nodal forces and convergence is checked according to the general calculation scheme (Eq. 1). The formulation allows thus the method to maintain its computational advantages.

Structural analysis

Dynamic relaxation is used to analyze the structural response of the "simplex" module (stresses in the elements) under self-stress as well as under the combination of self-stress with vertical loading. In order to investigate the effect of bending in the "simplex" module, compressive elements are first modeled using struts (purely axially loaded elements) and then with strut-beam elements. Moreover, since tensegrity structures have pinned connections an initial curvature is also given to the strut-beam elements ($1/(10*l_{strut})$) to initiate bending action in them.

Figures 2 shows the stresses in the cables and struts of the "simplex" module for different self-stress levels (5%, 10% and 15% of the tensile strength of the cables) with strut elements (left) and strut-beam elements (right). The analysis shows that in both configurations cables are the most load bearing elements of the system. Furthermore, when strut-beam elements are employed stresses in cables reduce (up to 40%) while stresses in strut-beam elements increase (up to 44%).
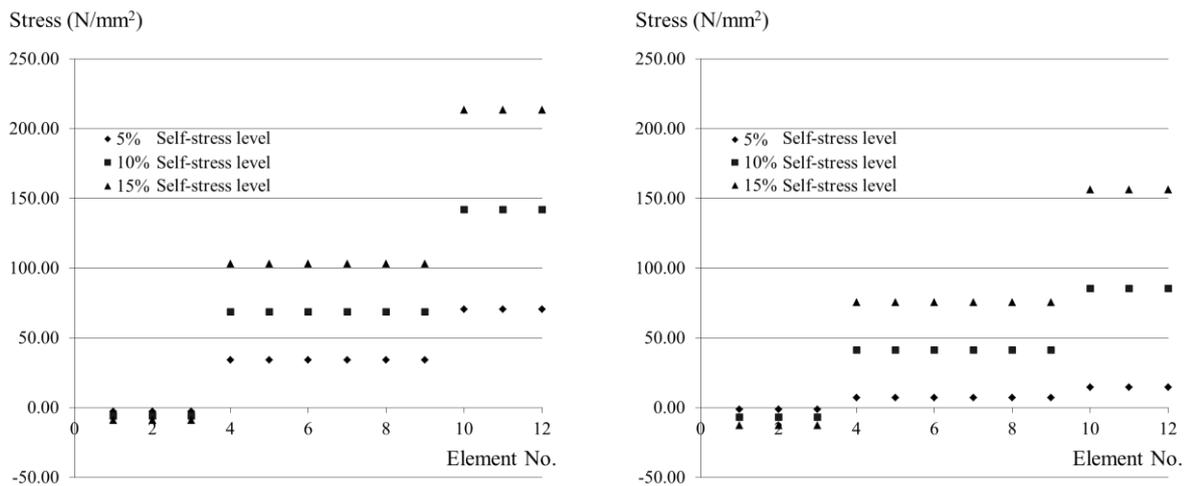
Figure 2. Stresses in the elements of the "simplex" module for different self-stress levels (5%, 10% and 15% of the tensile strength of the cables) with strut elements (left) and strut-beam elements (right).

Figures 3 shows the stresses in the elements of the "simplex" module for different vertical loads applied on the top nodes (10kN, 20kN and 30kN per top node) with strut elements (left) and strut-beam elements (right). A self-stress level of 5% is also applied in the cables to guarantee the stability and stiffness of the module. Similar to the self-stress study, the analysis reveals that cables are the load bearing elements and that when strut-beams are employed, stresses in these elements increase significantly (close to 100%) while stresses in cables remain in approximately the same level (decrease of 5%).
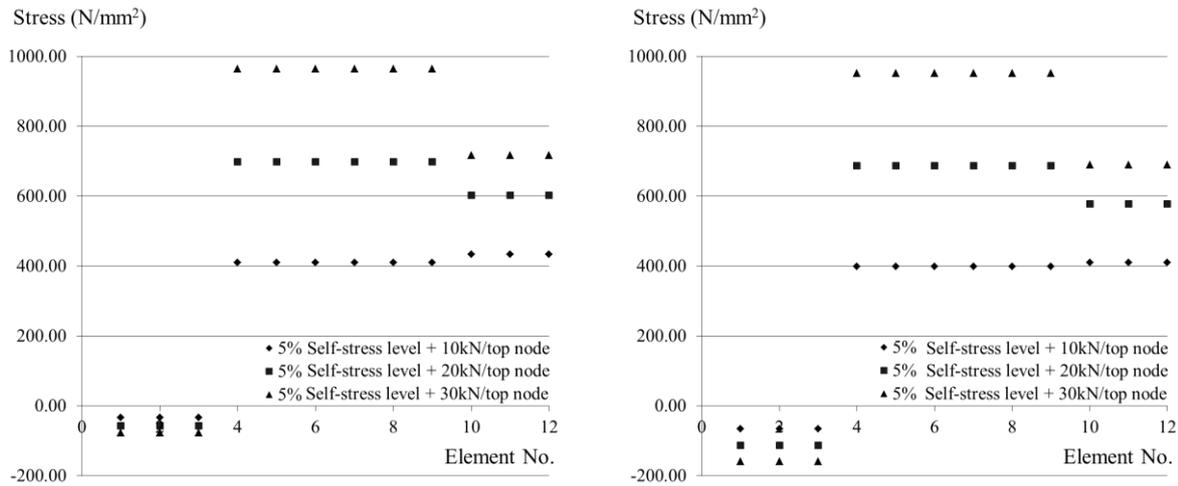
Figure 3. Stresses in the elements of the "simplex" module for different loads applied on the top nodes (0kN, 10kN and 20kN per top node) with struts modelled as axial elements (left) and bending elements (right).

Bending increases stresses in strut-beam elements. Consequently, it can lead to failure at lower loading levels than originally predicted with form-finding and analysis techniques that model compressive elements with struts. Therefore, bending should be taken into account when designing tensegrity systems especially for load-bearing applications.

Discussion

Current design theory holds that bending is undesirable in tensegrity elements. However, by integrating bending in the form-finding process, novel tensegrity structures constructed from flexible yet strong engineering materials that have low Young's modulus and high strength such as Fibre Reinforced Plastics (FRP) could be explored. Applying such materials reduces significantly bending stresses in the strut-beam elements avoiding failure and thus opening the door to the development of a whole new realm of novel tensegrity systems that can sustain large elastic deformations without failure similar to natural systems [Ingber et al. (2014)].

Conclusions

This paper investigates the effect of bending in a tensegrity "simplex" module. Dynamic relaxation is used to analyze a "simplex" module with strut and strut-beam elements. It is found that considering bending increases stresses in the elements which can lead to failure at lower loading levels than predicted with traditional form-finding and analysis techniques. Therefore, it is important to consider bending when designing tensegrity structures. Moreover, integrating bending in the form-finding process could lead to bending-active tensegrity systems and thus novel applications of tensegrity systems.

References

Motro, R. (2005). Tenségrité, Hermes Science.

Skelton, R. E. and M. C. de Oliveira (2009). Tensegrity systems, Springer.

Ingber, D.E., Wang, N., Stamenović, D. (2014). Tensegrity, cellular biophysics, and the mechanics of living systems. Reports on Progress in Physics, 77(4), 046603.

Ingber, D.E. (2003) Tensegrity: I. Cell structure and hierarchical systems biology. J Cell Sci 116: 1157-1173

Adriaenssens S., Barnes M.R. (2001). Tensegrity spline beam and grid shell structures, Eng Struct 23(1), 29-36.

Rhode-Barbarigos L., Bel Hadj Ali N., Motro R., Smith I.F.C. (2010). Designing tensegrity modules for pedestrian bridges. Eng Struct, 32(4): 1158-1167.

Rhode-Barbarigos L., Schulin C., Bel Hadj Ali N., Motro R., Smith I.F.C. (2012a). Mechanism-based approach for the deployment of a tensegrity-ring module. J of Struct Eng-ASCE, 138: 539-548.

Sultan, C. and Skelton R. (2003). Deployment of tensegrity structures. Int J of Solids and Struct 40(18): 4637-4657.

Rhode-Barbarigos L., Bel Hadj Ali N., Motro R., Smith I.F.C. (2012b). Design aspects of a deployable tensegrity-hollow-rope footbridge. Int J of Space Struct, 27(2): 81-96.

Paul C., Valero-Cuevas F.J., Lipson H. (2006). Design and Control of Tensegrity Robots for Locomotion, IEEE Trans. on Robotics, 22 (5), 944–957.

SunSpiral V., Gorospe G., Bruce J., Iscen A., Korbel G., Milam S., Agogino A., Atkinson D. (2013). Tensegrity Based Probes for Planetary Exploration: Entry, Descent and Landing (EDL) and Surface Mobility Analysis. Int J of Planetary Probes.

Tibert, A.G. and Pellegrino S. (2003). "Review of Form-Finding Methods for Tensegrity Structures." Int J of Space Struct 18: 209-223.

Masic, M., Skelton R., et al. (2005). "Algebraic tensegrity form-finding." Int J of Solids and Struct 42(16-17): 4833-4858.

Motro R. (2006) Tensegrity: Structural Systems for the Future. Butterworth-Heinemann, Boston.

Barnes, M.R. (1999). "Form Finding and Analysis of Tension Structures by Dynamic Relaxation." International Journal of Space Structures 14: 89-104.

Bel Hadj Ali N., Rhode-Barbarigos L., Smith I.F.C. (2011). Analysis of clustered tensegrity structures using a modified dynamic relaxation Algorithm. Int. J. of Solids and Structures, 48(5): 637-647.

Brew JS, Brotton DM. (1971). Non-linear structural analysis by dynamic relaxation. Int J Numer Methods Eng. 3(4), 463-83.

T.B. Belytschko, T.J.R. Hugues, (1983) Computational Methods for Transient Analysis: Computational Methods in Mechanics; Mechanics and Mathematical Methods, a Series of Handbooks. North-Holland, Amsterdam-New York.

P.A. Cundall, Numerical Methods in Engineering (Proc. of the EF Conf. on Numerical Methods in Geomechanics), 1, 132-150, (1976).

Appendix

Table 1. Nodal coordinates for the "simplex" module

Node	X [mm]	Y [mm]	Z [mm]
1	500	0	0
2	-250	433	0
3	-250	-433	0
4	-433	250	1000
5	0	-500	1000
6	433	250	1000

Table 2. Nodal connectivity for the "simplex" module

Element	Node	Node
Strut	1	4
Strut	3	6
Strut	2	5
Cable	1	2
Cable	2	3
Cable	3	1
Cable	4	5
Cable	5	6
Cable	6	4
Cable	1	6
Cable	3	5
Cable	2	4

Multiscale seamless domain method

for linear elastic and steady temperature problems

*Y. Suzuki[1]

[1] Department of Mechanical Sciences and Engineering, Tokyo Institute of Technology, Japan.

*Corresponding author: ysuzuki@ginza.mes.titech.ac.jp

Abstract

A new meshfree numerical analysis, termed the seamless domain method (SDM), is applied in a multiscale technology. The SDM requires only coarse-grained points and does not need a stiffness equation, mesh, grid, cell, or element. The SDM is composed of two analyses. The first is a microscopic analysis of local simulation domain to obtain shape functions for interpolation of dependent-variable distributions and influence coefficients for calculation of interaction between the points. These allow an SDM solution to represent a heterogeneous material without homogenization. The second step is a macroscopic analysis of a seamless global domain without mesh or grid. Various numerical examples illustrated that the method worked out both steady temperature fields and linear elastic fields. Every SDM solution using only a few hundreds of points was as accurate as that from conventional finite element analysis using more than 300 thousands of node points.

Keywords: Multiscale, Meshfree, Non-homogeneous media, Linear elasticity, Steady heat conduction

Introduction

Most analytical errors of numerical methods involve spatial (and temporal) discretization of a governing equation. Although the discretization in itself is a source of calculation error, conventional numerical techniques cannot produce an approximate solution without spatial division for simulation fields. Although modeling and analyzing the fields with fine node layout (a small mesh) can greatly reduce the error, a large calculation cost is required. In the conventional discretization, so-called shape functions interpolate dependent-variable distributions from the variable values at neighboring points such as nodes and grid points. The error is strongly associated with the precision of interpolating functions. Polynomial functions are generally used for interpolation but the simple functions cannot express a sufficiently smooth variable distribution. Especially in structural analyses, a coarse analytical model often causes problems of shear locking [Wang D et al. (2006); Li S et al. (2000)] and hourglass deformation [Joldes GR et al. (2008)] in conventional finite element methods (FEMs). In general, a variable profile in each finite-element mesh does not obey governing equations because types of shape functions are determined almost regardless of the type of the equations. Additionally, the dependent-variable gradients are discontinuous on grid lines, interfaces between meshes, and interfaces between control volumes. This is because the variable distribution is interpolated from only variable values at nodes in the mesh. Even multiscale FEMs [Chua J et al. (2008); Ilic S et al. (2009)] and multiscale finite volume methods [Jenny P et al. (2004); Lunati I and Jenny P (2006)] cannot generate a continuous variable gradient throughout the field. The above problems can occur when analyzing either a homogeneous field or a heterogeneous field.

There is another demand for multiscale analysis that can model and analyze heterogeneous materials quickly at low computational cost. Heterogeneous materials, such as alloyed metals, plywood, reinforced concrete, and advanced composite structures have been used in many different fields. Composites are composed of multiple constituents and have different physical or chemical properties from the individual components. Although the composite materials can be optimally designed for different purposes in accordance to the use and application, the mechanical, thermal, and electromagnetic tendencies are quite difficult to estimate without a lot of experiments. High-precision numerical analysis that can precisely predict the properties of the composites without performing any experiments would shorten a product development period. Therefore, multiscale schemes, such as homogenization method [Kaczmarczyk Ł et al. (2010)], multiscale FEM [Chua J et al. (2008); Ilic S et al. (2009)], and multiscale finite volume method [Jenny P et al. (2004); Lunati I and Jenny P (2006)], coupling macroscopic and microscopic models are necessary to achieve both high analytical accuracy and low computational cost. However, as mentioned above, these conventional techniques cannot give an exact solution satisfying a governing equation everywhere.

A new meshfree multiscale analysis is developed here to overcome all the above problems [Suzuki Y et al. (2014)]. The scheme is termed the seamless domain method (SDM) because a global field (the entire structure) is modeled as a "seamless" simulation domain that has no element, mesh, or grid. Although the SDM model has only a small number of coarse-grained points (CPs), the solution can represent a heterogeneous material with microscopic constituents without homogenization. Instead of spatial discretization, a local domain (i.e., a part of the global domain) is analyzed before conducting the global analysis. This enables dependent-variable profile in the SDM model to satisfy the governing equations almost perfectly throughout the global domain. Additionally, both the variable and its gradient are continuous everywhere.

To investigate the practical efficacy of the proposed scheme, two kinds of fundamental numerical examples were solved by the SDM: a steady temperature field (scalar field); static linear-elastic fields (vector field). Heterogeneous fiber composite materials were analyzed and both the analytical precision and computational time were compared with those of ordinal FEM analysis. In addition, occurrence of shear locking is also investigated in a bending-dominated problem. As a result, every SDM solution using only a few hundreds of points was shear-locking-free and as accurate as that from conventional finite element analysis using more than 300 thousands of node points.

Formulation of the seamless domain method (SDM)

This section illustrates theoretical framework of the SDM. Cartesian vectors and matrices are used throughout the paper. The proposed SDM couples microscopic local domain and microscopic global domain.

- Local analysis that generates interpolating functions and influence coefficient matrix
- Global analysis that provides a macroscopic solution (i.e., dependent-variable values of all coarse-grained points (CPs)) using the influence coefficient matrix, and the microscopic solution from the macroscopic solution using the interpolating functions

In the local analysis, a small domain composed of a few units to several tens of units is analyzed to obtain high-precision interpolating functions and influence coefficients for the second steps. This objective can be achieved with conventional numerical analysis by finely discretizing governing equation(s) of the local domain spatially (and temporally). The second analysis is of the entire global domain that is represented by CPs. Dependent variable at every CP in the global domain can be obtained using the influence coefficients. After that, the detailed information on the local-variable distributions in the global domain is calculated from the CPs' variable values using the interpolation function obtained in the first step.

The following subsections will explain the details of each analysis by taking steady temperature problem as an example. Therefore, dependent variable, its gradient, flux, and flow rate are the temperature, temperature gradient, heat flux, and heat flow, respectively.

Global analysis

This subsection illustrates how to conduct global analysis with a composite material as an example global domain. As presented in Fig. 1, the composite is assumed to have a perfectly periodic microstructure, called the unit, which has single carbon fiber in the center of the unit. Ω_G ($\subset \mathbf{R}^d$) is a d-dimensional ($d = 2$ in this example) global domain whose boundary is Γ_G. The global analysis of the entire composite structure is represented by CPs that are shown as black circles in Fig. 1. The CPs endowed with temperature as a dependent variable are placed at four corners of each unit in this example, which can significantly reduce the number of degrees of freedom (DOF) in comparison with an exact FEM domain that models fiber and polymer separately.

Figure 2 presents a partial magnified image of Fig. 1. A red dashed frame encircling CPs 1–8 is defined as a "region of influence", Ω_{R1}, where neighboring CPs interact with each other. At first, temperature distribution in the red area is calculated. Let us consider a case where temperature at CP 0 is unknown and temperatures at all the CPs remaining in the red frame (i.e., CPs 1–8) are known. The m CPs ($m = 8$ in this case) denoted $\mathbf{u}_{(m)}^R$ on the region's boundary Γ_{R1} and the temperature vector on these CPs is denoted

$$\mathbf{u}_{(m)}^R = \left(u_1^R \quad \cdots \quad u_m^R \right)^T \text{ on } \Gamma_{R1}, \tag{1}$$
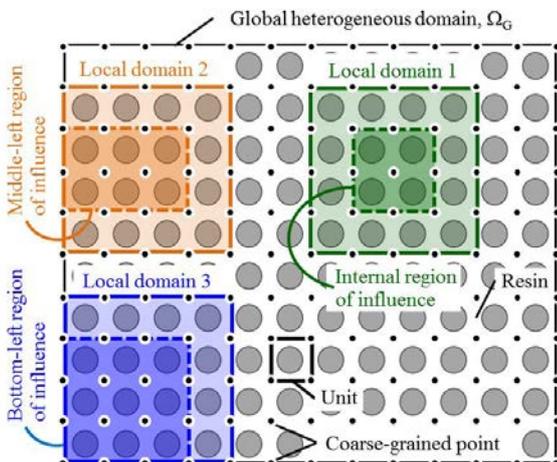
Figure. 1 Heterogeneous global domain (fiber composite) that is assumed to have a perfectly periodic microstructure
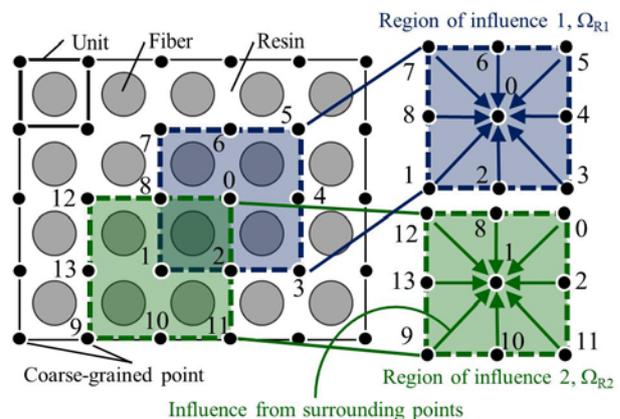
Figure. 2 Regions of influence in partial magnified image of the heterogeneous global domain as shown in Fig. 1

where u_i^R is temperature at CP i. The superscript "T" refers to the transposition of a matrix or vector. Unless heat enters from outside the system, the correct temperature contour in the blueframe can be calculated from the temperature profile on the frame. However, there are infinitely many points and temperature DOFs on the frame, and this estimation approach is thus unrealistic. Therefore, the temperature distribution in the blue frame is estimated referring to temperatures at CPs 1–8.

$$u_{(m)}(\mathbf{x}) = \mathbf{N}_{(m)}(\mathbf{x})\mathbf{u}_{(m)}^R \text{ in } \Omega_{R1}, \tag{2}$$

where $u_{(m)}(\mathbf{x})$ is temperature at position $\mathbf{x} = (x, y)^T \in \Omega_{R1}$ and $\mathbf{N}_{(m)}$ is a special interpolating function matrix derived in local analysis. Details of the local analysis will be illustrated in the next subsection. The number of CPs, m need not necessarily be 8. A sufficient number of reference CPs gives a true temperature distribution. Even if the global domain is not homogeneous, there is no necessity to separately model the microscopic constituents but the structural heterogeneity is exactly taken into account. After obtaining the temperature distribution in the blue frame, $u_{(m)}(\mathbf{x})$, temperature at CP 0 is thus determined by substituting $\mathbf{x} = \mathbf{0}$ into Eq. 2.

$$u_0^R = u_{(m)}(\mathbf{0}) = \mathbf{N}_{(m)}(\mathbf{0})\mathbf{u}_{(m)}^R = \mathbf{a}_{(m)}\mathbf{u}_{(m)}^R \tag{3}$$

where

$$\mathbf{a}_{(m)} = \mathbf{N}_{(m)}(\mathbf{0}) = \begin{pmatrix} a_1 & \cdots & a_m \end{pmatrix} \tag{4}$$

is a influence coefficient matrix. As it is, however, the global domain is spatially discretized by the blue frame and becomes a patchy domain having many seams. There is no guarantee that temperature and its gradient on the seam (i.e., frame) are continuous. To eliminate the seam, we should follow the procedure explained below. Temperature at CP 1 is estimated referring to the remaining CPs in the green dashed frame (i.e., CPs 0, 2, 8–13 in Fig. 2) using $\mathbf{a}_{(m)}$ in Eq. 4. It is of importance that both the blue frame (Ω_{R1}) and green frame (Ω_{R2}) contain CPs 0, 1, 2, 8. These shared CPs works for making the temperature contour in the shared region ($\Omega_{R1} \cap \Omega_{R2}$) interpolated referring to temperature at CPs 1–8 in the blue frame correspond with that interpolated from CPs 0, 2, 8–13 in the green frame. When the shared region has a sufficient number of shared CPs, the two temperature distributions match exactly. As this improves continuity of the temperature and its gradient on the shared frame greatly, the shared frame is no longer a seam. By describing frames centered at each of the CPs, the global domain is filled with shared regions. There is no region or seam belonging to an unshared and isolated frame, and the global domain finally becomes "seamless" and has continuous distributions of temperature and its gradient satisfying the governing equation(s) almost exactly.

Note that the SDM calculates temperature at a CP referring to temperature at surrounding CPs. There is no necessity to formularize and solve a stiffness equation (i.e., nodal equation of equilibrium) that determines the relation between nodal heat flow and the nodal temperature. When n CPs represent a two-dimensional global domain, Ω_G, there is n DOFs in Ω_G. The vector including all the CPs' temperature in Ω_G, $\mathbf{u}_{(n)}^G$, has n components

$$\mathbf{u}_{(n)}^G = \begin{pmatrix} u_1^G & \cdots & u_n^G \end{pmatrix}^T \text{ in } \Omega_G. \tag{5}$$

We can formularize the same number of relational expressions between the temperature of the center CP and those of surrounding CPs like Eq. 3 as that of the components of $\mathbf{u}_{(n)}^{\mathrm{G}}$. Consequently, the temperatures of all the CPs in Ω_{G}, $\mathbf{u}_{(n)}^{\mathrm{G}}$, can be determined uniquely. The $\mathbf{u}_{(n)}^{\mathrm{G}}$ satisfies the following algebraic equations.

$$\mathbf{a}_{(n)}^{\mathrm{G}}\mathbf{u}_{(n)}^{\mathrm{G}} = \mathbf{u}_{(n)}^{\mathrm{G}}, \tag{6}$$

where matrix $\mathbf{a}_{(n)}^{\mathrm{G}}$ is established by assembling all the influence coefficient matrices, $\mathbf{a}_{(m)}$. $\mathbf{a}_{(n)}^{\mathrm{G}}$ is a band matrix whose rows and columns are the same as the number of the CPs, n :

$$\mathbf{a}_{(n)}^{\mathrm{G}} = \begin{bmatrix} a_{1,1}^{\mathrm{G}} & a_{1,2}^{\mathrm{G}} & \cdots & \cdots & a_{1,n}^{\mathrm{G}} \\ a_{2,1}^{\mathrm{G}} & \ddots & \cdots & \cdots & \vdots \\ \vdots & \cdots & a_{i,j}^{\mathrm{G}} & \cdots & \vdots \\ \vdots & \cdots & \cdots & \ddots & \vdots \\ a_{n,1}^{\mathrm{G}} & \cdots & \cdots & \cdots & a_{n,n}^{\mathrm{G}} \end{bmatrix}, \tag{7}$$

where $a_{i,j}^{\mathrm{G}}$ is the component in the i-th row and j-th column of $\mathbf{a}_{(n)}^{\mathrm{G}}$. By imposing boundary conditions into Eq. 6, the solution of $\mathbf{u}_{(n)}^{\mathrm{G}}$ is calculated uniquely. For example, let us consider a Dirichlet boundary problem where a CP's temperature u_i^{G} is given and the other temperatures are unknown. At first, all the unknown temperatures on the right side of Eq. 6 are transposed to the left side:

$$\begin{bmatrix} a_{1,1}^{\mathrm{G}}-1 & a_{1,2}^{\mathrm{G}} & \cdots & \cdots & \cdots & \cdots & a_{1,n}^{\mathrm{G}} \\ \vdots & \ddots & \cdots & \cdots & \cdots & \cdots & \vdots \\ a_{i-1,1}^{\mathrm{G}} & \cdots & a_{i-1,i-1}^{\mathrm{G}}-1 & a_{i-1,i}^{\mathrm{G}} & a_{i-1,i+1}^{\mathrm{G}} & \cdots & a_{i-1,n}^{\mathrm{G}} \\ a_{i,1}^{\mathrm{G}} & \cdots & a_{i,i-1}^{\mathrm{G}} & a_{i,i}^{\mathrm{G}} & a_{i,i+1}^{\mathrm{G}} & \cdots & a_{i,n}^{\mathrm{G}} \\ a_{i+1,1}^{\mathrm{G}} & \cdots & a_{i+1,i-1}^{\mathrm{G}} & a_{i+1,i}^{\mathrm{G}} & a_{i+1,i+1}^{\mathrm{G}}-1 & \cdots & a_{i+1,n}^{\mathrm{G}} \\ \vdots & \cdots & \cdots & \cdots & \cdots & \ddots & \vdots \\ a_{n,1}^{\mathrm{G}} & a_{n,2}^{\mathrm{G}} & \cdots & \cdots & \cdots & a_{n,n-1}^{\mathrm{G}} & a_{n,n}^{\mathrm{G}}-1 \end{bmatrix} \begin{bmatrix} u_1^{\mathrm{G}} \\ \vdots \\ u_{i-1}^{\mathrm{G}} \\ u_i^{\mathrm{G}} \\ u_{i+1}^{\mathrm{G}} \\ \vdots \\ u_n^{\mathrm{G}} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ u_i^{\mathrm{G}} \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \tag{8}$$

Subsequently, 1 and 0 are substituted into $a_{i,i}^{\mathrm{G}}$ and the remaining components in the i-th row, respectively:

$$\begin{bmatrix} a_{1,1}^{\mathrm{G}}-1 & a_{1,2}^{\mathrm{G}} & \cdots & \cdots & \cdots & \cdots & a_{1,n}^{\mathrm{G}} \\ \vdots & \ddots & \cdots & \cdots & \cdots & \cdots & \vdots \\ a_{i-1,1}^{\mathrm{G}} & \cdots & a_{i-1,i-1}^{\mathrm{G}}-1 & a_{i-1,i}^{\mathrm{G}} & a_{i-1,i+1}^{\mathrm{G}} & \cdots & a_{i-1,n}^{\mathrm{G}} \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ a_{i+1,1}^{\mathrm{G}} & \cdots & a_{i+1,i-1}^{\mathrm{G}} & a_{i+1,i}^{\mathrm{G}} & a_{i+1,i+1}^{\mathrm{G}}-1 & \cdots & a_{i+1,n}^{\mathrm{G}} \\ \vdots & \cdots & \cdots & \cdots & \cdots & \ddots & \vdots \\ a_{n,1}^{\mathrm{G}} & a_{n,2}^{\mathrm{G}} & \cdots & \cdots & \cdots & a_{n,n-1}^{\mathrm{G}} & a_{n,n}^{\mathrm{G}}-1 \end{bmatrix} \begin{bmatrix} u_1^{\mathrm{G}} \\ \vdots \\ u_{i-1}^{\mathrm{G}} \\ u_i^{\mathrm{G}} \\ u_{i+1}^{\mathrm{G}} \\ \vdots \\ u_n^{\mathrm{G}} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ u_i^{\mathrm{G}} \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \tag{9}$$

By multiplying both sides of Eq. 9 by the inverse of the matrix on the left side of Eq. 9, we can determine all the components of $\mathbf{u}_{(n)}^{\mathrm{G}}$. After solving the above equation, local-temperature

distributions can be easily obtained by interpolating from $\mathbf{u}_{(n)}^{G}$ using $\mathbf{N}_{(m)}(\mathbf{x})$ in Eq. 2. Although the above illustration shows how to solve a Dirichlet boundary problem, the SDM can also solve Neumann boundary problems and mixed boundary problems. Note that the number of reference CPs in the region of influence, m, is arbitrary in Eqs. 1–4.

Local analysis

Before conducting global analysis, local analysis is done to calculate the interpolating function matrix $\mathbf{N}_{(m)}$ in Eq. 2 and influence coefficient matrix $\mathbf{a}_{(m)}$ in Eq. 4. This objective can be achieved by analyzing a local domain Ω_{L} ($\subset \mathbf{R}^{d}$) with boundary Γ_{L} that is extracted from the global domain, Ω_{G}, as shown in Fig. 3. In this case, Ω_{L} is composed of 16 units. Although the local analysis is done employing the traditional FEM, other conventional numerical techniques can be also used. By dividing Ω_{L} into fine meshes having a sufficient number of nodes, $\mathbf{N}_{(m)}$ and $\mathbf{a}_{(m)}$ for each region of influence Ω_{R} can be obtained.

Let us illustrate how to construct $\mathbf{N}_{(m)}$ and $\mathbf{a}_{(m)}$ for the internal region of influence. As depicted in Fig. 3(a), Ω_{L} includes the internal region, Ω_{R}, composed of four units ($\Omega_{\mathrm{R}} \subset \Omega_{\mathrm{L}}$). The Ω_{R} has m CPs on the edge, Γ_{R}, and another CP (CP 0) in the center. Figure 3(a) illustrates a case where m is equal to 8. In the global analysis, temperature profile in the Ω_{R} is determined referring to the temperature values at the m CPs using $\mathbf{N}_{(m)}$. Encircling Ω_{R} with the outer 12 units is termed the oversampling method [Henning P et al. (2013); Efendiev Y et al. (2013)]. It is known that oversampling reduces boundary effect on Ω_{R}. The dimension of the domain, d, is equal to 2 in this case. The m CPs denoted $\mathbf{u}_{(m)}^{R}$ in Eq. 1 and the m nodes denoted $\mathbf{u}_{(m)}^{L}$ are put on the region's boundary Γ_{R} and the local domain's boundary Γ_{L}, respectively. The temperature vector $\mathbf{u}_{(m)}^{L}$ on those points are denoted

$$\mathbf{u}_{(m)}^{L} = \left(u_1^L \quad \cdots \quad u_m^L \right)^{T} \text{ on } \Gamma_{\mathrm{L}}. \tag{10}$$

As shown in Fig. 3, a Dirichlet boundary is put on Γ_{L} so that the Γ_{L} has given linear temperature profile. The temperature of the center CP 0 and that at an arbitrary point $\mathbf{X}$ ($\in \Omega_{\mathrm{R}}$) are denoted u_0^R,
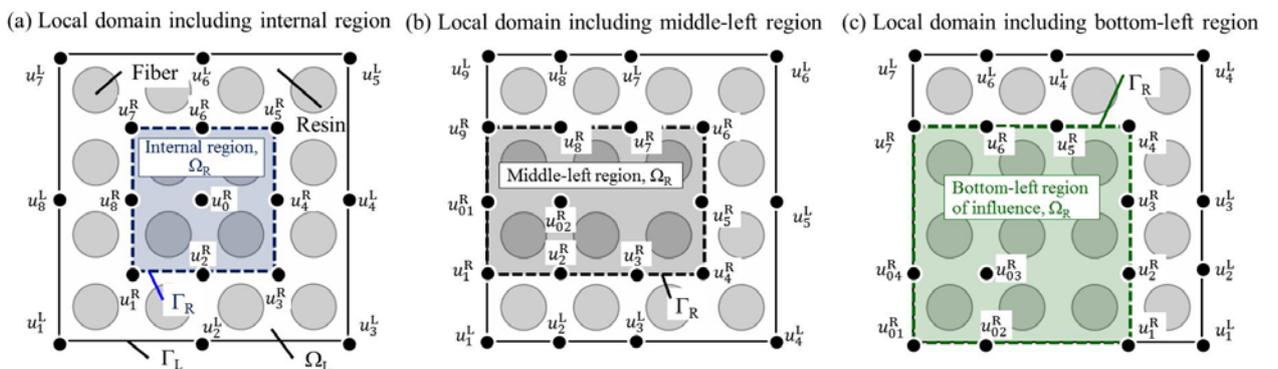
Figure 3. Layout of coarse-grained points (CPs) in local domains for the SDM analysis: (a) shows local domain including the internal region of influence; (b) shows local domain including the middle-left region; (c) shows local domain including the bottom-left region

$u_{(m)}(\mathbf{x})$, respectively. The objective of the local analysis is to calculate the relations between $\mathbf{u}_{(m)}^{\mathrm{L}}$ and the other temperatures u_0^{R} and $u_{(m)}(\mathbf{x})$:

$$
\begin{aligned}
u_0^{\mathrm{R}} &= \mathbf{A}_{(m)}^{\mathrm{R}} \mathbf{u}_{(m)}^{\mathrm{L}} \\
u_{(m)}(\mathbf{x}) &= \mathbf{A}_{(m)}(\mathbf{x}) \mathbf{u}_{(m)}^{\mathrm{L}}
\end{aligned}
\tag{11}
$$

All the entries of $\mathbf{A}_{(m)}^{\mathrm{R}}$ and $\mathbf{A}_{(m)}(\mathbf{x})$ can be derived from only a single analysis of local domain. The $\mathbf{u}_{(m)}^{\mathrm{L}}$ and $\mathbf{u}_{(m)}(\mathbf{x})$ are expressed in the form

$$
\begin{aligned}
\mathbf{u}_{(m)}^{\mathrm{L}} &= \left(\mathbf{A}_{(m)}^{\mathrm{R}}\right)^{-1} \mathbf{u}_{(m)}^{\mathrm{R}} \\
u_{(m)}(\mathbf{x}) &= \mathbf{A}_{(m)}(\mathbf{x})\left(\mathbf{A}_{(m)}^{\mathrm{R}}\right)^{-1} \mathbf{u}_{(m)}^{\mathrm{R}} = \mathbf{N}_{(m)}(\mathbf{x}) \mathbf{u}_{(m)}^{\mathrm{R}}
\end{aligned}
\tag{12}
$$

where

$$
\mathbf{N}_{(m)}(\mathbf{x}) = \mathbf{A}_{(m)}(\mathbf{x})\left(\mathbf{A}_{(m)}^{\mathrm{R}}\right)^{-1},
\tag{13}
$$

is the interpolating function matrix that was stated in the previous subsection. When the Ω_{L} has a sufficient number of nodes, $\mathbf{N}_{(m)}(\mathbf{x})$ provides accurate temperature profile. By substituting $\mathbf{x} = \mathbf{0}$ into $\mathbf{N}_{(m)}(\mathbf{x})$ in Eq. 12, the $\mathbf{a}_{(m)}$ can be calculated like Eqs. 3 and 4. Equation 3 shows that u_0^{R} can be written as a linear combination of $\mathbf{a}_{(m)}$ and $\mathbf{u}_{(m)}^{\mathrm{R}}$. Each component of $\mathbf{a}_{(m)}$ is the influence of a surrounding CP on CP 0 as shown by the arrows in Fig. 2. An increase in the number of CPs m improves the calculation accuracy of u_0^{R}.

The above illustration shows how to make the internal region of influence. Respectively different $\mathbf{N}_{(m)}$ and $\mathbf{a}_{(m)}$ need to be prepared for the other outer regions depicted in Figs. 3(b) and (c) because the outer regions are adjacent to a free edge and receive a boundary effect through one or more aspects of the region. For instance, 2-by-3 units located at the left side of the local domain are defined as the middle-left region of influence. In this region, unknown temperatures at CPs 01 and 02, $u_{01}^{\mathrm{R}}, u_{02}^{\mathrm{R}}$, are calculated referring to temperatures at 9 CPs (CPs 1–9). Another region located at the corner of the local domain has 3-by-3 units as shown in Fig. 3(c). In the region, temperatures at CPs 01–04, $u_{01}^{\mathrm{R}}, u_{02}^{\mathrm{R}}, u_{03}^{\mathrm{R}}, u_{04}^{\mathrm{R}}$, are estimated referring to temperatures at 7 CPs (CPs 1–7). $\mathbf{N}_{(m)}$ and $\mathbf{a}_{(m)}$ of these outer regions can be obtained in a similar way to those of the internal region. Constructing respective $\mathbf{N}_{(m)}$ and $\mathbf{a}_{(m)}$ for each region enables the SDM solution to take the boundary effect into account. In preparing the outer regions, we only have to move the CPs so that the region has contact with the local domain's boundary as shown in Figs. 3(b) and (c). In addition, both the number and arrangement of CPs in the local domain are arbitrary in Eqs. 10–13 and Figs. 1–3. Consequently, all the regions can be constructed from a result of single local analysis. In other words, there is no necessity to conduct FEM analysis of the local domain twice or more [Suzuki Y et al. (2014)].

Numerical examples

Outline

Numerical example problems for testing the practical efficacy of the SDM technique are outlined. To investigate the simulation accuracy and calculation cost, several heterogeneous fields such as steady temperature field and linear elastic fields are analyzed employing the following two numerical techniques:

·   the SDM scheme with quadrangular regions of influence, and

·   the conventional FEM analysis with four-node linear isoparametric meshes.

It is mathematically guaranteed that an FEM solution converges to exactly correct one when the finite-element mesh is infinitesimally fine. The two methods are compared in terms of analytical accuracy and computational cost. In all the examples, local analyses of the SDM are conducted employing a traditional FEM. The target heterogeneous fields are fiber composite materials that have carbon fibers in polymer matrix as depicted in Figs. 4 and 5. The composites, which have a huge number of fibers about 5–15 μm in diameter in the polymer matrix, have been used as a structural material in aircraft and aerospace components because they exhibit superior mechanical properties, higher fatigue strength characteristics, lower mass density, and higher corrosion resistance than conventional metallic materials. A heterogeneous material is assumed to have a perfectly periodic microstructure, termed the unit, which has a carbon fiber as shown in Figs. 4 and 5.

In local analysis of the SDM, a small local domain composed of 4-by-4 units extracted from the entire global field is analyzed to calculate an interpolation function matrix $\mathbf{N}_{(m)}$ in Eq. 2 and influence coefficient matrix $\mathbf{a}_{(m)}$ in Eq. 3 for each region of influence. For the numerical examples, we need prepare one kind of internal region of influence and 8 types of outer regions for the subsequent global analysis. The internal region represents part of the global domain that is located inside the global domain and not adjacent to the global domain's boundary. Figure 3(a) illustrates how to construct $\mathbf{a}_{(m)}$ and $\mathbf{N}_{(m)}$ for the internal region. The 2-by-2 units that are located at the center of the local domain and enclosed in the dashed frame are targeted as the internal region; the area encircled by the bold frame forms a surrounding region to remove the boundary effect on the target. On the other hand, the other 8 types of regions represent outer parts of the global domain. One or two surfaces of the outer regions are adjoined to the global boundary. In calculating $\mathbf{N}_{(m)}$
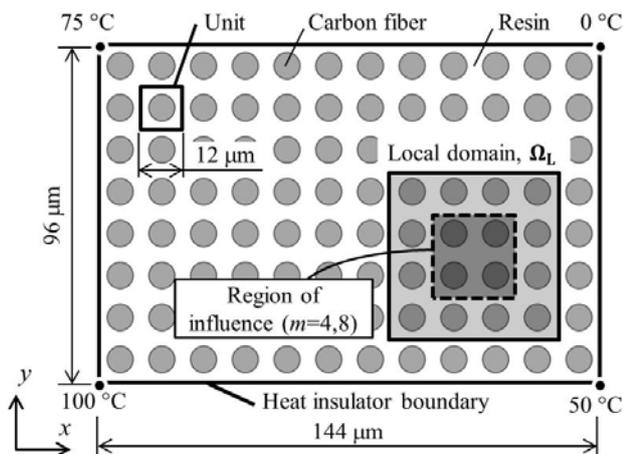
Figure 4. Fiber composite global domain for the steady temperature example problem
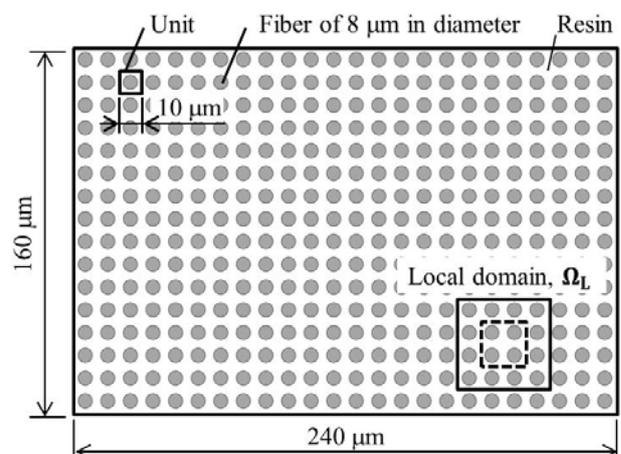
Figure 5. Heterogeneous global domain for the linear-elastic numerical example

and $\mathbf{a}_{(m)}$ for the outer regions, CPs are arranged and the regions of influence are extracted as shown Figs. 3(b) and (c). Once the two regions (i.e., middle-left and bottom-left regions) are constructed, the other six kinds of regions (i.e., middle-right, top-middle, bottom-middle, bottom-right, top-left, and top-right regions) can be easily generated by geometric symmetries. Again, note that all the regions can be prepared from the result of the single local analysis.

After the FEM analysis of the local domain is finished, we move on to the SDM analysis of the global domain. The SDM global field does not require individual fiber and polymer models. Because the heterogeneous material has a periodic unit, the global domain has evenly spaced CPs. After assembling and solving algebraic equations like Eq. 9 to determine dependent-variable values of all CPs, $\mathbf{u}_{(n)}^{\mathrm{G}}$ in Eq. 5, the variable distribution in the global domain is interpolated from $\mathbf{u}_{(n)}^{\mathrm{G}}$ using $\mathbf{N}_{(m)}(\mathbf{x})$ in Eq. 2.

Problem statement for steady-state heat conduction

A steady temperature heterogeneous field as shown in Fig. 4 is analyzed by the SDM and ordinal FEM. The conditions are: two-dimensional; steady-state. The vector of heat flux, matrix of thermal conductivity, and temperature of an arbitrary point $\mathbf{x}$ are denoted $\mathbf{q}(\mathbf{x}) \in \mathbf{R}^2$, $\mathbf{C}(\mathbf{x}) \in \mathbf{R}^{2 \times 2}$, and $u(\mathbf{x}) \in \mathbf{R}$, respectively.

$$\mathbf{q}(\mathbf{x}) = -\mathbf{C}(\mathbf{x})\nabla u(\mathbf{x}), \tag{14}$$

where

$$\mathbf{D}(\mathbf{x}) = \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix}, \tag{15}$$

$u(\mathbf{x})$ then satisfies the governing equation

$$\mathbf{C}(\mathbf{x})\Delta u(\mathbf{x}) = \mathbf{C}(\mathbf{x})\nabla^2 u(\mathbf{x}) = C_{11}\frac{\partial^2 u(\mathbf{x})}{\partial x_1^2} + C_{22}\frac{\partial^2 u(\mathbf{x})}{\partial x_2^2} = 0 \text{ for } {}^{\forall}\mathbf{x} = (x, \quad y)^{\mathrm{T}} \in \Omega, \tag{16}$$

throughout the simulation field. If the domain is isotropic and has the same thermal conductivities in the x and y axes,

$$C_{11} = C_{22} \neq 0,$$

then

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0. \tag{17}$$

As depicted in Fig. 4, Dirichlet boundary conditions are imposed on the target carbon fiber/epoxy resin composite material. The temperatures at top-left, top-right, bottom-left, and bottom-right corners are fixed at 75, 0, 100, and 50 °C, respectively. The outer edge of the global domain except for the four corners is a heat insulation wall. Thermal conductivity coefficients used for the analysis are listed in Table 1.

Table 1. Thermal conductivity values used in analyses of the examples

Material	Thermal conductivity [W·m⁻¹·K⁻¹]		
	C_{11}	C_{22}	$C_{12} = C_{21}$
Carbon fiber	500	500	0
Epoxy resin	0.26	0.26	0

In the SDM analysis, we test two cases where $m = 4$ (Fig. 6(a)) and $m = 8$ (b). m is the number of reference CPs for temperature. The local domain, Ω_L, is divided into fine finite-element meshes and analyzed with Dirichlet boundary conditions of $u_1^L, u_2^L, \cdots u_m^L$ on the m CPs on the boundary of Ω_L and Γ_L, to obtain the interpolating function matrix, $\mathbf{N}_{(m)}(\mathbf{x})$ in Eq. 2, and influence coefficient vector, $\mathbf{a}_{(m)}$ in Eq. 3. Ω_L has 3,649 nodes. The global domain is represented by 59 CPs in the case that $m = 4$ and 117 CPs in the cases that $m = 8$. After constructing and solving algebraic equations like Eq. 9 to determine temperature values of all CPs, $\mathbf{u}_{(n)}^G$ in Eq. 5, the temperature contours in the global domain is interpolated from $\mathbf{u}_{(n)}^G$ using $\mathbf{N}_{(m)}(\mathbf{x})$. Conversely, in the traditional FEM, fiber and polymer are modeled separately and divided into sufficiently fine meshes (more than 20 thousands of node points).

Problem statement for linear elasticity

The static elasticity fields as described in Fig. 7 are analyzed. The conditions are: two-dimensional; static; linear elastic; plane stress. The vectors of stress, engineering strain, and displacement of an arbitrary point $\mathbf{x}$ ($\in \Omega$) are denoted $\boldsymbol{\sigma}(\mathbf{x}) \in \mathbf{R}^3$, $\boldsymbol{\varepsilon}(\mathbf{x}) \in \mathbf{R}^3$, and $\mathbf{u}(\mathbf{x}) \in \mathbf{R}^2$, respectively. The constitutive law for the domain can be written in the form

$$\boldsymbol{\sigma}(\mathbf{x}) = \mathbf{D}(\mathbf{x})\boldsymbol{\varepsilon}(\mathbf{x}), \tag{18}$$

where

$$\boldsymbol{\varepsilon}(\mathbf{x}) = \begin{bmatrix} \varepsilon_x \\ \varepsilon_y \\ \gamma_{xy} \end{bmatrix} = \begin{bmatrix} \partial u / \partial x \\ \partial v / \partial y \\ \partial u / \partial y + \partial v / \partial x \end{bmatrix}, \tag{19}$$

and u, v are displacements in the x and y axes at $\mathbf{x}$, respectively.
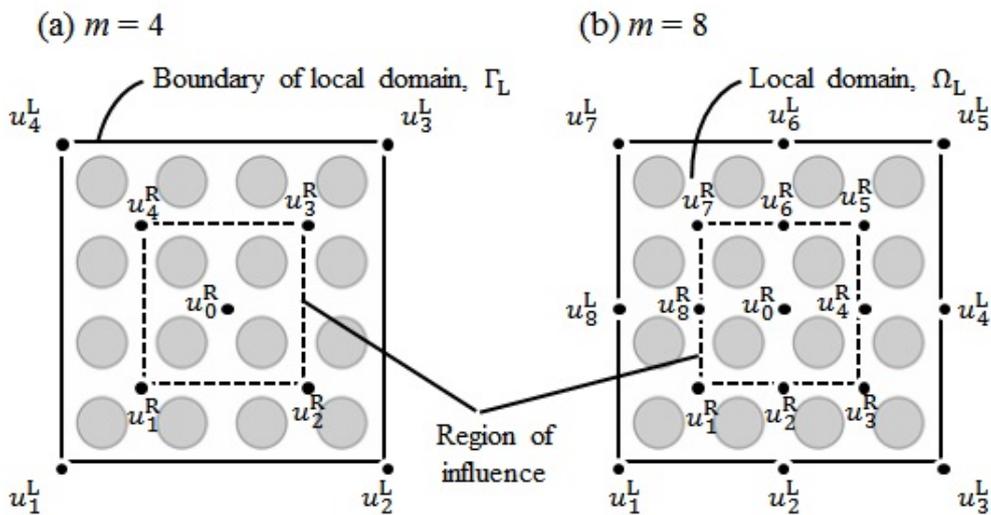
Figure 6. Layout of coarse-grained points (CPs) in local domains for the SDM analysis: (a) shows the case that m = 4; (b) shows the case that m = 8

Figure 7 describes three carbon fiber (T650-35)/ Polyimide resin (PMR-15) composites consisting of 16-by-24 units which have single fiber in the center under: (a) a uniaxial tensile

Table 2. Linear elastic properties of the carbon fiber and polyimide resin models [Rupnowski P et al. (2005)]

Material	Young's modulus in the x	Young's modulus in the y	Poisson ratio, v_{xy}
Carbon fiber (T650-35)	13.8 GPa	13.8 GPa	0.37
Polyimide resin (PMR-15)	4.0 GPa	4.0 GPa	0.29

load; (b) a bending load; (c) a shearing load. Perfect bonding between the fibers and resin matrix is assumed. The elastic properties are listed in Table 2 [Rupnowski P et al. (2005)].

In the example, we test single case where $m = 8$. The local domain, Ω_L, has 14,864 nodes. Global analysis of the SDM is conducted using $\mathbf{N}_{(m)}$ and $\mathbf{a}_{(m)}$ obtained in the local analysis. The three global domains shown in Figs. 7(a)–(c) have the same structure but different boundary conditions are imposed. Each global domain of the SDM is represented by 425 CPs, while more than 300 thousands of nodes are in the FEM model. This is because the SDM global domains do not require individual fiber and polymer models. After constructing and solving algebraic equations like Eq. 9 to determine displacement values of all the CPs, $\mathbf{u}_{(n)}^G$ in Eq. 5, the displacement distribution in the global domain is interpolated from $\mathbf{u}_{(n)}^G$ using $\mathbf{N}_{(m)}(\mathbf{x})$ in Eq. 12.

Results for steady-state heat conduction

Figures 8 present the temperature distribution, u, and temperature gradients, $\partial u/\partial x$ and $\partial u/\partial y$, in the heterogeneous domain. In the traditional FEM, fiber and polymer are modeled separately and divided into sufficiently fine meshes. Therefore, the FEM solutions as depicted in Fig. 8(a) are believed to be exactly correct. Therefore, when the SDM solution (Figs. 8(b) and (c)) accords with the FEM result, calculation error of the SDM is regarded as zero.

A sharp change in temperature from 0 to 100 °C is generated over the global domain. The temperature distribution calculated with the SDM having 59 CPs (Fig. 8(b)) is partially discontinuous. This means that the number of reference CPs, $m = 4$ is insufficient to give the exact solution. The temperature distributions of the SDM consisting of 117 CPs (Fig. 8(c)) almost completely matches that obtained with the conventional FEM with more than 20 thousands of nodes. Therefore, it is believed that $m = 8$ is enough for solving the example problem. The contours of the temperature gradient in Fig. 8(c) show that high resolution of localized steep gradients can be
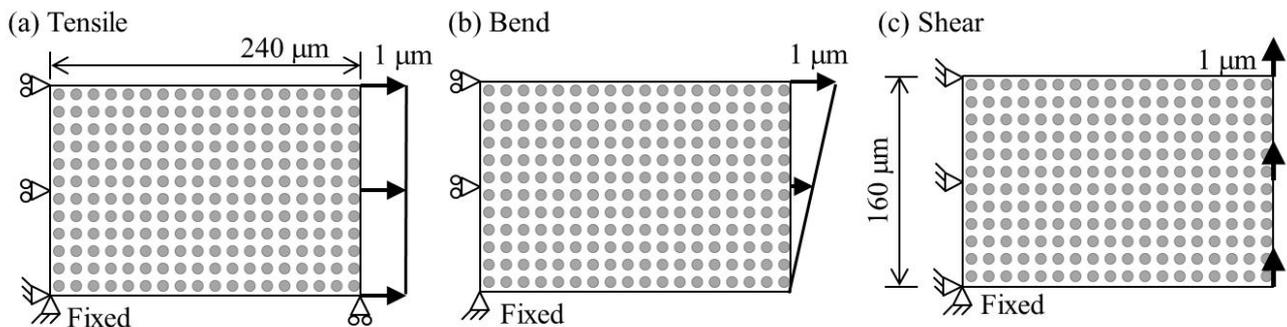
Figure 7. Three carbon fiber composite structures: (a)–(c) show the analysis model under a uniaxial tensile load, a bending load, and a shearing load, respectively

achieved by the SDM.
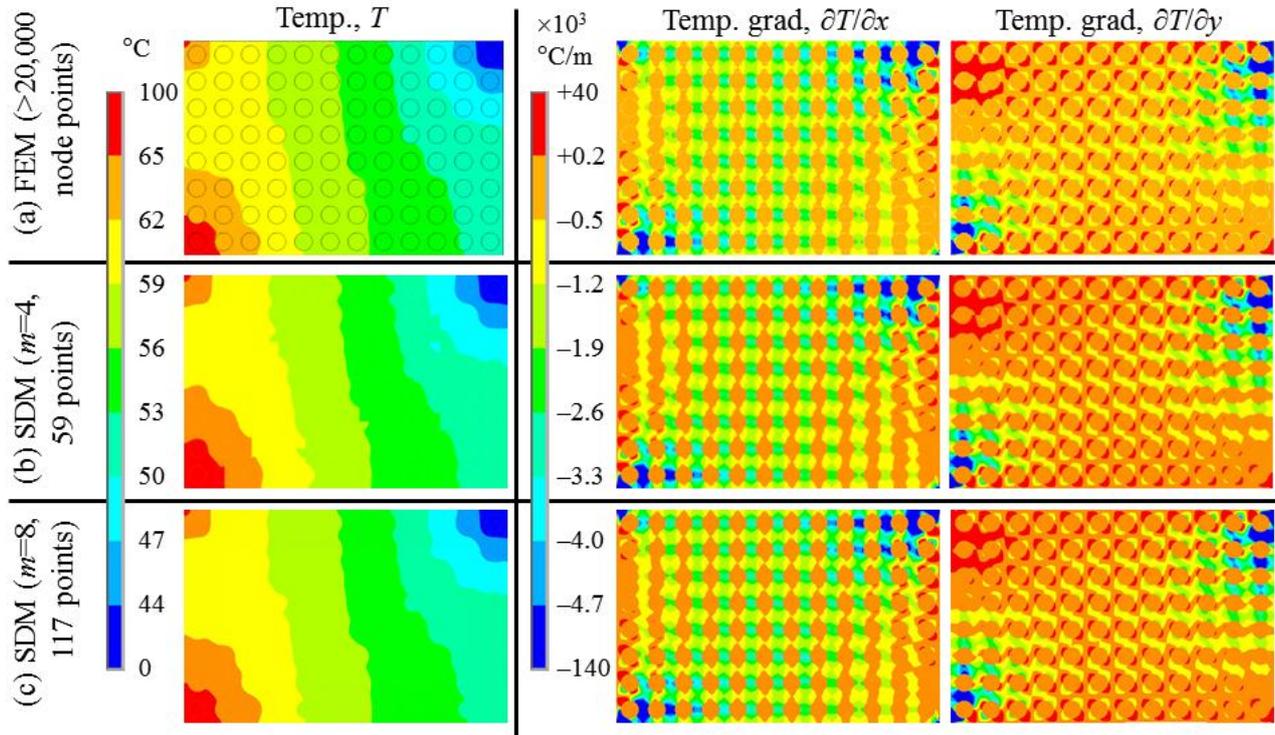
Figure 8. Calculated steady-state temperature and temperature gradient distributions in the composite material: (a)–(c) describe the results of FEM with >20,000 nodes, those of the SDM with 59 points (*m*=4), and those of the SDM with 117 points (*m*=8), respectively

Table 3. Comparison of analytical accuracies and calculation times

Item	DOF	Average difference in temp.	Average difference in temp. gradient		Total calculation time
Symbol		u	$\partial u/\partial x_1$	$\partial u/\partial x_2$	
Unit		Degree	$\times 10^4$ °C/m		s
FEM	>20,000	0	0	0	8.10
SDM (*m* = 4)	59	1.35	12.1	6.34	1.12
SDM (*m* = 8)	117	0.00308	4.42	0.781	1.14

Table 3 gives the average values of the absolute difference in temperature, *u*, and that in temperature gradients, $\partial u/\partial x$ and $\partial u/\partial y$, when comparing with the result of the conventional FEM. The average difference in *u* over all CPs in the case of 59 CPs ($m = 4$) and in the case of 117 CPs ($m = 8$) are 1.35, 0.00308 °C, respectively. Although the temperature result for $m = 4$ is not exactly correct in comparison with the temperature difference of 100 °C arising in the global domain, there is no temperature difference at all for $m = 8$. The average differences in $\partial u/\partial x$ and $\partial u/\partial y$ for $m = 8$ are 4.42×10^4 and 0.781×10^4 °C/m, respectively. These errors are sufficiently small because the maximum temperature gradients of $\partial u/\partial x$ and $\partial u/\partial y$ at all CPs are both 1.40×10^7 °C/m. As shown in Figs. 15 and 16, employing 117 CPs in the SDM ($m = 8$) gives highly continuous distributions of $\partial u/\partial x$ and $\partial u/\partial y$ that are almost the same as those of the exact result obtained with the FEM and >20,000 of nodes.

The total computational times are compared in Table 3. The results include the time expended in meshing domains and solving the inverse matrix of matrix in Eq. 9. For the SDM, total computational time of the local and global analyses is presented. In the SDM, increase of the

number of CPs enhances greatly the simulation precision but requires little additional calculation time. The conventional FEM with more than 20 thousands of nodes has 7.1 times the calculation cost of the SDM with 117 CPs in producing a solution as accurate as that of the SDM.

Results for linear elasticity

Figure 9 presents calculated displacement and strain contours in the composite models under: (a) a uniaxial tensile load; (b) a bending load; (c) a shearing load. As stated above, all the FEM results in Figure 9 are supposed to give the almost exact solution. Each SDM model is represented by 425 CPs. Therefore, if the SDM result accords with the FEM result, analytical error of the SDM is regarded as zero. As shown in Figs. 9(a)–(c), all the contours of displacement and strain obtained with the SDM are sufficiently continuous and appear to exactly match those obtained with the conventional FEM. Table 4 presents the average values of the absolute difference in displacement over all CPs between the SDM solutions and the FEM solutions. All of the average difference in displacement in the x axis, u, and that in the y axis, v, are less than 0.007 μm. The three composite materials in Figs. 7(a)–(c) receive respectively different types of loads but the same maximum displacement of 1 μm occurs in them. Therefore, these differences are sufficiently small compared with the maximum displacement in the domain and practically negligible.

As a traditional FEM generally interpolates displacement contour in each mesh from displacement at nodes in the mesh using a simple polynomial function, coarse-mesh layout cannot reproduce a complicated displacement profile. Especially, FEM models with fully integrated first-order quadrangular solid meshes (or hexahedral solid meshes in a case of three-dimensional analysis) tend to overestimate the stiffness in bending-dominated problems. This problem is called shear locking and occurs because a first-order mesh cannot represent a curved surface. However, as shown in Fig. 9(b), employing only 425 CPs in the SDM reproduce smooth bending deformation and does not cause shear locking because the displacement is interpolated by exact functions obtained by the prior analysis of fine-mesh local domain. For the similar reason, the SDM model is free of hourglass deformation. Hourglassing is another major problem of the FEM caused in certain coarse-grained domains. This is zero-energy deformation mode where a mesh deforms without consuming strain energy because all kinds of stresses at each integration point in the mesh are zero.

Table 5 presents comparison of the total computational times required to solve the problems. In giving a solution as accurate as that of the SDM, the conventional FEM requires more than one hundred times the calculation cost of the SDM. As the size of the global domain becomes large, the FEM has huge cost and becomes inefficient. In contrast, the SDM can analyze a large global domain without considerable cost and its competitiveness and effectiveness become higher as the global domain increases in size.

Table 4. Differences in displacement between the SDM and the FEM when analyzing three FRP models

Item	Ave. of absolute difference in displacement		Max. of absolute difference in displacement	
Symbol	u in the x axis	v in the y axis	u in the x axis	v in the y axis
Unit	μm			
Tensile	0.00179	0.00691	0.00490	0.00909
Bend	0.00090	0.00127	0.00357	0.00308

Table 5. Comparison of total calculation time

Load	Method	Calculation time
Tensile	FEM	425 sec
	SDM	3.46 sec
Bend	FEM	425 sec
	SDM	3.46 sec
Shear	FEM	425 sec
	SDM	3.46 sec

| Shear | 0.00019 | 0.00038 | 0.00180 | 0.00085 |

(a) Disp., u [μm]

(i) FEM with >300,000 nodes   (ii) SDM with 425 points (m = 8)

Normal strain, ε_x [×10⁻⁴]

(iii) FEM with >300,000 nodes   (iv) SDM with 425 points (m = 8)

(b) Disp., u [μm]

(i) FEM with >300,000 nodes   (ii) SDM with 425 points (m = 8)

Normal strain, ε_x [×10⁻⁴]

(iii) FEM with >300,000 nodes   (iv) SDM with 425 points (m = 8)

(c) Disp., v [μm]

(i) FEM with >300,000 nodes   (ii) SDM with 425 points (m = 8)

Normal strain, γ_{xy} [×10⁻⁴]

(iii) FEM with >300,000 nodes   (iv) SDM with 425 points (m = 8)

Figure 9. Calculated results of the FRP consisting of simple RVEs when applying: (a) a

uniaxial tensile load; (b) a bending tensile load; (c) a shearing load.

Conclusions

This manuscript proposes a new high-cost-effectiveness multiscale technique, termed the SDM, which is applicable to numerical analyses by modeling a macroscopic field as a "seamless" global domain that has no mesh, grid, cell, or control volume and requires only coarse-grained points. The SDM involves two numerical analyses. The first is an analysis of the local domain to construct accurate interpolating functions and influence coefficients for the second analysis. The second analysis is of the seamless global domain using the special interpolating functions and influence coefficients. There are two novelties of the SDM.

· Dependent-variable distributions interpolated by the special functions satisfy governing equations of the domain almost exactly.
· The seamless domain can enhance continuity of the dependent variable and its gradient.

Therefore, an accurate global solution of the entire field can be obtained by the SDM. This paper explains theoretical aspects of the SDM that are useful in the analysis of domains with a strong boundary effect, anisotropic fields, and heterogeneous materials, as well as isotropic homogenous fields. The analytical precision and computational cost of the SDM technique were investigated for two numerical example problems of two-dimensional periodic heterogeneous materials: stationary temperature field; linear elastic fields. The SDM models with eight reference points (i.e., $m = 8$) gave much more exact solutions in both example problems than the ordinal FEM.

As a result of the steady-state heat conduction analysis, the average differences in temperature, u, and temperature gradients, $\partial u/\partial x$ and $\partial u/\partial y$, between the solution of the SDM model represented by only 117 points and that of the conventional FEM model with 20 thousand nodes were 4.17×10^{-5} °C, 0.0305 °C/m, and 0.00649 °C/m, respectively. The temperature and its gradient were highly continuous over the entire material. This verified the feasibility of the SDM for a scalar temperature field. The ordinal FEM required 7.1 times the calculation cost of the SDM in producing a solution as accurate as that of the SDM.

Additionally, we also applied the SDM for linear elastic analysis of vector displacement fields. As a result, the SDM with 117 points provided as accurate solutions as those from conventional FEM using more than 300 thousands of nodes. For all the elastic examples, the average difference of displacement between the SDM solution and the FEM solution with respect to the maximum displacement was less than 0.7 %. The displacement and strain were sufficiently continuous throughout the entire global domains. The SDM solutions were shear-locking-free and hourglass-free because the global domains maintain high analytical resolution and the displacement distributions were flexible enough to fit the true ones. The FEM required more than 100 times the computational cost of the SDM in generating a solution at the same level of accuracy as that of the SDM.

References

Wang D, Chen JS. (2006) A locking-free meshfree curved beam formulation with the stabilized conforming nodal integration conforming nodal integration. *Comput Mech* **39**, 83–90.
Li S, Hao W, Liu WK. (2000) Numerical simulations of large deformation of thin shell structures using meshfree methods. *Comput Mech*; **25**, 102–116.
Joldes GR, Wittek A, Miller K. (2008) An efficient hourglass control implementation for the uniform strain hexahedron using the Total Lagrangian formulation. *Commun Numer Meth Engng* **24**, 1315–1323.
Chua J, Efendievb Y, Gintingc V, Hou TY. (2008) Flow based oversampling technique for multiscale finite element methods. *Advances in Water Resources* **31(4)**, 599–608.
Ilic S, Hackl K. (2009) Application of the multiscale FEM to the modeling of nonlinear multiphase materials. *J Theor Appl Mech* **47(3)**, 537–551.

Jenny P, Lee SH, Tchelepi HA. (2004) Adaptive multiscale finite-volume method for multiphase flow and transport in porous media. *Multiscale Mode. Simul* **3(1)**, 50–64.

Lunati I, Jenny P. (2006) Multiscale finite-volume method for compressible multiphase flow in porous media. *Journal of Computational Physics* **216(2)**, 616–636.

Kaczmarczyk Ł, Pearce CJ, Bićanić N. (2010) Studies of microstructural size effect and higher-order deformation in second-order computational homogenization. *Comput Struct* **88**(23–24), 1383–1390.

Henning P, Peterseim D. (2013) Oversampling for the multiscale finite element method. *arXiv(mathe NA)* arXiv:1211.5954.

Efendiev Y, Galvis J, Li G, Presho M. (2013) Generalized multiscale finite element methods. Oversampling strategies. *arXiv(mathe NA)* arXiv:1304.4888.

Suzuki Y, Soga K, Nakamura Y. (2014) Seamless domain method: a meshfree multiscale numerical analysis. *Int J Numer Meth Eng* (under review).

Rupnowski P, Gentz M, Sutter JK, Kumosa M. (2005) An evaluation of the elastic properties and thermal expansion coefficients of medium and high modulus graphite fibers. *Composites Part A* **36**(3): 327–338.

Calculation of underwater acoustic scattering problems in unbounded domain using the alpha finite element method

W. Li[1,2], *Y.F. Li[1] and Y.B. Chai[1]

[1]Department of Naval Architecture and Ocean Engineering, Huazhong University of Science and Technology, Wuhan City, P. R. China 430074

[2]Hubei Key Laboratory of Naval Architecture & Ocean Engineering Hydrodynamics (HUST)

*Corresponding Author: lyf198961@gmail.com

Abstract It is well-known that the classical finite element method fails to provide accurate results to the Helmholtz equation due to the dispersion error, which is rooted at the "overly-stiff" feature of the FEM model. By combining the "smaller wave number" model of FEM and the "larger wave number" model of NS-FEM, an alpha finite element method(α-FEM) can obtain accurate solutions. In this paper, the α-FEM has been applied to analyze 2D underwater exterior scattering problems in the unbounded domain. The non-reflecting boundary condition is imposed as an artificial boundary to model exterior acoustic problems. Several two-dimensional underwater exterior scattering problems with known exact solutions have been chosen as numerical examples. Results demonstrate the excellent properties of α-FEM.

Keywords: Alpha finite element method(α-FEM), Acoustic Scattering, Unbounded Domain, Non-reflecting boundary

Introduction

For several decades, many numerical methods have been introduced to compute the approximate solutions of acoustic problems [Suleau et al.(2000); Harari and Magoules(2004); Babuska et al(1999)]. The standard finite element method (FEM) is one of the most widely-used numerical methods in solving these acoustic problems governed by the Helmholtz equation. However it is known that the FEM fails to provides reliable predictions in high frequency range. Many studies have been done to improve resolve this defect. But such efforts have difficulties because of the well-known "pollution error".

   Various numerical methods have been proposed，They are the stabilized FEM [Harari and Huhes(1992; Thompson and Pinsky(1995)], higher order methods [Petersen et al.(2006)]，meshless method [Bouillard and Suleau(1998)] and so on. They all get better solutions. However, "softened" stiffness for the discrete model is more effective [Liu et al.(2009)]. The wave number in the FEM model is smaller than the actual one, leading to the so-called numerical dispersive error. The FEM model based on the standard Galerkin weak form behaviors stiffer than the continuous system. In order to "soften" the numerical system, Liu [Liu(2008; 2009)] has proposed generalized gradient smoothing technique and applied if in the meshfree setting to formulate the node-based smoothing point interpolation method (NS-PIM)

and node-based finite element method (NS-FEM). But NS-PIM and NS-FEM model both behave "overly-soft". An alpha finite element method (α-FEM) was then proposed [Liu et al.(2008)] by combining the "overly-stiffness" of the FEM and the "overly-soft" of the NS-FEM through a parameter α, resulting in a numerical model with very close-to-exact stiffness.

In this work, by introducing the DtN artificial boundary condition [Givoli(1988); Givoli and Keller(1989)], the accuracy and convergence of the α-FEM is studied. Initially the scattering problem is described. Next, the weak form of α-FEM and DtN boundary condition for the two dimensions case is derived. Finally, a comparison between the α-FEM solution, the FEM solution and the numerical solution shows the performance of the α-FEM for a rigid sphere as an example.

Mathematical model of acoustic problem

Consider an acoustic problem domain Ω with boundary Γ. The acoustic wave equation can be written as following form:

$$\nabla^2 p - \frac{1}{c^2} \frac{\partial^2 p}{\partial^2 t} = 0 \tag{1}$$

where p denotes the acoustic field pressure and c be the speed of sound traveling in fluid, ∇^2 and t denote the Laplace operator and time. In the frequency domain, the acoustic can be expressed as:

$$p = p_0 e^{j\omega t} \tag{2}$$

where p_0 and ω denote the amplitude of the acoustic wave and the angular frequency, $j = \sqrt{-1}$. Then p satisfies the well-known Helmholtz equation:

$$\nabla^2 p + k^2 p = 0 \tag{3}$$

where k is the wave number expressed as:

$$k = \frac{\omega}{c} \tag{4}$$

besides, the particle velocity υ in an ideal fluid is proportional to the gradient of the pressure:

$$\nabla p + j\rho\omega v = 0 \tag{5}$$

For unbounded domain problems, an artificial boundary condition is introduced. In this paper we introduce a so-called "Dirichlet-to-Neumann" boundary condition [Keller and Givili(1989)] on the outer boundary of the domain discretized with finite elements:

$$\nabla p \cdot n = -\mathbf{M} \cdot p \qquad (6)$$

where M is the DtN operator. In two dimensional problems, the DtN operator is expressed as;

$$M_n = -\frac{k}{\pi} \sum_{n=0}^{\infty} \frac{H_n^{(2)'}(kR)}{H_n^{(2)}(kR)} cosn(\theta - \vartheta) \qquad (7)$$

where R is the radius of the outer boundary, $H_n^{(2)}$ is the Hankel function of the second kind, θ and ϑ are azimuth angle.

Formulation of the $\alpha - $ FEM
In the standard FEM, the discretized system equation can be written in the following matrix form:

$$\left[\mathbf{K}^{\text{FEM}} - k^2\mathbf{M} + j\rho\omega\mathbf{C}\right]\{\mathbf{P}\} = -j\rho\omega\{\mathbf{F}\} \qquad (8)$$

where

The acoustic stiffness matrix: $\mathbf{K}^{\text{FEM}} = \int_{\Omega}(\nabla\mathbf{N})^{\text{T}}(\Delta\mathbf{N})d\Omega \qquad (9)$

The acoustic mass matrix: $\mathbf{M} = \int_{\Omega}\mathbf{N}^{\text{T}}\mathbf{N}d\Omega \qquad (10)$

The acoustic damping matrix: $\mathbf{C} = \int_{\Gamma}\mathbf{N}^{\text{T}}\mathbf{N}A_n d\Gamma \qquad (11)$

The vector of nodal acoustic forces: $\mathbf{F} = \int_{\Gamma}\mathbf{N}^{T}v_n d\Gamma \qquad (12)$

The nodal acoustic pressure: $\mathbf{P}^{\text{T}} = \left[p_1, p_2, \ldots, p_n\right] \qquad (13)$

In the NS-FEM, the problem is first divided into N_e elements with of N_n nodes as same as those used in the FEM. Then the problem is further divided in N_n node-based

smoothing domains on top of the generated cells. For 2D problems, the smoothing domain Ω_k is created by connecting sequentially the mid-edge-point to the centroids of the surrounding triangles of node k. The boundary of the smoothing domain Ω_k is labeled as Γ_k and the union of all Ω_k forms exactly the global domain Ω.

In the NS-FEM, the field variable is constructed using the linear FEM shape functions in the same way as those in the FEM. The difference is that the gradient component ∇N is replaced by the smoothing item $\overline{\nabla N}$ obtained using the node-based gradient smoothing operation [Liu et al.(2005); Zhang et al(2007)].The acoustic stiffness matrix in the NS-FEM is expressed as:

$$\mathbf{K}^{\text{NS-FEM}} = \int_{\Omega} (\overline{\nabla \mathbf{N}})^{T} (\overline{\nabla \mathbf{N}}) d\Omega \tag{14}$$

The above integration is evaluated base on the summation of all the node-based smoothing domains as:

$$\mathbf{K}^{\text{NS-FEM}} = \sum_{k=1}^{N_n} \mathbf{K}^{(k)} \tag{15}$$

where the $\mathbf{K}^{(k)}$ is the local smoothed stiffness matrix associated with node k, and can be calculated:

$$\mathbf{K}^{(k)} = \int_{\Omega_k} \overline{\mathbf{B}}^{\text{T}} \overline{\mathbf{B}} d\Omega = \overline{\mathbf{B}}^{\text{T}} \overline{\mathbf{B}} A_k \tag{16}$$

where A_k is the area of the smoothing domain for node k in 2D problems, and

$$\overline{\mathbf{B}}_i(\mathbf{x}_k) = \begin{bmatrix} \overline{b}_{i1} & \overline{b}_{i2} \end{bmatrix} \tag{17}$$

$$\overline{b}_{ip} = \frac{1}{A_k} \int_{\Gamma_k} N_i(\mathbf{x}) n_p(\mathbf{x}) d\Gamma \tag{18}$$

where is the FEM shape function for node i.

In the $\alpha - \text{FEM}$, each triangular element is divided into four parts with a scaled factor α: three quadrilaterals associated with three vertexes with equal area of $\frac{1}{3}\alpha A_e$ and the remaining Y-shaped part in the middle of the element with a area of $(1-\alpha)A_e$, where the A_e is the area of the triangular element. The NS-FEM and the FEM formulations are constructed respectively in the three quadrilaterals and the Y-shaped area for each element. Then the $\alpha - \text{FEM}$ will be the assembly from the entries of both the NS-FEM and FEM with the following form:

$$\mathbf{K}_{IJ}^{\alpha-\text{FEM}} = \alpha^2 \sum_{m=1}^{N^e} \mathbf{K}_{IJ(m)}^{\text{FEM}} + \left(1-\alpha^2\right) \sum_{n=1}^{N^n} \mathbf{K}_{IJ(n)}^{\text{NS-FEM}} \qquad (19)$$

In unbounded domain, according to Givoli and Kaller [Givoli(1988); Givoli and Keller(1989)], the stiffness matrix $\mathbf{K}$ consists of two matrices:

$$\mathbf{K} = \mathbf{K}^{\text{FEM}} + \mathbf{K}^b \qquad (20)$$

where $\mathbf{K}^b$ is the DtN artificial boundary matrix, it contains the operator M and the shape functions used in the FEM:

$$\mathbf{K}_{ij}^b = \int_{\Gamma} \mathbf{N}_i M \mathbf{N}_j d\Gamma \qquad (21)$$

Finally, the discretized system equations can be obtained and written in the following form:

$$\left[\mathbf{K}^{\alpha-\text{FEM}} - k^2 \mathbf{M} + j\rho\omega\mathbf{C} + \mathbf{K}^b\right]\{\mathbf{P}\} = -j\rho\omega\{\mathbf{F}\} \qquad (22)$$

Numerical example

In this paper, to illustrate the performance and ability of $\alpha-$ FEM for acoustic problems, the scattering problem on the exterior domain of a rigid sphere is dealt with. The radius of the sphere is 0.2, the radius of the artificial boundary is 1.

Consider a wave propagates in the exterior domain with two boundary condition described as follow:

on the artificial boundary:　$\nabla\mathbf{p}\bullet\mathrm{n} = -\mathbf{M}\bullet\mathbf{p}$ \qquad (23)

on the boundary of the rigid sphere:　$v_n = -v$ \qquad (24)

The problem has an analytical solution as follow:

$$p = -p_0 e^{j\omega t} \sum_{n=0}^{\infty} (-j)^n \varepsilon_n \frac{\dfrac{dJ_n(ka)}{dka}}{\dfrac{dH_n^{(2)}(ka)}{dka}} H_n^{(2)}(kr)\cos(n\theta), \quad \varepsilon_n = \begin{cases} 1, n=0 \\ 2, n=1,2,3 \end{cases} \qquad (25)$$

Three different α values $\alpha = 0.7, \alpha = 0.8, \alpha = 0.9$ have been employed to compare the influence of α with element size of 0.02. The numerical results of

acoustic pressure using $\alpha - FEM$ and exact solution are plotted in Fig1. It can be seen from plot that when $\alpha = 0.8$, the numerical solution is in agreement with exact solution. So $\alpha = 0.8$ is used in the following computation.
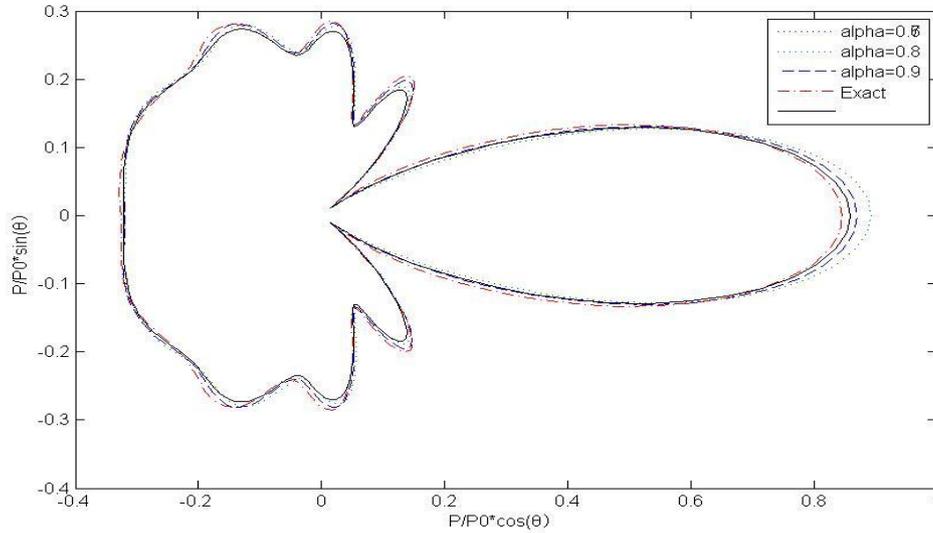
Figure 1. Analytical solution and α-FEM solution with different α values
$$\alpha = 0.7, \alpha = 0.8, \alpha = 0.9$$
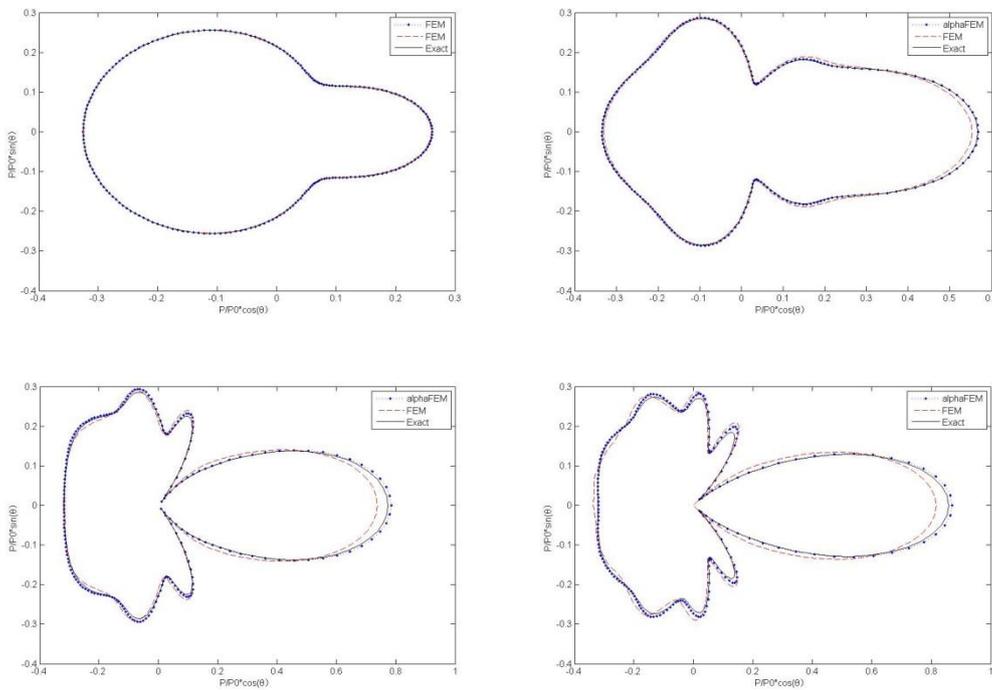
**Figure. 2 Analytical solution, FEM solution and α-FEM solution
on the artificial boudary for k=5, k=15, k=25, k=30**

Four different wave number values k=5, k=10, k=15, k=20are employed to study the accuracy of $\alpha - FEM$ on the artificial boundary in Fig2. These plots show that for lower wave number, $\alpha - FEM$ and FEM all can provide close-to-exact solution. But $\alpha - FEM$ is more close to exact solution than FEM solution when comparing them in the forward scattering. For higher wave number, the advantage of $\alpha - FEM$ is more obvious, $\alpha - FEM$ solution is still close to the exact solution on the artificial boundary, but FEM solution depart more from the exact solution.
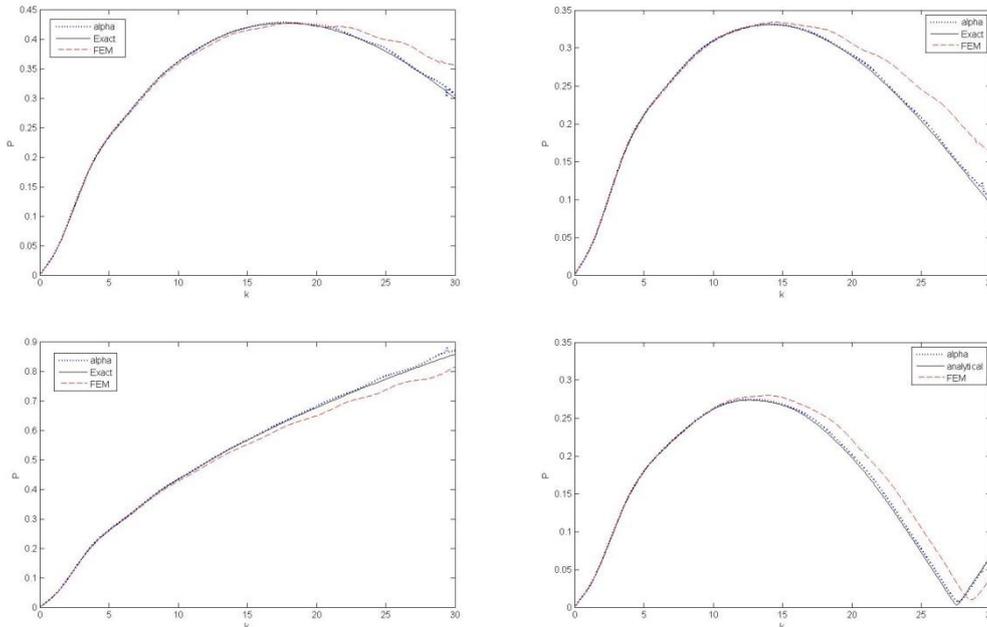
Figure 3. Calculation of $\alpha - FEM$ solution, FEM solution and analytical solutionas the wave number increasing at four nodes $\theta = 0, \theta = 0.5386, \; \theta = 0.7630, \; \theta = 1.0322$

In Fig3, we compare $\alpha - FEM$ and FEM solution with exact solution as the wave number increasing at four nodes $\theta = 0, \theta = 0.5386, \; \theta = 0.7630, \; \theta = 1.0322$. From these plots we can find that as wave number increasing, $\alpha - FEM$ and FEM both lose their accuracy, but the error of $\alpha - FEM$ is much smaller than the erroe of FEM.

Conclusions and discussions

In this work, the alpha finite element method (α-FEM) for solving scattering problems of the Helmholtz equation in two dimensions has been presented. By combining the "overly-stiff" FEM model with the "overly-soft" NS-FEM model, the α-FEM is obtained by a scaled factor $\alpha \in [0,1]$. Calculations of the scattering of a rigid sphere show the following conclusions:

1. The scaled factor α has a giant effect on the accuracy of the α-FEM.
2. The results indicate that the DtN boundary condition is a good alternative to other methods in solving scattering problems in infinite domains.

3. The α –FEM and Fem use the same mesh, which means the α-FEM model can be get from the FEM model with little change.

4. By using the gradient smoothing technology and the optimal alpha, the α-FEM appropriately softened the stiffness matrix and reduces the dispersion error. Numerical example demonstrate that the accuracy and convergence of the α-FEM is better than the FEM.

References

Suleau. S, Deraemaeker. A, Bouillard. Ph.(2000) Dispersion an pollution of meshless solution for the Helmholtz equation, Comput. Meth. Appl. Mech. Engrg. 190, 639-657.

Harari. I, Magoules. F.(2004)Numerical investigations of stabilized finite element computation for acoustics, Wave Motion 39, 339-329.

Deraemaeker. A, Babuska. I, Bouillard. Ph.(1999) Dispersion an pollution of the FEM solution for the Helmholtz equation in one, two and three dimension, Int. J. Numer. Meth. Engrg. 46, 471-499.

Harari. I, Hughes. TJR(1992)Galerkin/last squares finite element methods for the reduced wave equation with non-reflecting boundary conditions in unbounded domains. Comp Methods Appl Mech Eng 98, 411-454.

Thompson. LL, Pinksy. PM(1995)A Galerkin least squares finite element method for the two-dimensional Helmholtz equation. Int. J. Numer. Meth. Engrg. 38, 371-397.

Petersen. S, Dreyer. D, Estorff. Ov(2006)Assessment of finite and spectral element shape function or efficient iterative simulations of interior acoustics. Comp Methods Appl Mech Eng 195, 1171-1188.

Bouillard. Ph, Suleau. S(1998)Element-free Galerkin solutions for Helmholtz problems: formulation and numerical assessment of the pollution effect. Comp Methods Appl Mech Eng 162, 317-335.

He. Z. C, Liu G. R, Zhong. Z. H, Wu. S. C, Zhang. G. Y, Cheng. A. G(2009)An edge-based smoothed finite element method for analyzing three-dimensional acoustic problems. Comp Methods Appl Mech Eng 199, 20-33.

Liu. G. R.(2008)A generalized gradient smoothing technique and the smoothed bilinear form for Galerkin formulation of wide class of computational methods, Inyrtnational Journal of Computational Methods 5, 199-236.

Liu. G. R. (2008)Meshfree Methods:Moving Beyond the Finite Element Method, second ed, CRC Press, Boca Raton, USA.

Liu. G. R. Nguyen. T. T, Lan. K. Y(2009)A novel alpha finite element method for exact solution to mechanic problems using triangular and trtrahedral element. Comp Methods Appl Mech Eng 197, 3883-3897.

Givoli. D.(1988)A finite element method for large domain problems, Thesis, Stanford University.

Keller. J, Givoli. D.(1989)Exact non-reflecting boundary conditions, Journal of Computational Physics 82, 172-192.

Liu. G. R, Zhang. G. Y, Dai. K. Y, Wang. Y. Y, Zhong. Z. H, Li. G. Y, Han. X,(2005)A linear conforming point interpolation method for two-dimensional solid mechanics problems, Int J Comput Methods 2(4),645-665.

Liu. G. R, Zhang. G. Y, Dai. K. Y, Wang. Y. Y, Zhong. Z. H, Li. G. Y, Han. X.(2007) A linear conforming point interpolation method for three-dimensional elasticity problem, Int J Numer Methods Eng 72(13), 1524-1543.

Engineering the mechanical properties of graphene nanotube hybrid structures through structural modulation

K. Xia[1], H.F. Zhan[1], Y. Wei[1, 2], S.B. Sang[2], and *Y.T. Gu[1]

[1]School of Chemistry, Physics and Mechanical Engineering, Queensland University of Technology, Brisbane, Australia
[2]MicroNano System Research Center, Key lab of advanced Transducers and Intelligent Control System of the Ministry of Education & College of information Engineering, Taiyuan University of Technology, Taiyuan, China

*Corresponding author: yuantong.gu@qut.edu.au

Abstract

The excellent multi-functional properties of carbon nanotube (CNT) and graphene have enabled them as appealing building blocks to construct 3D carbon-based nanomaterials or nanostructures. The recently reported graphene nanotube hybrid structure (GNHS) is one of the representatives of such nanostructures. This work investigated the relationships between the mechanical properties of the GNHS and its structure basing on large-scale molecular dynamics simulations. It is found that increasing the length of the constituent CNTs, the GNHS will have a higher Young's modulus and yield strength. Whereas, no strong correlation is found between the number of graphene layers and Young's modulus and yield strength, though more graphene layers intends to lead to a higher yield strain. In the meanwhile, the presences of multi-wall CNTs are found to greatly strengthen the hybrid structure. Generally, the hybrid structures exhibit a brittle behavior and the failure initiates from the connecting regions between CNT and graphene. More interestingly, affluent formations of monoatomic chains and rings are found at the fracture region. This study provides an in-depth understanding of the mechanical performance of the GNHSs while varying their structures, which will shed lights on the design and also the applications of the carbon-based nanostructures.

Keywords: Graphene, Carbon nanotube, Tension, Molecular dynamics simulations

Introduction

In the past decade, low dimensional carbon-based nanomaterials, carbon nanotube (CNT) and graphene, have fascinated the scientific community with their excellent mechanical, electrical and thermal properties [Wang (2005); Zhu et al. (2010)]. Their appealing properties, e.g., Young's modulus up to 1 TPa, high fracture strength and low mass density, have enabled them with a wide promising applications in electronics, photonics, composite materials, energy storage, sensors, and bio-applications [Novoselov et al. (2012); De Volder et al. (2013); Dellinger et al. (2013)]. For example, a CNT-based mechanical mass sensor has been reported with a resolution of 1.7 yg (1 yg=10^{-24} g, corresponds to the mass of one proton) [Chaste et al. (2012)]. The enticing properties of carbon materials have driven the exploration of three-dimensional nanomateirals or nanostructures, such as the so-called "super" nanotubes or graphene that are made from CNTs [Coluci et al. (2006)], CNT bundles [Kis et al. (2004)], graphene/nanowire sandwich structures [Liu and Kuo (2013)], and the graphene nanotube hybrid structure (GNHS) [Zhu et al. (2012)].

The GNHS, as one of the representatives of the three-dimensional nanomaterials, has been envisioned with various potential applications in adsorption, biosensors, batteries, nanocomposite, drug delivery, and so on [Zhao et al. (2012); Odedairo et al. (2014)]. For instance, it is found that the GNHS can store up to 41 g H_2/L under ambient conditions with lithium cations doping [Dimitrakakis et al. (2008)]. Experiments show that the GNHS has a high specific capacitance and remarkable rate capability, which significantly outperform many electrode materials currently used

in the state-of-the-art supercapacitors [Du et al. (2011)]. Hernandez-Ferrer *et al.* [Hernández-Ferrer et al. (2014)] shows that the multi-walled CNT/graphene hybrid materials possess higher heterogeneous charge transfer constant and sensitivity towards H_2O_2 reduction, which is promising for the applications in electrochemical sensing.

Series of studies have been carried out on GNHSs to facilitate their diverse applications. For example, using first principle calculations, the conductance of the GNHS is found to nearly independent of the CNT length, but changes strongly with the link structure [Novaes et al. (2010)]. Molecular dynamics simulations have shown that the thermal transport of the GNHS is governed by the minimum CNT's distance and length, due to the scattering of phonons occurred at the CNT-graphene junctions [Varshney et al. (2010)], and the long-wavelength out-of-plane modes contribute significantly to thermal conduction [Loh et al. (2011)]. Using the phonon wave packet method, Lee *et al.* [Lee et al. (2012)] found that the graphene interface distributes the incoming phonon energy to both sides of the interface almost equally, and provides a strong diffusive scattering site. It is noticed that majority of current studies have focused on the electrical and thermal properties of the GNHS, and their mechanical properties are still lacking of investigation. One recent work reported the mechanical properties of four different GNHS basing on the structural molecular mechanics approach [Sihn et al. (2012)]. Though our pervious works have probed the mechanical properties of GNHS with different CNT allocations [Zhan et al. (2013)] and dopants [Xia et al. (2014)], a more comprehensive understanding of the structural impacts on the associated mechanical properties are still required.

Therefore, in this work, we will investigate how the tensile properties of the GNHS can be tailored through the structure variation by the large-scale molecular dynamics (MD) simulations. Experiments have already reported the controllable fabrication of the graphene nanotube hybrid structures *via* one-step chemical vapor deposition process [Paul et al. (2010)]. Emphasis will be laid on the mechanical properties (Young's modulus, yield strength and yield strain) of the GNHS under tensile loading. Different hybrid structures will be constructed by either varying the length of the single-wall CNTs (SWNTs) or the number of graphene layers.

Computational details

A series of large-scale MD simulations were performed using the open-source LAMMPS code [Plimpton (1995)]. All tested hybrid structures have a similar width around 5.6 nm. The hybrid structures are constructed by connecting an armchair (4,4)-CNT with graphene through heptagons (see Figure 1) [Matsumoto and Saito (2002); Dimitrakakis, Tylianakis et al. (2008); Varshney, Patnaik et al. (2010)]. To acquire the tensile properties of different hybrid structures, two groups of models have been established, including the GNHSs either with different CNT lengths or with different number of graphene layers. For discussion convenience, the GNHSs with different CNT lengths are denoted as GNHS-L#, where # ranges from 1 to 6 representing the CNT length ranging from 24.9 Å to 51.9 Å (these models possess two graphene layers). The GNHSs with different number of graphene layers are denoted as GNHS-N#, where # runs from 1 to 3 representing three, four and five graphene layers (see Figure 1b). All models will be constructed by single-wall CNTs (SWNTs), as shown in Figure 1.
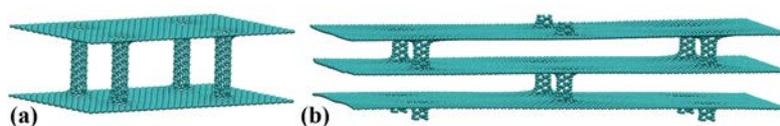
Figure 1. Simulation models of the graphene nanotube hybrid structures with: (a) two graphene layers; (b) three graphene layers.

To describe the interactions between carbon atoms, the popularly applied reactive empirical bond order (REBO) potential [Brenner et al. (2002)] was adopted, which has been shown to well represent the binding energy and elastic properties of graphene and CNT [Zhang et al. (2011)]. The whole simulation is carried out in three steps, including initial energy minimization using the conjugate gradient algorithm, relaxation under the NVT ensemble using Nose-Hoover thermostat [Nosé (1984); Hoover (1985)], and then tension of the sample. To minimize the thermal influence, a relatively low temperature, 1 K is used during the simulation. A time step of 1 fs is chosen, and the hybrid structure is relaxed for 500 ps. During the tensile deformation, one end of the GNHS was fixed, and a constant velocity of 0.001 Å/ps was applied to the other end along the length direction. The equations of motion are integrated with time using a Velocity Verlet algorithm [Verlet (1967)], with no periodic boundary conditions being applied.

During the simulation, the virial stress is applied to calculate the atomic stress $\Pi^{\alpha\beta}$ during tension, which is expressed as [Diao et al. (2004)]

$$\Pi^{\alpha\beta} = \frac{1}{\Omega}\sum_i \varpi_i \pi_i^{\alpha\beta} \ , \ \pi_i^{\alpha\beta} = \frac{1}{\varpi_i}\left(-m_i v_i^\alpha v_i^\beta + \frac{1}{2}\sum_{j\neq i} F_{ij}^\alpha r_{ij}^\beta\right) \tag{1}$$

Here $\pi_i^{\alpha\beta}$ is the atomic stress associated with atom i. ϖ_i is the effective volume of the ith atom and Ω is the volume of the whole system. m_i and v_i are the mass and velocity of the ith atom, respectively. F_{ij} and r_{ij} are the force and distance between atoms i and j, respectively, and the indices α and β denote the Cartesian components. The engineering strain is used to estimate the strain as $\varepsilon = (l - l_0)/l_0$, where l and l_0 are the instantaneous and initial length of the sample.

Results and discussions

Impacts from the nanotube's length

In the beginning, we assess the tensile properties of GNHSs containing two layers that are pillared by SWNTs with different lengths. Six cases are considered, with the SWNT's length ranging from 24.9 to 58.2 Å (see Figure 1). Figure 2a shows the stress-strain curves obtained from MD simulations. Generally, all stress-strain curves share an identical changing trend, i.e., the stress increases linearly at the beginning, and decreases suddenly after passing a specific stress value. According to the continuum mechanics [Gere and Timoshenko (1999)], such stress-strain profile signifies that the GNHS exhibits a brittle behavior. Following previous researchers [Zhan and Gu (2011)], Young's modulus is derived directly from the stress-strain curve with the strain <3% using linear regression. The yield strength is referred to the peak stress after which the failure of the structure is observed. The corresponding strain is denoted as yield strain. For the GNHS with the CNT length of 24.9 Å, a Young's modulus of 187.9 GPa is estimated, with the yield strength and yield strain calculated as 19.2 GPa, and 10.7%, respectively.

It is observed from Figure 2a that the longer length of the SWNT, the lower Young's modulus and yield strength of the GNHS. For example, the GNHS constructed by the shortest SWNT (length of 24.9 Å) shows the highest Young's modulus which is almost two times of its counterpart with 51.9 Å SWNTs. According to Figure 2b, the effective Young's modulus of the GNHSs exhibits a linear relationship with the length of the SWNT. This phenomenon is understandable, as the effective atom volume is a linear function of the SWNT's length, which will thus lead to a linear reduction to the atomic stress, see Eq. (1). Therefore, it is reasonable to observe a linearly decreased Young's modulus and yield strength with the increasing SWNT's length. Additionally, we find that the yield strain for all considered samples appears almost the same (around 10.6%, consistent with the results

reported by other researchers [Xu et al. (2012)]), implying that the SWNT's length exerts ignorable influence on the yield strain. Such observations can be explained by considering that the elastic tensile deformation of different GNHS is dominated by the two (upper and lower, see Figure 1) graphene layers. In other words, regardless how the length of the SWNT changes, the graphene layers will absorb majority of the tensile strain energy while the GNHS is elongated along *the* length direction, and thus induce a similar yield strain.
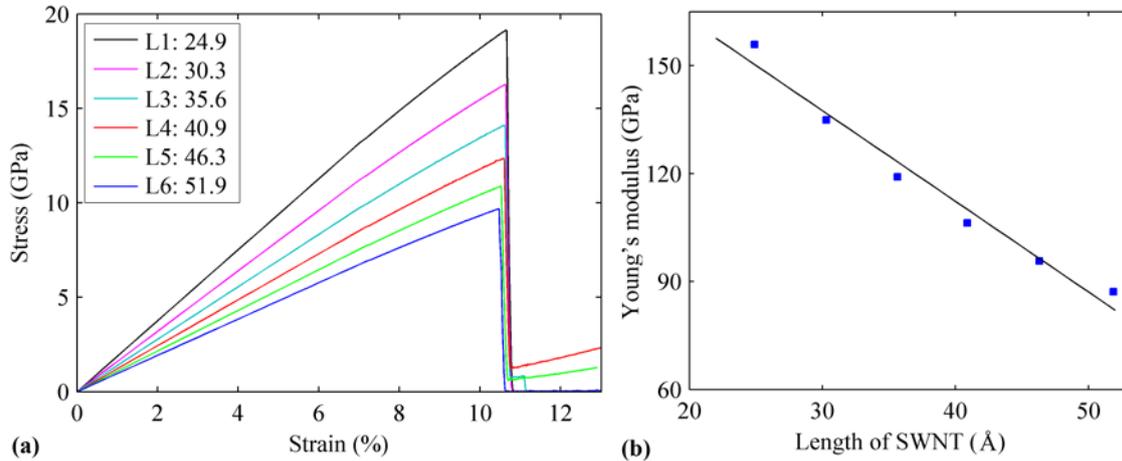
Figure 2. (a) Stress-strain curves of GNHS containing two graphene layers pillared by SWNT with different lengths; (b) The effective Young's modulus as a function of the SWNT's length.

To further investigate the impacts on the tensile properties of GNHS from the SWNT's length, we acquire the atomic configurations of the GNHS at different strains. In agreement with the stress-strain curves presented in Figure 2a, the atomic structure for all GNHSs show marginal changes during the elastic deformation period (before reaching the peak stress). After yielding, bond breakings are found. Generally, GNHS-L1, L2 and L3 share the same deformation behaviors. As illustrated in Figure 3a, during the elastic deformation period, the C-C bonds are stretched along the loading direction. With the increasing strain, fracture of the structure is observed, which is initiated from two CNT-graphene junctions (one on the top layer, the other on the bottom layer) between graphene layer and CNT near the two ends (Figure 3b). From Figure 3c, the structure is found to failure around two connecting regions with further extension. Particularly, after failure, local buckling of the graphene sheet due to the release of the strain as well as the inclined CNTs is observed. It is interesting to mention that at the front of the failure region, the formations of monoatomic chains (inset of Figure 3b) and monoatomic rings (inset of Figure 3c) are observed.
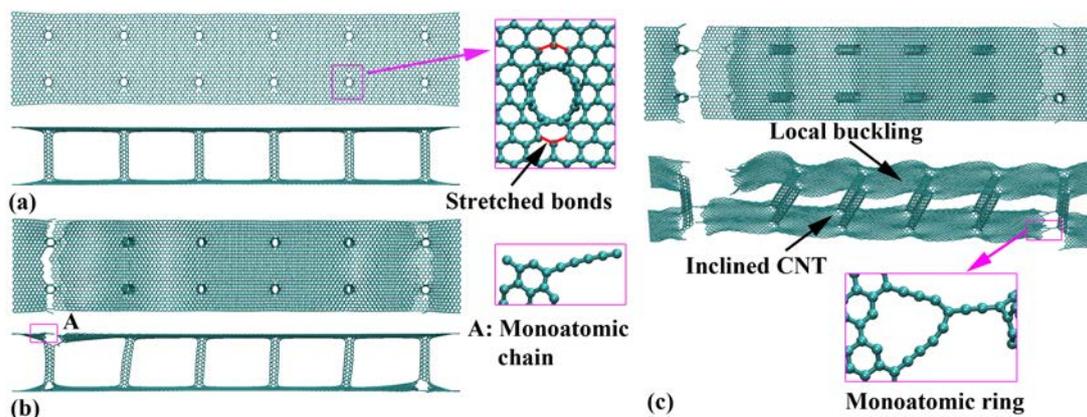
Figure 3. GNHS-L3 at the strain of: (a) 10.59%, inset shows the stretched C-C bonds around the CNT-graphene junction; (b) 10.65%, inset shows the formation of monoatomic chain; (c) 10.76%, inset shows the existence of monoatomic ring.

Similar deformation procure is also observed for the hybrid structure with longer CNTs (i.e., GNHS-L4 and GNHS-L5). As illustrated in Figure 4a, the CNT pillars appear a larger inclined angles due to the extension of the top and bottom graphene layers. The fracture is nucleated from the connecting regions and the graphene sheet is bulked after failure. The existence of monoatomic chains and rings are also observed (Figure 4b). Particularly, unlike the atomic configurations in Figure 3c, the top and bottom graphene layers are still connected by the inclined CNTs after failure for the GNHS-L4. Such phenomenon is consistent with the stress-strain curves in Figure 2a, from which the stress does not reduce directly to zero after failure. For the hybrid structure with the longest CNTs considered in this paper (GNHS-L6), a different deformation scheme is observed besides the monoatomic chains and rings emerged around the fracture region. As shown in Figure 4c, the fracture initiates from one end of the structure, which results in an early separation at the fracture region. In particular, the CNTs are still vertically aligned after yielding and a more evident local buckling of the graphene sheet is observed.
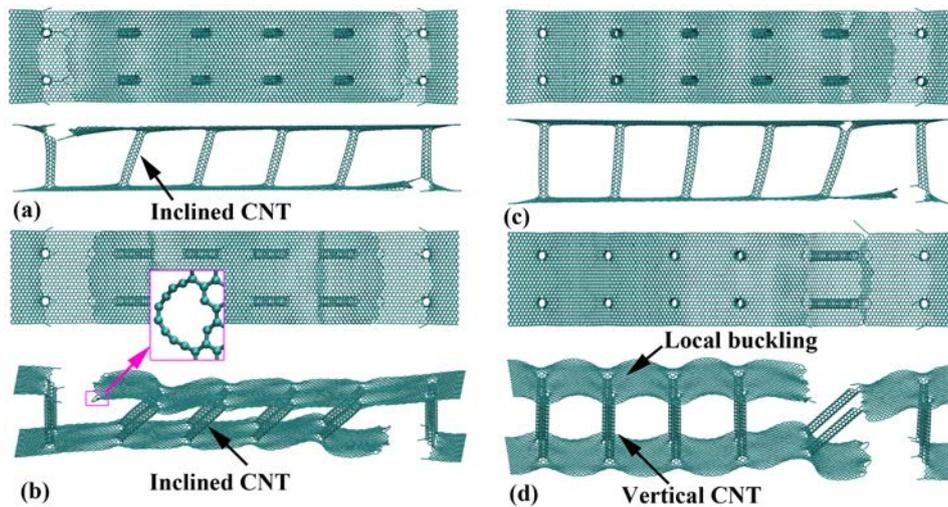
Figure 4. GNHS-L4 at the stain of: (a) 10.66%, (b) 10.74%, inset shows the formation of monoatomic chain; GNHS-L6 at the strain of: (a) 10.54%, (b) 10.64%.

Influence of the graphene layer number

Above discussions suggest that the increase of SWNT's length will result a direct reduction to the effective Young's modulus and yield strength of the hybrid structure. In this section, we investigate the influence on the tensile properties of GNHS from the number of graphene layers, i.e., the GNHS contains three, four and five graphene layers pillared by SWNTs (see Figure 1).

From Figure 5, the tensile properties of the GNHS vary with different number of graphene layers, and no explicit correlation is found between them. For example, comparing with the hybrid structure with two graphene layers, the GNHS-N1 (three-layers) shows a much smaller Young's modulus and yield strength, around 116.3 and 12.8 GPa, respectively. Whereas, increasing the number of graphene layers from three to four (GNHS-N2) yields a larger Young's modulus (139.6 GPa) and yield strength (16.2 GPa), and the hybrid structure with more graphene layers (GNHYS-N3) exhibit a similar tensile properties as estimated from GNHS-N2. In general, the yield strain is found to increase with the increasing number of graphene layers.
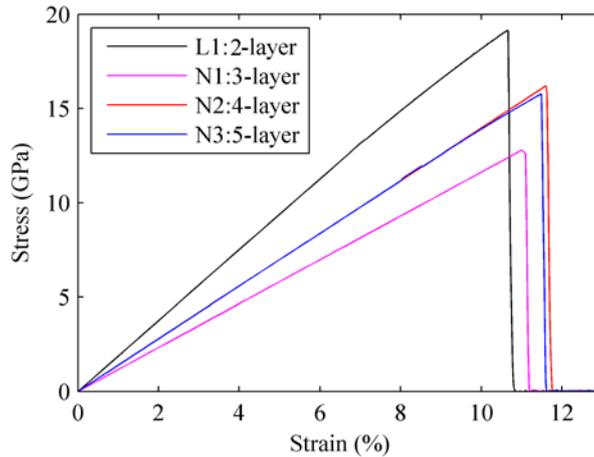
Figure 5. Stress-strain curves of GNHSs with different number of graphene layers.

The atomic configurations of different GNHSs have been tracked to identify the in-depth influence from the number of graphene layers. Similar as the above results, during the elastic deformation period, only stretched C-C bonds around the connecting regions are found (Figure 6a). When the stress passes the peak value, some of those stretched bonds are found to break, which leads to the failure of the structure (Figure 6b). The formation of monoatomic chains and rings and also inclined CNTs are found upon failure. In particular, the dangling graphene layers are found to adhere to each other forming a bi-layer or triple-layer graphene structure. As highlighted in Figure 6c, the distance between different dangling graphene layers is around 3.4 Å, which equals to the layer distance observed in multi-layer graphene. The emergence of the multi-layer graphene is originated from the shorter CNTs being utilized while establishing the model. After yielding, the dangling graphene layers will vibrate freely and the short distance between the two layers makes them easier to be adhered to each other through *van der Waals* force.
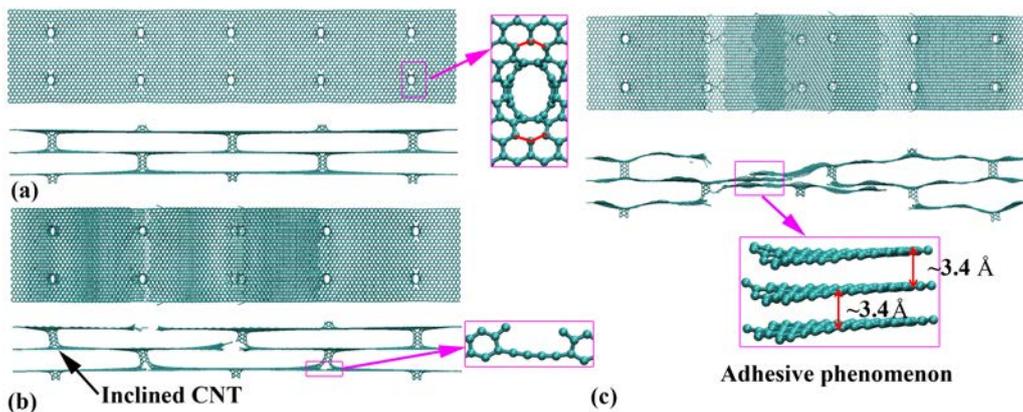
Figure 6. GNHS-N1 at the strain of: (a) 10.60%, inset shows the stretched carbon bonds; (b) 11.13%; (c) 11.30%, inset shows the adhered dangling graphene layers.

The incorporation of more graphene layers induces more complex deformation after yielding. In all three cases with three, four and five graphene layers, we found the fracture starts from the top layer and then propagates to the underneath layer. The existences of graphene layers have led to an easy formation of multi-layered graphene after failure (Figure 7b and 7d). For example, the residual atomic chain between graphene and CNT is found to be bended by the strong *van der Waals* interactions between the graphene layers (inset of Figure 7b).

To further investigate the structure influence on the mechanical properties of the hybrid structure, we also examine the tensile properties of GNHS with multi-walled CNTs (see inset of Figure 8a).

For the sample with a similar size as GNHS-L3 (length of 25.5 nm, width × height = 5.8 × 3.5 nm^2), a much higher Young's modulus (256.9 GPa) and yield strength (25.4 GPa) are estimated. As shown in Figure 8b, the failure of the structure is observed around the two outer CNTs in one end of the GNHS. After yielding, many monoatomic chains are formed at the fracture region. Particularly, it is found that only the outer CNTs have been split due to the tensile strain and the inner CNTs appears almost unchanged. Unlike the hybrid structures with SWNTs, no local bucking of the graphene layers is observed for the sample with multi-walled CNTs after yielding.
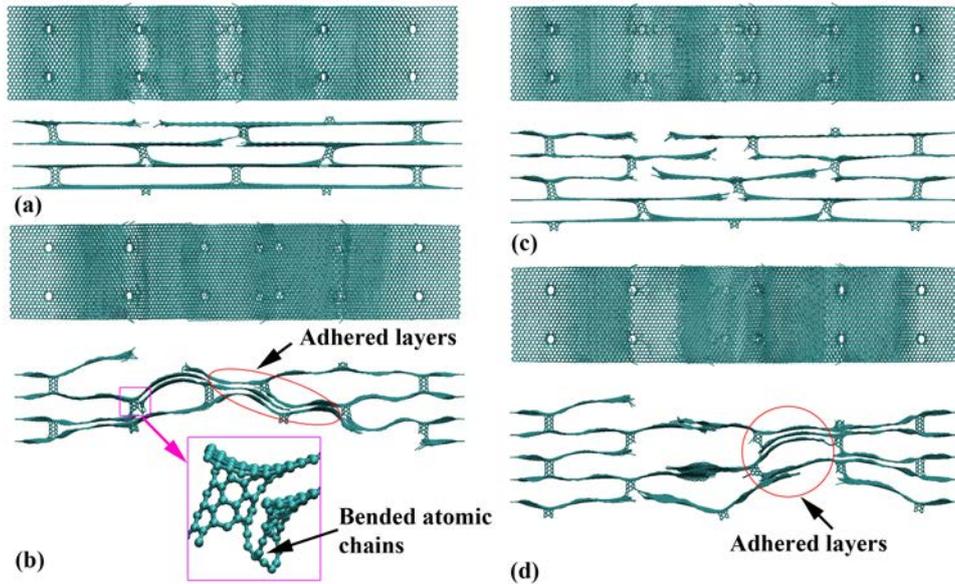
Figure 7. GNHS-N2 at the strain of: (a) 11.65%, (b) 12.63%, inset shows the bended atomic chains due to the *van der Waals* interactions between the adhered graphene layers; GNHS-N3 at the strain of: (c) 11.59%, (d) 11.87%.
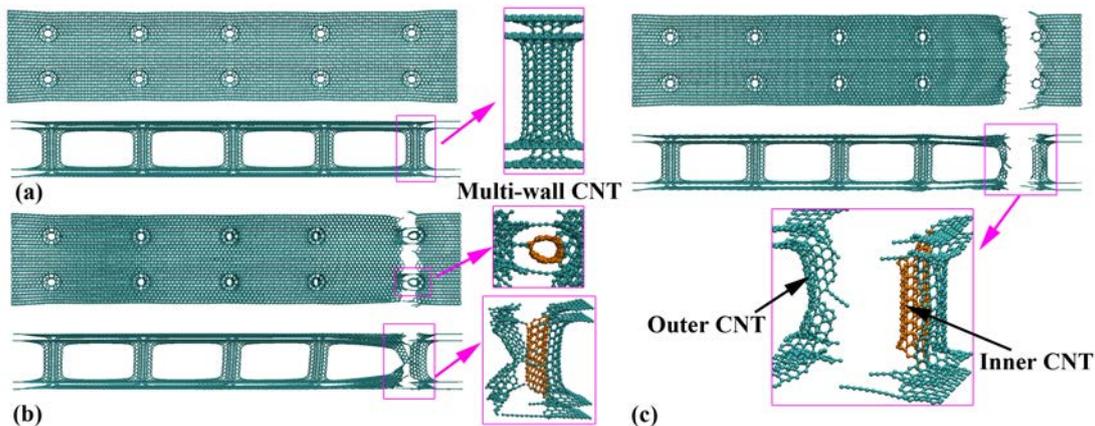
Figure 8. GNHS with multi-walled CNTs at the strain of: (a) 9.4%, inset shows a multi-walled CNT-graphene junction; (b) 10.1%, inset shows the deformation around the multi-walled CNT-graphene junctions; (c) 10.15%, inset shows the fracture of the outer CNTs.

Before concluding, we summarize the estimated tensile properties from all tested hybrid structures as listed in Table 1. Basically, Young's modulus and yield strength decrease with the increasing CNT length but have no strong relation with the number of graphene layers. In the other hand, the CNT length exerts marginal influence to the yield strain, and the increase of graphene layer number intends to increase the yield strain. In addition, comparing with its counterpart with SWNTs, the GNHS with multi-walled CNTs exhibits a much higher Young's modulus and yield strength, and a lower yield strain.

Table 1. Summary of Young's modulus, yield strength and yield strain for all considered graphene nanotube hybrid structures.

Different GNHSs		Young's modulus (GPa)	Yield strength (GPa)	Yield strain (%)
Bi-layer; SWNT's lengths L1-L6 (Å)	L1: 24.9	187.9	19.2	10.7
	L2: 30.3	159.9	16.3	10.6
	L3: 35.6	138.8	14.1	10.6
	L4: 40.9	121.7	12.4	10.6
	L5: 46.3	107.6	10.9	10.5
	L6: 51.9	96.2	9.7	10.5
Graphene layer number N1-N3	N1: 3	116.3	12.8	11.0
	N2: 4	139.6	16.2	11.6
	N3: 5	138.6	15.9	11.5
Multi-wall CNTs		256.9	25.4	9.9

Conclusions

Basing on large-scale MD simulations, we have examined the mechanical properties of graphene nanotube hybrid structures. It is found that the increase of CNT's length will induce a higher Young's modulus and yield strength, together with ignorable impacts to the yield strain of the hybrid structure. Whereas, no strong correlation is found between the number of graphene layers and Young's modulus and yield strength, though more graphene layers intends to lead to a higher yield strain. The presence of multi-wall CNTs are found to greatly strengthen the hybrid structure, which results in a much higher Young's modulus and yield strength. All studied hybrid structures exhibit a brittle behavior and the stretched carbon bonds are observed around the connecting regions between graphene and CNT during the elastic deformation period. Specifically, affluent formations of monoatomic chains and rings are found at the fracture region. For the hybrids structures with short CNT pillars, the dangling graphene layers will adhere to each other to form a multi-layered graphene after failure driven by *van der Waals* force. This study provides an in-depth understanding of relationship between the mechanical performance of the hybrid structures and their structures, which will shed lights on the design and also the applications of carbon-based nanostructures.

Acknowledgment

References

Brenner, D. W., O. A. Shenderova, J. A. Harrison, S. J. Stuart, B. Ni and S. B. Sinnott (2002). A second-generation reactive empirical bond order (REBO) potential energy expression for hydrocarbons *Journal of Physics: Condensed Matter* **14**, 783.

Chaste, J., A. Eichler, J. Moser, G. Ceballos, R. Rurali and A. Bachtold (2012). A nanomechanical mass sensor with yoctogram resolution *Nature Nanotechnology* **7**, 301-304.

Coluci, V. R., D. S. Galvao and A. Jorio (2006). Geometric and electronic structure of carbon nanotube networks:'super'-carbon nanotubes *Nanotechnology* **17**, 617.

De Volder, M. F., S. H. Tawfick, R. H. Baughman and A. J. Hart (2013). Carbon nanotubes: present and future commercial applications *Science* **339**, 535-539.

Dellinger, A., Z. Zhou, J. Connor, A. Madhankumar, S. Pamujula, C. M. Sayes and C. L. Kepley (2013). Application of fullerenes in nanomedicine: an update *Nanomedicine* **8**, 1191-1208.

Diao, J., K. Gall and M. L. Dunn (2004). Atomistic simulation of the structure and elastic properties of gold nanowires *Journal of the Mechanics and Physics of Solids* **52**, 1935-1962.

Dimitrakakis, G. K., E. Tylianakis and G. E. Froudakis (2008). Pillared graphene: a new 3-D network nanostructure for enhanced hydrogen storage *Nano Letters* **8**, 3166-3170.

Du, F., D. Yu, L. Dai, S. Ganguli, V. Varshney and A. Roy (2011). Preparation of tunable 3D pillared carbon nanotube–graphene networks for high-performance capacitance *Chemistry of Materials* **23**, 4810-4816.

Gere, J. and S. Timoshenko (1999). Mechanics of Materials, Cheltenham, UK: Stanley Thornes.

Hernández-Ferrer, J., P. Laporta, F. Gutiérrez, M. D. Rubianes, G. Rivas and M. T. Martínez (2014). Multi-walled carbon nanotubes/graphene nanoribbons hybrid materials with superior electrochemical performance *Electrochemistry Communications* **39**, 26-29.

Hoover, W. G. (1985). Canonical dynamics: Equilibrium phase-space distributions *Physical Review A* **31**, 1695-1697.

Kis, A., G. Csanyi, J.-P. Salvetat, T.-N. Lee, E. Couteau, A. Kulik, W. Benoit, J. Brugger and L. Forro (2004). Reinforcement of single-walled carbon nanotube bundles by intertube bridging *Nature Materials* **3**, 153-157.

Lee, J., V. Varshney, J. S. Brown, A. K. Roy and B. L. Farmer (2012). Single mode phonon scattering at carbon nanotube-graphene junction in pillared graphene structure *Applied Physics Letters* **100**, 183111-183111-183114.

Liu, B.-T. and H.-L. Kuo (2013). Graphene/silver nanowire sandwich structures for transparent conductive films *Carbon* **63**, 390-396.

Loh, G. C., E. H. T. Teo and B. K. Tay (2011). Interpillar phononics in pillared-graphene hybrid nanostructures *Journal of Applied Physics* **110**, 083502.

Matsumoto, T. and S. Saito (2002). Geometric and electronic structure of new carbon-network materials: Nanotube array on graphite sheet *Journal of the Physical Society of Japan* **71**, 2765.

Nosé, S. (1984). A unified formulation of the constant temperature molecular dynamics methods *The Journal of Chemical Physics* **81**, 511.

Novaes, F. D., R. Rurali and P. Ordejón (2010). Electronic transport between graphene layers covalently connected by carbon nanotubes *ACS Nano* **4**, 7596-7602.

Novoselov, K., V. Fal, L. Colombo, P. Gellert, M. Schwab and K. Kim (2012). A roadmap for graphene *Nature* **490**, 192-200.

Odedairo, T., J. Ma, Y. Gu, J. Chen, X. S. Zhao and Z. Zhu (2014). One-pot synthesis of carbon nanotube-graphene hybrids via syngas production *Journal of Materials Chemistry A* **2**, 1418-1428.

Paul, R. K., M. Ghazinejad, M. Penchev, J. Lin, M. Ozkan and C. S. Ozkan (2010). Synthesis of a Pillared Graphene Nanostructure: A Counterpart of Three-Dimensional Carbon Architectures *Small* **6**, 2309-2313.

Plimpton, S. (1995). Fast parallel algorithms for short-range molecular dynamics *Journal of Computational Physics* **117**, 1-19.

Sihn, S., V. Varshney, A. K. Roy and B. L. Farmer (2012). Prediction of 3D elastic moduli and Poisson's ratios of pillared graphene nanostructures *Carbon* **50**, 603-611.

Varshney, V., S. S. Patnaik, A. K. Roy, G. Froudakis and B. L. Farmer (2010). Modeling of thermal transport in pillared-graphene architectures *ACS Nano* **4**, 1153-1161.

Verlet, L. (1967). Computer 'experiments' on classical fluids. I. Thermodynamical Properties of Lennard-Jones Molecules *Physical Review* **159**, 98-103.

Wang, J. (2005). Carbon-nanotube based electrochemical biosensors: A review *Electroanalysis* **17**, 7-14.

Xia, K., H. Zhan, Y. Wei and Y. Gu (2014). Tensile properties of a boron/nitrogen-doped carbon nanotube–graphene hybrid structure *Beilstein Journal of Nanotechnology* **5**, 329-336.

Xu, L., N. Wei, Y. Zheng, Z. Fan, H.-Q. Wang and J.-C. Zheng (2012). Graphene-nanotube 3D networks: intriguing thermal and mechanical properties *Journal of Materials Chemistry* **22**, 1435-1444.

Zhan, H. F. and Y. T. Gu (2011). Atomistic exploration of deformation properties of copper nanowires with pre-existing defects *CMES: Computer Modeling in Engineering & Sciences* **80**, 23-56.

Zhan, H. F., K. Xia and Y. T. Gu (2013). Tensile properties of graphene-nanotube hybrid structures: a molecular dynamics study *International Journal of Computational Materials Science and Engineering* **02**, 1350020.

Zhang, Y., C. Wang, Y. Cheng and Y. Xiang (2011). Mechanical properties of bilayer graphene sheets coupled by sp^3 bonding *Carbon* **49**, 4511-4517.

Zhao, M.-Q., X.-F. Liu, Q. Zhang, G.-L. Tian, J.-Q. Huang, W. Zhu and F. Wei (2012). Graphene/Single-Walled Carbon Nanotube Hybrids: One-Step Catalytic Growth and Applications for High-Rate Li–S Batteries *ACS Nano* **6**, 10759-10769.

Zhu, Y., L. Li, C. Zhang, G. Casillas, Z. Sun, Z. Yan, G. Ruan, Z. Peng, A.-R. O. Raji and C. Kittrell (2012). A seamless three-dimensional carbon nanotube graphene hybrid material *Nature Communications* **3**, 1225.

Zhu, Y., S. Murali, W. Cai, X. Li, J. W. Suk, J. R. Potts and R. S. Ruoff (2010). Graphene and graphene oxide: synthesis, properties, and applications *Advanced Materials* **22**, 3906-3924.

Complex modal analysis using undamped modes

*Yujin Hu, and Li Li

School of Mechanical Science and Engineering, Huazhong University of Science and Technology, China.

*Corresponding author: yjhu@mail.hust.edu.cn

Abstract

Based on undamped modes, a novel method is presented to efficiently calculate complex eigenpairs by combining the Neumann series and the reduced basis technique. To avoid the modal truncation problem, the reduced basis is calculated by using a Neumann series expansion and only requires the undamped eigenpair of interest. The sufficient condition for the convergent Neumann series is derived and the computational complexity of the proposed method is discussed. Useful characteristics on the accuracy and the advantages of the proposed method over the exact state-space method, as well as over the common approximate procedure of ignoring the modal coupling, are shown and discussed in terms of some case studies. It is shown that the complex eigenpairs can be calculated by simply postprocessing of undamped eigenpairs.

Keywords: Complex modal analysis; Eigensolution; Undamped modes; Non-classically damped systems; Modal coupling; Frequency response function

Introduction

The equation of motion of an N DOF linear viscously damped system can be given by

$$\mathbf{M}\ddot{\mathbf{q}}(t) + \mathbf{C}\dot{\mathbf{q}}(t) + \mathbf{K}\mathbf{q}(t) = \mathbf{f}(t) \tag{1}$$

where $\mathbf{M}$, $\mathbf{C}$ and $\mathbf{K}$ are real mass, damping and stiffness matrices, respectively; $\mathbf{q}(t)$ and $\mathbf{f}(t)$ are displacement vector and force vector, respectively. In this paper, assume that $\mathbf{M}$ is a positive definite symmetric matrix, $\mathbf{K}$ and $\mathbf{C}$ are non-negative definite symmetric matrices. The inclusion of the influence of damping in structural and mechanical systems is extremely important if a model is to be applied in predicting vibration levels, transient responses, transmissibility and design problems dominated by energy dissipation. The eigenvalue problem of the viscously damped system can be written in matrix form as

$$\left(\lambda_j^2\mathbf{M} + \lambda_j\mathbf{C} + \mathbf{K}\right)\boldsymbol{\varphi}_j = \mathbf{0} \quad \forall j = 1,\ 2,\ \ldots, 2N \tag{2}$$

Here λ_j and $\boldsymbol{\varphi}_j$ denote the jth eigenvalue and eigenvector. Although several efficient approximation techniques were suggested, the most common is so-called the proportional approximation method (PAM), which is simply to ignore the mode coupling by using undamped modal shapes. The undamped mode shapes (normal modes) can be obtained by solving the undamped eigenproblem $\mathbf{K}\mathbf{u}_j = \omega_j^2\mathbf{M}\mathbf{u}_j$ where ω_j is the jth undamped frequency and in order of ascent; $\mathbf{u}_j$ denotes the mode shape corresponding to the jth frequency ω_j The PAM may be the most efficient approximate method, but the results of the PAM are not always with acceptable accuracy. It was shown by many studies (Hasselman, 1976; Warburton and Soni, 1977; Gawronski and Sawicki, 1997; Udwadia, 2009) that the light damping, the diagonal dominance of the transformed damping matrix

and the good separation property of normal modes are not any more the sufficient conditions for the accuracy of the proportional approximation method (these conditions were once believed to produce small errors for the PAM). When the non-proportional part of the damping is local, some method developed by (Özgüven, 1987; Denoël and Degée, 2009) can be used to efficiently calculate the frequency responses of a non-classically damped system in terms of the undamped modes. The complex modal analysis may be also a good choice to accurately calculate the frequency response (Adhikari, 2013). If only the lower modes are available, the frequency responses can be efficiently calculated by the method developed by (Li et al., 2014b)(Li et al., 2014a). In addition, complex modes can be used to transform any viscously damped system with *N* DOF into *N* independent second-order equations [see, e.g., (Kawano et al., 2013; Morzfeld et al., 2011; Ma et al., 2010) for details]. Note that in the dynamic response analysis, the primary computational effort is spent on the solution of the complex modes of the eigenproblem (2).

Real eigensolution techniques can be easily extended to handle the damped eigenproblem in terms of 2*N*-space (state-space) formulation, where *N* is the system dimension [see e.g., (Veletsos and Ventura, 1986) for details]. Although these state-space methods are exact in nature, they usually need heavy computational cost in practice due to the double size of system matrices. The state-space methods are not only computationally expensive, but also lack the physical insight provided by the superposition of the complex modes in the original physical space. To avoid the disadvantages mentioned previously, some efficiently computational methods in the original space were developed to compute the complex modes [see, e.g., (Kwak, 1993; Adhikari, 2011; Fischer, 2000; Holz et al., 2004; Rajakumar, 1993; Lee et al., 1998)].

In this paper, based on undamped modes, an efficient method is presented to calculate the complex eigenpairs by combining the Neumann series and the reduced basis technique. To avoid the modal truncation problem, the reduced basis is calculated by using a Neumann series expansion and only requires the undamped eigenpair of interest (i.e., it is not necessary to calculate all the undamped eigenpairs to hold the accuracy of engineering required). The sufficient condition for the Neumann series is derived and the computational complexity of the proposed method is discussed.

Solution of complex eigenvalue problem using classical normal modes

Premultiplying the damped eigenproblem by using $(\lambda_j^2 \mathbf{M})^{-1}$ yields

$$\left(\mathbf{I}_N + \kappa_j \mathbf{M}^{-1}\mathbf{C} + \kappa_j^2 \mathbf{M}^{-1}\mathbf{K}\right)\boldsymbol{\varphi}_j = \mathbf{0} \text{ with } \kappa_j = 1/\lambda_j \qquad (3)$$

Here $\mathbf{I}_N$ denotes the identity matrix of size *N*. The previous equation can be rewritten as

$$\left(\mathbf{I}_N + \kappa_j \mathbf{M}^{-1}\mathbf{C}\right)\boldsymbol{\varphi}_j = -\kappa_j^2 \mathbf{M}^{-1}\mathbf{K}\boldsymbol{\varphi}_j \qquad (4)$$

By using the Neumann series expansion, one obtains

$$\boldsymbol{\varphi}_j = -\kappa_j^2 \sum_{k=0}^{\infty} \left(-\kappa_j \mathbf{M}^{-1}\mathbf{C} \right)^k \mathbf{M}^{-1}\mathbf{K}\boldsymbol{\varphi}_j \tag{5}$$

On the condition that the mode shapes are not changed significantly, it may be convenient to approximate the eigenvectors using undamped mode shapes

$$\boldsymbol{\varphi}_j \approx -\sum_{k=0}^{\infty} \left(-\kappa_j \right)^{k+2} \mathbf{r}_k^{(j)} \tag{6}$$

where

$$\mathbf{r}_0^{(j)} = \mathbf{M}^{-1}\mathbf{K}\mathbf{u}_j \text{ and } \mathbf{r}_k^{(j)} = \mathbf{M}^{-1}\mathbf{C}\mathbf{r}_{k-1}^{(j)} \qquad \forall\, k \geq 1 \tag{7}$$

In view of $\mathbf{M}^{-1}\mathbf{K}\mathbf{u}_j = \omega_j^2 \mathbf{u}_j$, Eq. (7) can be further simplified as

$$\mathbf{r}_0^{(j)} = \omega_j^2 \mathbf{u}_j \text{ and } \mathbf{r}_k^{(j)} = \mathbf{M}^{-1}\mathbf{C}\mathbf{r}_{k-1}^{(j)} \qquad \forall\, k \geq 1 \tag{8}$$

The basis vectors can be efficiently calculated using the matrix decomposition of the mass matrix $\mathbf{M}$, which only needs to be obtained once for different eigenpairs. Once the matrix decomposition is available, the calculation of the series vectors by Eq. (8) involves only forward and backward substitutions, which are trivial. However, the coefficient of each basis vector is unknown.

In the reduced basis technique, the approximate result is searched within a subspace spanned by using reduced basis vectors. We define a reduced subspace $\mathbf{R}_j$ for each complex eigenpair in terms of the first r terms of the Neumann series expansion. That is

$$\mathbf{R}_j = \text{span}\left\{ \mathbf{r}_0^{(j)}, \quad \mathbf{r}_1^{(j)}, \quad \ldots, \quad \mathbf{r}_{r-1}^{(j)} \right\} \in \mathbb{R}^{N \times r} \tag{9}$$

One obtains the approximate complex eigenvectors as

$$\boldsymbol{\varphi}_j \approx \sum_{k=0}^{r-1} \alpha_k^{(j)} \mathbf{r}_k^{(j)} = \mathbf{R}_j \boldsymbol{\alpha}_j \text{ where } \boldsymbol{\alpha}_j = \left\{ \alpha_1^{(j)}, \quad \alpha_2^{(j)}, \quad \ldots, \quad \alpha_{r-1}^{(j)} \right\}^T \in \mathbb{R}^{r \times 1} \tag{10}$$

Since the approximate eigenvectors should satisfy the eigenproblem (2), we determine these unknown coefficients $\boldsymbol{\alpha}_j$ by substituting the approximate eigenvectors back into Eq. (2) and pre-multiplying it using $\mathbf{R}_j^T$, that is,

$$\left(\lambda_j^2 \mathbf{M}_R + \lambda_j \mathbf{C}_R + \mathbf{K}_R \right) \boldsymbol{\alpha}_j = \mathbf{0} \tag{11}$$

with

$$\mathbf{M}_R = \mathbf{R}_j^T \mathbf{M}\mathbf{R}_j \in \mathbb{R}^{r \times r}, \; \mathbf{C}_R = \mathbf{R}_j^T \mathbf{C}\mathbf{R}_j \in \mathbb{R}^{r \times r} \text{ and } \mathbf{K}_R = \mathbf{R}_j^T \mathbf{K}\mathbf{R}_j \in \mathbb{R}^{r \times r} \tag{12}$$

The robustness of the rectangular basis matrix $\mathbf{R}_j$ can be improved by using a Gram–Schmidt orthonormalization procedure. The ($r \times r$) reduced system matrices are dense matrices, but they are symmetric and much smaller in size than the ($N \times N$) original system matrices. On the condition that the mode shapes of the systems do not have significantly changed, the approximate modified eigenvalue λ_j can be chosen by finding the eigenvalue

of the reduced eigenproblem (11) such that the absolute largest value of $\omega_j - \lambda_j$ is minimized. Next, based on the frequency shifting technique, we give a method to obtain the eigenvalue λ_j. The reduced eigenproblem by Eq. (11) can be reformed as

$$\left(\tilde{\lambda}_j^2 \mathbf{M}_R + \tilde{\lambda}_j \overline{\mathbf{C}}_R + \overline{\mathbf{K}}_R \right) \boldsymbol{\alpha}_j = \mathbf{0} \text{ with } \tilde{\lambda}_j = \lambda_j - \omega_j \tag{13}$$

where the equivalent damping and stiffness matrices are

$$\overline{\mathbf{C}}_R = \left(\mathbf{C}_R + 2\omega_j \mathbf{M} \right) \text{ and } \overline{\mathbf{K}}_R = \left(\mathbf{K}_R + \omega_j \mathbf{C}_R + \omega_j^2 \mathbf{M}_R \right) \tag{14}$$

Once the first-order mode of the reduced eigenproblem (13) is solved. The approximate modified eigenvalue λ_j can be calculated. Once the vector $\boldsymbol{\alpha}_j$ is calculated, the modified eigenvectors can be obtained using Eq. (10). It is interesting to note that the first reduced basis vector is the undamped mode shape in nature. If the complex eigenpair can be calculated to satisfy suitable accuracy requirements by using one reduced basis vector, it means the system is close to the classically damped system and the PAM may produce small errors. If the complex eigenpairs need more than one reduced basis vector, under such circumstance, the PAM will produce unexpected errors and the proposed method can be used for suitable accuracy requirements.

In view of Eq. (2), the number r in Eq. (10) can be then determined if

$$e(\mathrm{i}\omega) = \left\| \left(\lambda_j^2 \mathbf{M} + \lambda_j \mathbf{C} + \mathbf{K} \right) \boldsymbol{\varphi}_j \right\|_2 < \varepsilon \tag{15}$$

where the parameter ε is a given accuracy for the absolute error. Alternatively, the number r in Eq. (10) may be determined by

$$\theta < \varepsilon_\theta \quad (\text{here } \cos\theta = \frac{\left| \mathbf{r}_r^H \mathbf{r}_{r-1} \right|}{\sqrt{\left(\mathbf{r}_{r-1}^H \mathbf{r}_{r-1} \right) \left(\mathbf{r}_r^H \mathbf{r}_r \right)}}) \tag{16}$$

Here the parameter ε_θ is a given accuracy. It means that increasing the reduced basis vectors can be stoped if the reduced basis vectors become linearly-dependent.

Computational considerations

For the convergent Neumann series expansion, one obtains the necessary and sufficient condition $\rho(\kappa_j \mathbf{M}^{-1}\mathbf{C}) < 1$. Here $\rho(\bullet)$ denotes the spectral radius of matrix $(\bullet)$. It means that all the eigenvalues of matrix $\kappa_j \mathbf{M}^{-1}\mathbf{C}$ have absolute values less than one. The maximal eigenvalues of matrix $\mathbf{M}^{-1}\mathbf{C}$ can be found by solving the minimal eigenvalues of $\mathbf{M}\mathbf{y}_j = s_j \mathbf{C}\mathbf{y}_j$. Once the minimal eigenvalue $s_{\min}$ is solved, the convergence condition can be given by

$$\left| s_{\min} \kappa_j \right| < 1 \quad \text{or} \quad \left| s_{\min} \right| < \left| \lambda_j \right| \tag{17}$$

which can be approximated determined using the known undamped frequencies

$$\left| s_{\min} \right| < \left| \omega_j \right| \tag{18}$$

When the size N is large, the operation count to solve the undamped eigenproblem is $O(N^3)$ (Adhikari, 2013). The operation count for the $\mathbf{LDL}^T$ decomposition of the mass matrix $\mathbf{M}$ is $0.5Nb^2$ (Bathe, 1996), where b is the semi-bandwidth of the mass matrix. The process of forward and backward substitutions for these basis vectors given by Eq. (7) is $2LrNb$. Since the number of the basis vectors $r \ll N$ and the number of the calculated complex modes $L \ll N$, the other operation count is trivial. Therefore, the operation count of the proposed method is $O(N^3+0.5Nb^2+2LrNb)$. Since the half-bandwidth b is roughly proportional to $N^{0.5}$ (Bathe, 1996), the flop can be simplified as $O(N^3+0.5N^2+2LrN^{1.5})$ for the consistent mass matrix and $O(N^3+2LrN^{1.5})$ for the lumped mass matrix. For non-classically damped systems, the state-space method (the size of the state-space matrix is $2N$) to calculate the complex eigenproblem is $O(8N^3)$ (Adhikari, 2013). Therefore, the proposed method shows a clear advantage over the state-space method in engineering applications as $O(8N^3) > O(N^3+0.5N^2+2LrN^{1.5})$ in the case the consistent mass matrix and $O(8N^3) > O(N^3+2LrN^{1.5})$ in the case of the lumped mass matrix.

Examples and discussions

Example 1: Three-DOF viscously damped system. A three DOF viscously damped system is considered here. The mass, damping and stiffness matrices are

$$\mathbf{M} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix}, \mathbf{C} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0.175 & -0.175 \\ 0 & -0.175 & 0.175 \end{pmatrix}, \mathbf{K} = \begin{pmatrix} 4 & -2 & 0 \\ -2 & 4 & -2 \\ 0 & -2 & 4 \end{pmatrix} \quad (19)$$

To illustrate the accuracy of the proposed method, two cases for distinct damping matrices are considered.

Case 1: consider the damping matrix given by Adhikari (Adhikari, 2011). In this case, the system does not satisfy the convergence condition for the Neumann series expansion and the assumption condition that the mode shapes are not changed significantly. **Table 1** lists the undamped frequencies and the complex eigenvalues using the exact state-space method and the proposed method for case 1. The damping coefficient of non-classically damped systems can be defined as (McLean, 2010)

$$g_j = -2\alpha_j / |\omega_{dj}| \quad (\text{here } \lambda_j = \alpha_j + i\omega_{dj})$$

which is popular in the aeroelastic filed. The damping coefficient g_j is approximately twice times the value of the conventional modal damping ratio. The MAC between the undamped mode shapes and exact complex mode shapes is 0.9953 for the first mode, 0.7734 for the second mode and 0.6587 for the third mode (the MAC value close to unity denotes the similarity whereas it close to zero denotes no similarity). As can be seen, the MAC of the last two modes is not much similarity in comparison with that of the first mode and the last two undamped frequencies are close. Therefore the system shows a significant modal coupling. The high modal coupling means that the system does not satisfy the assumption condition that the mode shapes are not changed significantly. The minimal eigenvalue of $\mathbf{M}y_j = s_j\mathbf{C}y_j$ is 0.8571 ($s_{\min}=0.8571$). It means that the convergence

condition for the Neumann series expansion is not satisfied. Although the convergence problem and high modal coupling exist in the special case, the proposed method with two basis vectors shows a good accuracy (the maximal error is 0.7469% and the minimum MAC is 0.9822). Therefore, although the convergence condition given by (17) is the necessary and sufficient condition for the Neumann series expansion, it may be only a sufficient condition for the proposed method. In addition, the condition, which assumes the mode shapes are not changed significantly, maybe not a necessary condition for the proposed method. Now, we use these calculated modes to calculate the frequency response functions (FRFs). For the sake of comparison, the FRFs are also calculated using the direct frequency response method (DFRM) and the PAM. The DFRM, which requires a matrix decomposition of the dynamic stiffness matrix at each excitation frequency, is considered as an exact result. Two typical FRFs are shown in **Figure 1**. The damping coefficient is 0.0328 for the first mode, 0.0771 for the second mode and 0.7546 for the third mode. Only two resonance peaks are visible since the damping coefficient of the third mode is high. It is shown that an unacceptable error is produced in the FRFs obtained by the PAM. However, the proposed method improves the results.

Table 1. Eigenvalues using state-space method and the proposed method for case 1.

Mode	Complex eigenvalue (exact)	Undamped frequency	One basis vector (r=1)			Two basis vectors (r=2)		
			Eigenvalues	Error (%)	MAC	Eigenvalues	Error (%)	MAC
1	-0.0103 + 0.6298i	0.6249	-0.0125 + 0.6248i	0.8734	0.9953	-0.0107 + 0.6295i	0.0715	0.9995
2	-0.0478 + 1.2407i	1.1547	-0.1458 + 1.1455i	11.0078	0.7734	-0.0406 + 1.2350i	0.7469	0.9822
3	-0.5252 + 1.2890i	1.5087	-0.4250 + 1.4476i	13.4770	0.6587	-0.5250 + 1.2825i	0.4708	0.9917
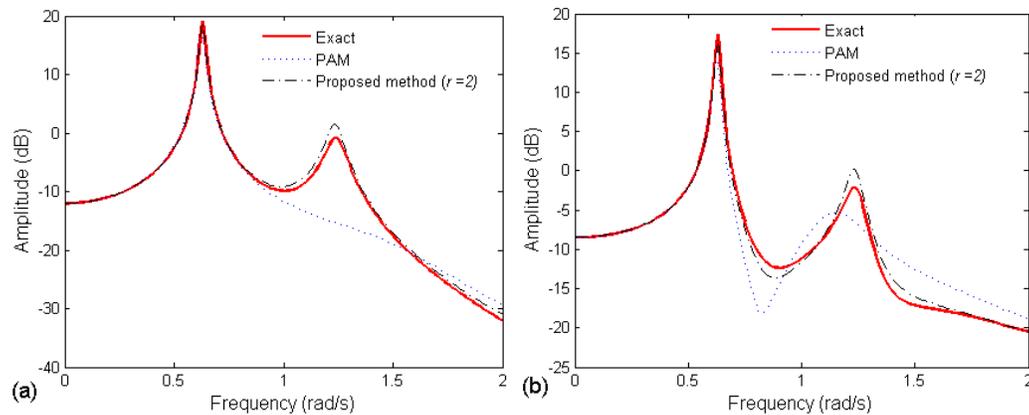
Figure 1. FRFs for case 1. (a) The FRF excited at the first DOF and measured at the second DOF. (b) The FRF excited at the three DOF and measured at the three DOF.

Case 2: the damping matrix is considered as Eq. (19). In this case, the system satisfies the assumption condition that the mode shapes are not changed significantly but does not satisfy the convergence condition for the Neumann series expansion. **Table 2** lists the eigenvalues for case 2. The MAC of the undamped mode shapes and exact complex mode shapes is 0.9999 for the first mode, 0.9962 for the second mode and 0.9935 for the third mode. These MAC values close to unity show the system satisfy the assumption condition that the mode shapes are not changed significantly. The minimal eigenvalue of $\mathbf{M}y_j = s_j\mathbf{C}y_j$ is 8.5714 (s_{min}= 8.5714), which means that the system does not satisfy the

convergence condition for the Neumann series expansion. It is shown that, although the system does not satisfy the convergence condition in the special case, the proposed method with one basis vectors shows a good accuracy (the maximum error is 0.1378 and the minimum MAC is 0.9935) and the results of the proposed method with two basis vectors almost coincides with the exact results. As can be seen, the assumption condition can significantly affect the accuracy of the proposed method. In this case, two typical FRFs are shown in **Figure 2**. The damping coefficient is 0.0040 for the first mode, 0.0251 for the second mode and 0.0565 for the third mode. It is shown that, although the damping coefficient is relatively light, an unacceptable error is also produced in the FRFs obtained by the PAM. However, the proposed method improves the results and its result shows a good agreement with that of the DFRM.

Table 2. Eigenvalues using state-space method and the proposed method for case 2.

Mode	Complex eigenvalue (exact)	One basis vector (r=1)			Two basis vectors (r=2)		
		Eigenvalues	Error (%)	MAC	Eigenvalues	Error (%)	MAC
1	-0.0012 + 0.6250i	-0.0013 + 0.6249i	0.0095	0.9999	-0.0012 + 0.6250i	0.0008	1.0000
2	-0.0145 + 1.1561i	-0.0146 + 1.1546i	0.1281	0.9962	-0.0145 + 1.1563i	0.0161	0.9995
3	-0.0426 + 1.5060i	-0.0425 + 1.5081i	0.1378	0.9935	-0.0425 + 1.5061i	0.0049	0.9999
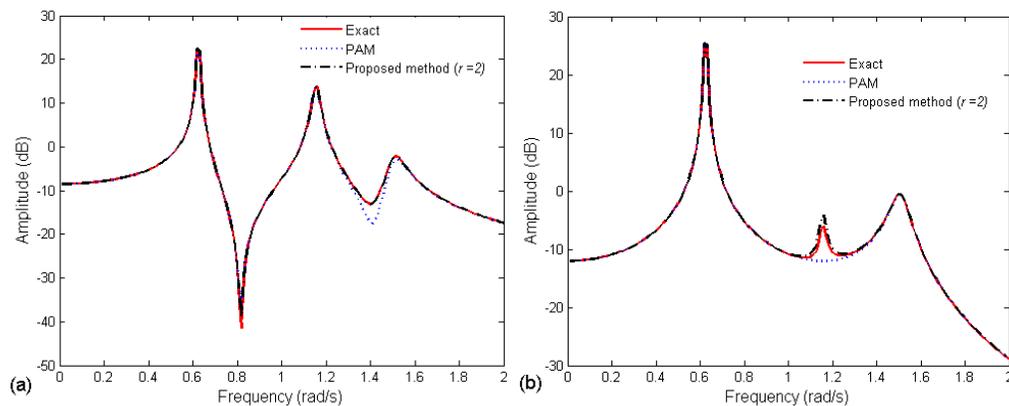
Figure 2. FRFs for case 2. (a) The FRF excited at the first DOF and measured at the first DOF. (b) The FRF excited at the first DOF and measured at the second DOF.
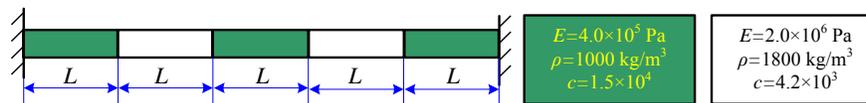
Figure 3. A damped truss structure.

Example 2: Damped truss structure. A simple but representative truss structure, shown in **Fig. 3**, is considered. For an element e, the elementary matrices are defined as follows:

$$\mathbf{M}_e = \frac{A\rho l_e}{2}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \mathbf{K}_e = \frac{EA}{l_e}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \text{ and } \mathbf{C}_e = \frac{cA}{l_e}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$$

where A is the cross-section area, ρ is the density, E is the elastic modulus, c is the damping coefficient and l_e is the length of the truss element (l_e=0.5 m). here L=10 m and

A=0.0001 m^2. The DOF of the truss structure is 99. **Figure 4** shows the first DOF of the FRF excited at the first DOF and measured at the first DOF. The complex modal superposition method should be used to accurately calculate the FRF since the results of the PAM do not give an acceptable accuracy. **Table 3** lists the undamped frequencies and the complex eigenvalues using the exact state-space method and the proposed method. Two typical mode shapes calculated by using the exact state-space method and the proposed method with three basis vectors are shown in **Figure 5**. It is shown that the proposed method shows a good agreement with the exact results. The computational time of obtaining the first five complex eigenpairs is 3.2396e-2 seconds for the proposed method and 4.8613e-2 seconds for the state-space method. It means that the proposed method is faster than the state space method. If the DOF becomes larger, the time of the state-space method will increase rapidly and the proposed method will show a clear advantage over the state-space method as it discussed previously.
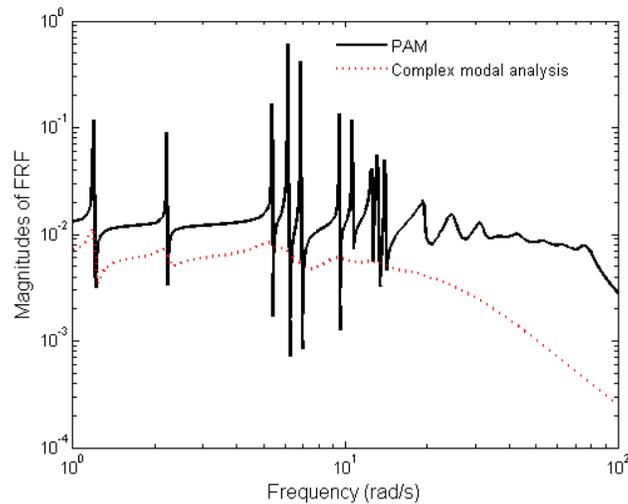
Figure 4. The FRF calculated by the PAM and the complex modal analysis.

Table 3. Eigenvalues using state-space method and the proposed method.

Mode	Complex eigenvalue (exact)	Undamped frequency	One basis vector (*r*=3)		Two basis vectors (*r*=4)	
			Eigenvalues	Error (%)	Eigenvalues	Error (%)
1	-2.4839e-2 +1.1883i	1.1885	-2.4836e-2 +1.1882i	5.2116e-3	-2.4836e-2 +1.1882i	5.1716e-3
2	-8.6932e-2 +2.2064i	2.2078	-8.6913e-2 +2.2061i	1.2450e-2	-8.6913e-2 +2.2061i	1.2273e-2
3	-3.5758e-1 +5.3648i	5.3479	-3.5588e-1 +5.3385i	4.8983e-1	-3.5592e-1 +5.3388i	4.8479e-1
4	-5.7556e-1 +6.1127i	6.1088	-5.7054e-1 +6.0848i	4.6164e-1	-5.7060e-1 +6.0851i	4.5686e-1
5	-7.9690e-1 +6.7949i	6.8224	-7.9366e-1 +6.7809i	2.1052e-1	-7.9375e-1 +6.7812i	2.0528e-1
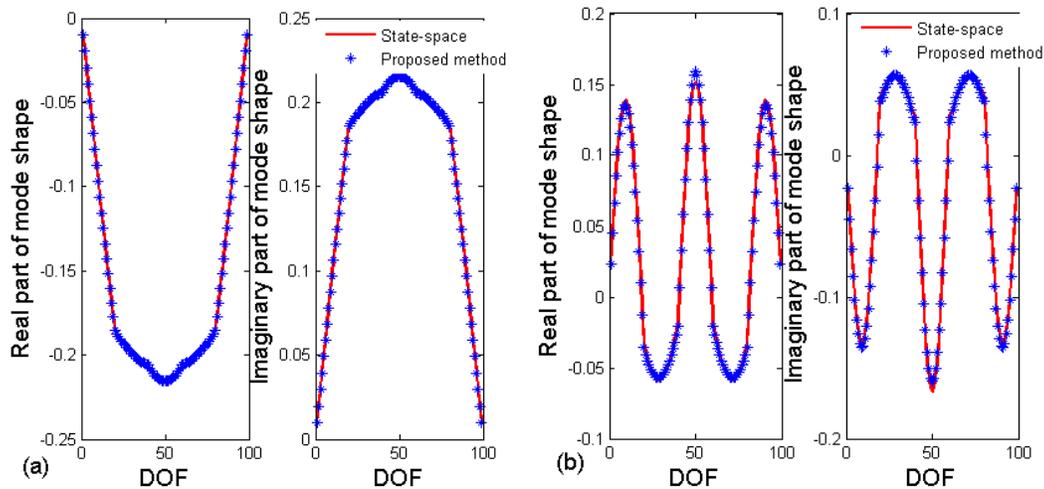
Figure 5. Complex mode shape. (a) First-order mode. (b) Fifth-order mode.

Conclusions

Based on undamped modes, an efficient method is presented to calculate complex eigenpairs by combining the Neumann series and the reduced basis technique. The method only requires the undamped eigenpair of interest. Some interest characteristic is shown and discussed in terms of numerical examples. It is shown that the complex eigenpairs can be calculated by simply postprocessing of undamped eigenpairs. Note that the first reduced basis vector is the undamped mode shape in nature. If complex eigenpair can be calculated by using one reduced basis vector to satisfy suitable accuracy, it means the system is close to the classically damped system and the proportional approximation method (PAM) may produce small errors. Therefore, it can be concluded that the proposed method can be reduced to the PAM when only the first reduced basis vector is used. If complex eigenpairs need more than one reduced basis vector to satisfy suitable accuracy requirements, under such circumstance, the PAM will produce unexpected errors and the proposed method can be used for suitable accuracy. Therefore, the proposed method is also developed as a criterion for choosing the PAM.

Acknowledgments

References

Adhikari S. (2011) An iterative approach for nonproportionally damped systems, *Mechanics Research Communications* **38**, 226-230.

Adhikari S. (2013) *Structural Dynamic Analysis with Generalized Damping Models: Analysis*, John Wiley & Sons.

Bathe K-J. (1996) Finite element procedures. New Jersey: Prentice Hall.

Denoël V and Degée H. (2009) Asymptotic expansion of slightly coupled modal dynamic transfer functions, *Journal of Sound and Vibration* **328**, 1-8.

Fischer P. (2000) Eigensolution of nonclassically damped structures by complex subspace iteration, *Computer Methods in Applied Mechanics and Engineering* **189**, 149-166.

Gawronski W and Sawicki JT. (1997) Response errors of non-proportionally lightly damped structures, *Journal of Sound and Vibration* **200**, 543-550.

Hasselman T. (1976) Modal coupling in lightly damped structures, *AIAA Journal* **14**, 1627-1628.

Holz UB, Golub GH and Law KH. (2004) A subspace approximation method for the quadratic eigenvalue problem, *SIAM Journal on Matrix Analysis and Applications* **26**, 498-521.

Kawano DT, Morzfeld M and Ma F. (2013) The decoupling of second-order linear systems with a singular mass matrix, *Journal of Sound and Vibration* **332**, 6829-6846.

Kwak M. (1993) Perturbation method for the eigenvalue problem of lightly damped systems, *Journal of Sound and Vibration* **160**, 351-357.

Lee I-W, Kim M-C and Robinson A. (1998) Efficient solution method of eigenproblems for damped structural systems using modified Newton-Raphson technique, *Journal of Engineering Mechanics* **124**, 576-580.

Li L, Hu YJ and Wang XL. (2014a) Accurate method for harmonic responses of non-classically damped systems in the middle frequency range, *Journal of Vibration and Control* **in press**.

Li L, Hu YJ, Wang XL and Lü L. (2014b) A hybrid expansion method for frequency response functions of non-proportionally damped systems, *Mechanical Systems and Signal Processing* **42**, 31-41.

Ma F, Morzfeld M and Imam A. (2010) The decoupling of damped linear systems in free or forced vibration, *Journal of Sound and Vibration* **329**, 3182-3202.

McLean DM. (2010) *MD Nastran Dynamic Analysis User's Guide*. Santa Ana: MSC.Software Corporation.

Morzfeld M, Ma F and Parlett BN. (2011) The transformation of second-order linear systems into independent equations, *SIAM Journal on Applied Mathematics* **71**, 1026-1043.

Özgüven H. (1987) A new method for harmonic response of non-proportionally damped structures using undamped modal data, *Journal of Sound and Vibration* **117**, 313-328.

Rajakumar C. (1993) Lanczos algorithm for the quadratic eigenvalue problem in engineering applications, *Computer Methods in Applied Mechanics and Engineering* **105**, 1-22.

Udwadia FE. (2009) A note on nonproportional damping, *Journal of Engineering Mechanics* **135**, 1248-1256.

Veletsos AS and Ventura CE. (1986) Modal analysis of non‐classically damped linear systems, *Earthquake Engineering & Structural Dynamics* **14**, 217-243.

Warburton G and Soni S. (1977) Errors in response calculations for non-classically damped structures, *Earthquake Engineering & Structural Dynamics* **5**, 365-376.

Homogenization for composite material properties using smoothed finite element method

Eric Li[1*], Zhongpu Zhang[1], CC Chang[1], GR Liu[2], Q Li[1]

[1]*School of Aerospace, Mechanical and Mechatronic Engineering, The University of Sydney, Sydney, NSW 2006, Australia*

[2]*School of Aerospace Systems, University of Cincinnati, Cincinnati, OH 45221-0070, USA*

Abstract

Numerical homogenization is an efficient way to determine effective material properties of composite materials. Conventionally, the finite element technique has been widely used in implementing the homogenization. However, the standard finite element method (FEM) leads to an overly-stiff model which gives poor accuracy especially using triangular elements in 2D or tetrahedral elements in 3D with coarse mesh. In this paper, the smoothed finite element methods (S-FEMs) are developed to analyse the effective mechanical properties of composite materials. Various examples, including modulus with multiphase composites and permeability of tissue scaffold, have demonstrated that smoothed finite element method is able to provide more accurate results using the same set of mesh compared with the standard finite element method. In addition, the computation efficiency of smoothed finite element method is also much better than the FEM counterpart.

Keywords: smoothed finite element methods, homogenization, composite material, Tissue Scaffold

---

[1*]Corresponding author.

Email address: ericsg2012@gmail.com

1. Introduction

Attributable to its more versatile and tuneable material properties, a range of composites has been widely used in aerospace, marine, vehicle and biomedical industry as shown in Fig. 1. Their different microstructures with two or more constituents allow achieving desirable properties such as multi-functionality and lightweight. To control the material properties, the spatial layout of the microstructure and/or the compositions of the constituent phases are extremely important [1]. However, it is always cumbersome to tweak a most appropriate microstructure and characterize its effective properties effectively [2].

Substantial research has been done in the evaluation of effective (or namely bulk) elastic properties of composite materials. An early attempt for the design of composite material is perhaps the bounds of material property. Using variational principles, Hashin and Shtrikman developed theoretical bounds for the bulk (K) and shear (G) moduli of two-phases, well-ordered materials [3]. Following this, some improvements of material bounds have been extended to multi-phase, multi-dimensional composites for various physical properties [4-5]. Although property bounds give the constraint of effective properties and provide some clue in choosing different phases for composite design, these all require additional information regarding the geometric layout of the microstructure [6].

Numerical homogenization is an effective way to quantify the material properties based on an asymptotic expansion of the governing equations [7-11]. In the numerical homogenization, it is assumed that the representative volume element (RVE) or unit cell is locally repeated with very small microstructure compared with the overall 'macroscopic' dimensions of the structure of interest, in which the different materials are bonded in the RVE. The homogenization method is based on a rigorous mathematical theory [12], and it can provide us with a reasonable solution for some material design problems, where experimental

techniques may be very costly or unavailable, apart from the determination of theoretical bounds for an estimation purpose [13].

Currently, FEM is the most popular method in numerical homogenization for composite material properties [14-18]. However, FEM has some inherent drawbacks which limit the development of numerical homogenization. The first issue is its "overly stiff" phenomenon of a fully compatible FEM model [19, 20]; the second concerns with the mesh distortion related problems such as the significant accuracy loss when the element mesh is heavily distorted; the third is the poor accuracy in the stress solution using triangular in 2D or tetrahedral elements in 3D.

Due to this reason, Liu and his co-authors have established a weakened weak (W2) formulation using the generalized gradient smoothing technique to unify all the developed numerical methods [21]. The W2 formulation seeks solutions in the so-called G space, which includes both continuous and discontinuous functions. Hence, it works for both compatible and incompatible displacements in the framework of the finite element and meshfree methods. Using the generalized strain smoothing technique, the Smoothed Point Interpolation Methods (S-PIMs) [22] have been developed, which offers a number of outstanding features. With the strain smoothing technique [23], the smoothed finite element methods (S-FEMs) have been proposed to overcome the shortcoming of FEM model [24-25]. The S-FEMs can be viewed as the simplest linear version of S-PIMs and has the advantages of simplicity and yet outstanding performance and important features.

Compared with overly-stiff FEM, S-FEMs provides a softened system model for high convergence and accuracy. According to smoothing algorithm, several different formulations, namely cell-based smoothed FEM (CS-FEM) [26-27], node-based smoothed FEM (NS-FEM) [28], and edged-based smoothed FEM (ES-FEM) [29-30], have been proposed respectively. These methods introduce the strain smoothing operations to the standard FEM procedures,

and worked fairly effectively for a range of engineering problems such as solid mechanics [31], heat transfer [32-35] and acoustics [36-37]. The implementation of such smoothed finite element methods is quite straightforward without additional parameters involved. The study of S-FEMs has also clearly shown that the smoothing operation on strains allows to manipulate the assumed strain field in a proper fashion to ensure the stability (boundness), thus ensuring the (monotonic) convergence, and giving the S-FEMs some very good features.

Lured by the special properties of smoothed finite element methods, this study for the first time attempted to formulate homogenization problem for different composite materials. The objective to develop S-FEMs for homogenization is to improve the numerical accuracy, computational efficiency, as well as to study the applicability of S-FEMs. The above mentioned various smoothed finite element techniques are formulated here to characterize the effective properties for two or more materials. This paper is organized as follows: Section 2 briefs numerical homogenization using finite element methods. Section 3 presents the formulation of smoothed finite element methods in numerical homogenization. The 2D and 3D demonstrative examples are shown in Section 4 to verify the applicability of smoothed finite element method in numerical homogenization. The conclusions are drawn in Section 5.

2. Brief of numerical homogenization using FEM

2.1 Concept of Periodic Representative Volume Element (RVE)

In this study, we consider periodic composites which comprise repetitive identical unit cells in the microscopic level of material structure. For the sake of simplicity, only two-dimensional problem is illustrated here as shown Fig. 2. To characterize one piece of material sample, the computational cost for a full finite element (FE) model can be extremely high as discretization of the whole sample solid becomes enormous in order to represent detailed structure of the microscopic material constitutions. Such issue will be more serious and could

become prohibitive in three-dimensional problems. Fortunately, the homogenization method provides an efficient way to predict the mechanical behaviour of macrostructure without modeling the entire macroscopic structure of multiphase composites [12].

In general, the selection of representative volume element (RVE) is crucial in the homogenization process in order to accurately predict the effective mechanical properties. The selective RVE must make sure repetitive unit identifiable in the domain carry all the geometric features necessary to fully define the medium [38]. Note that while the RVE is not uniquely defined in the homogenization; the effective mechanical properties from different unit cells should be ideally the same on the given scale. Additionally, when deciding the size of the RVE, the geometrical and material symmetries of the structure can be considered in order to simplify the implementation of numerical code [17].

2.2 Effective Elastic Moduli

For more effective discussion, we first brief on the standard formulation of FEM in numerical homogenization [14, 15], as some of these formulae will be used repetitively in later sections. In the homogenization, two levels of coordinate systems are used: one is the global coordinate system of macrostructure y_i and another is the local coordinate system of microstructure x_i. In the following analysis, linear elastic constitutive law is assumed. The relationship between the local coordinate system x_i for the RVE and global coordinate system of sample macrostructure y_i can be written as follows [14, 15]:

$$x_i = \frac{y_i}{\varepsilon} \tag{1}$$

where ε is the small scaling parameter between these two length scales.

Based on asymptotic expansion, it is reasonable to approximate $\mathbf{u}^\varepsilon$ in the following form with respect to parameter ε

$$\mathbf{u}^{\varepsilon}(y) = \mathbf{u}^{0}(x,y) + \varepsilon\mathbf{u}^{1}(x,y) + \varepsilon^{2}\mathbf{u}^{2}(x,y) + \ldots \tag{2}$$

where the function $\mathbf{u}_{i}^{0}, \mathbf{u}_{i}^{1}, \mathbf{u}_{i}^{2} \ldots$ are X-periodic with respect to the local coordinate x.

The strain-displacement and stress-strain relationships can be accordingly expressed as follows

$$\varepsilon(\mathbf{y}) = \frac{1}{2}\left(\frac{\partial\mathbf{u}_{k}^{\varepsilon}}{\partial y_{l}} + \frac{\partial\mathbf{u}_{l}^{\varepsilon}}{\partial y_{k}}\right) \tag{3}$$

$$\sigma(\mathbf{u}^{\varepsilon}) = \mathbf{D}\varepsilon(\mathbf{u}^{\varepsilon}) \tag{4}$$

where $\mathbf{D}$, ε, $\mathbf{u}^{\varepsilon}$ σ are the elasticity matrix for base material, strain, displacement and stress.

As $\mathbf{u}_{1}$ is the first order variation from the average displacement, this variation can be considered to be proportional to the average strain ε_{0} [14]:

$$\mathbf{u}_{1} = -\xi(\mathbf{x})\varepsilon_{0}(\mathbf{y}) \tag{5}$$

where ξ is the characteristic displacement function of the microstructure. In other words, the characteristic displacement is scaled directly by the average train because it is the displacement for the unit strain of the macrostructure.

From Eqs. (3)-(5), the total potential energy is formulated as follows [14]:

$$\Gamma\Gamma(\mathbf{u}^{\varepsilon}) = \frac{1}{|X|}\int_{D}(\mathbf{L}_{dy}\mathbf{u}_{0})^{\mathrm{T}}\int_{X}(\mathbf{I}-\mathbf{L}_{dx}\xi)^{\mathrm{T}}\mathbf{D}(\mathbf{I}-\mathbf{L}_{dx}\xi)(\mathbf{L}_{dy}\mathbf{u}_{0})dX(\mathbf{L}_{dy}\mathbf{u}_{0})dD$$
$$- \int_{\Gamma_{T}}\mathbf{u}_{0}^{t}\mathbf{t}d\Gamma \tag{6}$$

where $|X|$ is the area (volume in 3D) of RVE, $\mathbf{I}$ is the identity, and $\mathbf{L}_{d}$ is a matrix of differential operator defined as:

$$\mathbf{L}_{dx} = \begin{vmatrix} \partial/\partial\mathbf{x}_{1} & 0 \\ 0 & \partial/\partial\mathbf{x}_{2} \\ \partial/\partial\mathbf{x}_{2} & \partial/\partial\mathbf{x}_{1} \end{vmatrix}, \qquad \mathbf{L}_{dy} = \begin{vmatrix} \partial/\partial\mathbf{y}_{1} & 0 \\ 0 & \partial/\partial\mathbf{y}_{2} \\ \partial/\partial\mathbf{y}_{2} & \partial/\partial\mathbf{y}_{1} \end{vmatrix} \tag{7}$$

Define

$$\mathbf{D}^{H} = \frac{1}{|X|} \int_{X} \left(\mathbf{I} - \mathbf{L}_{dx}\xi\right)^{\mathrm{T}} \mathbf{D}\left(\mathbf{I} - \mathbf{L}_{dy}\xi\right) dX \tag{8}$$

and substitute Eq. (8) into (6) leads to

$$\Gamma\Gamma\left(\mathbf{u}^{\varepsilon}\right) = \int_{D} \left(\mathbf{L}_{dy}\mathbf{u}_{0}\right)^{\mathrm{T}} \mathbf{D}^{H} \left(\mathbf{L}_{dy}\mathbf{u}_{0}\right) dD - \int_{\Gamma_{T}} \mathbf{u}_{0}^{t}\mathbf{t}d\Gamma \tag{9}$$

The homogenized elasticity matrix $\mathbf{D}^{H}$ in Eq. (8) is obtained [14] through discretization of unit cell using finite element technique. In the numerical implementations, the FEM divides the domain Ω into a number of elements, and the following trial functions are used:

$$\xi\left(\mathbf{x},\mathbf{d}\right) = \mathbf{N}_{i}\left(\mathbf{x}\right)\mathbf{d}_{i} \tag{10}$$

$$\mathbf{L}_{dx}\xi = \mathbf{B}_{i}\mathbf{d}_{i} \tag{11}$$

where $\mathbf{d}_{i}$ is the vector of nodal displacements, and $\mathbf{N}_{i}\left(\mathbf{x}\right)$ is a matrix of shape function in the microstructure. The strain matrix $\mathbf{B}_{i}$ is defined as follows:

$$\mathbf{B}_{i} = \begin{vmatrix} \partial\mathbf{N}_{i}/\partial\mathbf{x}_{1} & 0 \\ 0 & \partial\mathbf{N}_{i}/\partial\mathbf{x}_{2} \\ \partial\mathbf{N}_{i}/\partial\mathbf{x}_{1} & \partial\mathbf{N}_{i}/\partial\mathbf{x}_{2} \end{vmatrix} \qquad \text{2D} \tag{12}$$

$$\mathbf{B}_{i} = \begin{vmatrix} \partial\mathbf{N}_{i}/\partial\mathbf{x}_{1} & 0 & 0 \\ 0 & \partial\mathbf{N}_{i}/\partial\mathbf{x}_{2} & 0 \\ 0 & 0 & \partial\mathbf{N}_{i}/\partial\mathbf{x}_{3} \\ \partial\mathbf{N}_{i}/\partial\mathbf{x}_{1} & \partial\mathbf{N}_{i}/\partial\mathbf{x}_{2} & 0 \\ 0 & \partial\mathbf{N}_{i}/\partial\mathbf{x}_{2} & \partial\mathbf{N}_{i}/\partial\mathbf{x}_{3} \\ \partial\mathbf{N}_{i}/\partial\mathbf{x}_{1} & 0 & \partial\mathbf{N}_{i}/\partial\mathbf{x}_{3} \end{vmatrix} \qquad \text{3D} \tag{13}$$

Substitute Eq. (10) into Eq. (8), the effective mechanical properties can be written as follows:

$$\mathbf{D}^{H} = \frac{1}{|X|} \int_{X} \left(\mathbf{I} - \mathbf{B}_{i}\mathbf{d}_{i}\right)^{\mathrm{T}} \mathbf{D}\left(\mathbf{I} - \mathbf{B}_{i}\mathbf{d}_{i}\right) dX \tag{14}$$

In order to obtain an equilibrium state, one can set variational to zero,

$$\delta\mathbf{d}_{i} \left(\int_{X} \mathbf{B}_{i}\mathbf{D}dX - \int_{X} \mathbf{B}_{i}^{T}\mathbf{D}\mathbf{B}_{i}dX\mathbf{d}_{i}\right) = 0 \tag{15}$$

By substituting the approximations $\mathbf{d}_i$ into Equation (14) and invoking the arbitrariness of virtual nodal displacements, we have the standard discretized algebraic system equation:

$$\mathbf{Kd}^n = \mathbf{F}^n \tag{16}$$

where $\mathbf{K}$ is an analogized element stiffness matrix, and $\mathbf{F}^n$ is the nodal force vector which is equivalent to the initially applied strain field.

$$\mathbf{K} = \int_X \mathbf{B}_i^T \mathbf{D} \mathbf{B}_i dX_e \tag{17}$$

$$\mathbf{F}^n = \int_X \mathbf{B}_i \mathbf{D} dX_e \tag{18}$$

For a 2D elastic problem, $n=1, 2, 3$, for 3D elastic problem, $n=1\ldots 6$. For a 2D heat conduction or fluidic permeability problem, $n=1, 2$, for 3D heat conduction or fluidic permeability problem, $n=1, 2, 3$. The boundary conditions for each case are listed in Tables 1 and 2. The detailed formulation of numerical homogenization method is available in many resources [14, 15, 17, 18]. Although the above process is derived from elasticity problem, the effective fluidic permeability can be calculated in the same way.

3. Implementation of S-FEMs in numerical homogenization

In the formulation of S-FEMs homogenization, the critical step is that the smoothed strain, instead of compatible strain, is used. The pre-process of mesh and implementation of boundary conditions are exactly the same as the standard FEM based homogenization. Hence, the computation of a smoothed strain is crucial to formulate S-FEMs based homogenization. In this section, several different strain smoothness algorithms are presented for homogenization.

3.1 Edge-based smoothed finite element method (ES-FEM)

In the formulation of ES-FEM based homogenization, the first step is to construct the smoothing domain. The local smoothing domains are constructed with respect to the edge of

triangular elements such that $\Omega = \bigcup_{k=1}^{N_s} \Omega_k^s$ and $\Omega_i^s \cap \Omega_j^s = \varnothing$, $\forall i \neq j$, in which N_s is the number of smoothing domains and equals to the number of elemental edges in the scheme of ES-FEM. For the triangular elements in 2D or tetrahedral elements in 3D, the smoothing domain associated with edge k is created by connecting two endpoints of the edge to two centroids of the two adjacent elements as shown in Fig. 3(a). Extending the smoothing domain Ω_k^s in 3D problems, the sub-domain of the smoothing domain Ω_k^s for edge k located in the particular cell j can be obtained by connecting two end nodes of the edge to the centroids of the surface triangles and the centroid of cell j. The sub-smoothing-domain for edge k is one sixth region of this tetrahedral element.

With the edge-based smoothing technique, the smoothed strains can be computed using the compatible strains $\boldsymbol{\varepsilon} = \mathbf{L}\mathbf{u}$ from the following smoothing operation [39]:

$$\bar{\boldsymbol{\varepsilon}}_k = \frac{1}{A_k^s} \int_{\Gamma_k^s} \boldsymbol{\varepsilon}_k d\Gamma = \frac{1}{A_k^s} \int_{\Gamma_k^s} \mathbf{L}\mathbf{u} d\Gamma \tag{19}$$

where Γ_k^s is the boundary surface of the smoothing domain Ω_k^s, $A_k = \int_{\Omega_k^s} d\Omega$ is the area of the smoothing domain for edge k, and $\mathbf{u}$ is the displacement vector expressed in the following approximate form:

$$\mathbf{u} = \sum_{i=1}^{N} \mathbf{N}_i(\mathbf{x})\mathbf{d}_i = \mathbf{N}_s \mathbf{d} \tag{20}$$

where N is the number of field nodes per element and equals to 3 for the three-node triangular elements used in this work, $\mathbf{d}_i = \{d_{xi} \quad d_{yi}\}^{\mathrm{T}}$ is the nodal displacement vector, $\mathbf{d}$ is the vector with all the N nodal displacements in the element, and $\mathbf{N}_i$ is a matrix of shape functions.

Substituting Eq. (20) into (19) and applying the divergence theorem, the smoothed strains for the smoothing domain Ω_k^s can then be obtained as follows

$$\bar{\boldsymbol{\varepsilon}}\left(\mathbf{x}_k\right) = \sum_{i=1}^{M_k} \bar{\mathbf{B}}_i^{\text{ES-2D}}\left(\mathbf{x}_k\right)\mathbf{d}_i \tag{21}$$

where M_k is the total number of nodes containing the same edge i. For the inner edge, M_k is equal to 4, and M_k becomes 3 for boundary edge.

The smoothed strain matrix can be calculated numerically in the following way:

$$\bar{\mathbf{B}}^{\text{ES-2D}} = \frac{1}{A^{(k)}} \sum_{j=1}^{N_e^{(k)}} \frac{1}{3} A_e^{(j)} \mathbf{B}_j^{2D} \tag{22}$$

where $N_e^{(k)}$ is the number of elements around the edge k. For the boundary edge, no smoothing effect exists in the edge, hence $N_e^{(k)}=1$ for boundary edges. For all inner edges, there are only two elements sharing one edge, so $N_e^{(k)}=2$.

The smoothed strain in 2D ES-FEM can be very straightforward to extend to 3D tetrahedral elements:

$$\bar{\mathbf{B}}^{\text{ES-3D}} = \frac{1}{V^{(k)}} \sum_{j=1}^{N_e^{(k)}} \frac{1}{6} V_e^{(j)} \mathbf{B}_j^{3D} \tag{23}$$

where V_e and $\mathbf{B}_j$ are the volume and the compatible strain gradient matrix of the jth tetrahedral element around node k, respectively. $V^{(k)}$ is calculated by the following equation:

$$V^{(k)} = \sum_{j=1}^{N_e^{(k)}} \frac{1}{6} V_e^{(j)} \tag{24}$$

Based on the formulations of smoothed strain expressed in Eqs. (22) and (23), the smoothed stiffness and force matrix can be written in the following forms, respectively:

$$\bar{\mathbf{k}}_{ES} = \sum_{k \in N_e} \int_{\Omega^{(k)}} \left(\mathbf{B}_{ES}\right)^{T} \mathbf{D} \mathbf{B}_{ES} d\Omega \tag{25}$$

$$\bar{\mathbf{F}}_{ES} = \sum_{k \in N_e} \int_{\Omega^{(k)}} \left(\mathbf{B}_{ES}\right)^{T} \mathbf{D} d\Omega \tag{26}$$

For the multi-material formulation in ES-FEM, the process is similar to single material except the interface of different materials. As the material property is discontinuous along the interface, the associated smoothing domain will be separated into two regions as shown in Fig. 3(b), which is the same as it does along domain boundaries.

3.2 Node-based smoothed finite element method

Similar to ES-FEM, the smoothing domain is first constructed in the scheme of NS-FEM. For 2D problems with single material as shown in Fig. 3(a), the smoothing domain Ω_k^s for node k is constructed by connecting sequentially the mid-edge-points to the centroids of the surrounding triangles of node k. The smoothing domain can be easily extended to 3D problems, where the sub-domain of the smoothing domain for node k located in the particular cell j can be obtained by connecting the mid-edge-points, the centroids of the surface triangles and the centroid of cell j. Finding out other sub-domains located in cells which contain node k and the smoothing domain for node k can be constructed by uniting all the sub-domains. Hence, the smoothed strain in 2D using the node-based smoothing technique with triangular element is expressed as

$$\bar{\mathbf{B}}_{I}^{\text{NS-2D}} = \frac{1}{A^{(k)}} \sum_{j=1}^{N_e^{(k)}} \frac{1}{3} A_e^{(j)} \mathbf{B}_j \tag{27}$$

where $\mathbf{B}_j$ is the compatible strain computed by standard FEM, N_e is the number of elements surrounding the node k; A_e is the area the jth element around the node k.

The area $A^{(k)}$ is computed by:

$$A^{(k)} = \int_{\Omega^{(k)}} d\Omega = \frac{1}{3} \sum_{j=1}^{N_e^{(k)}} A_e^{(j)} \tag{28}$$

Note that with this formulation, only the area and usual "compatible" strain matrices $\mathbf{B}_j$ by Eq. (27) of triangular elements are needed to calculate the system stiffness matrix for the NS-FEM.

In 3D NS-FEM, the smoothed strain can be calculated in a similar way:

$$\overline{\mathbf{B}}_I^{\text{NS-3D}} = \frac{1}{V^{(k)}} \sum_{j=1}^{N_e^{(k)}} \frac{1}{4} V_e^{(j)} \mathbf{B}_j \tag{29}$$

where $\mathbf{B}_j$ is the compatible strain computed by standard FEM, the V_e is the volume of the jth tetrahedral element around the node k. The $V_{(k)}$ is computed by:

$$V^{(k)} = \int_{\Omega^{(k)}} d\Omega = \frac{1}{4} \sum_{j=1}^{N_e^{(k)}} V_e^{(j)} \tag{30}$$

Hence, in the NS-FEM formulation of numerical homogenization, the stiffness matrix and force matrix can be formulated respectively as:

$$\overline{\mathbf{k}}_{\text{NS}} = \sum_{k \in N_e} \int_{\Omega^{(k)}} \left(\mathbf{B}_{\text{NS}} \right)^{\text{T}} \mathbf{D} \mathbf{B}_{\text{NS}} d\Omega \tag{31}$$

$$\overline{\mathbf{F}}_{\text{NS}} = \sum_{k \in N_e} \int_{\Omega^{(k)}} \left(\mathbf{B}_{\text{NS}} \right)^{\text{T}} \mathbf{D} d\Omega \tag{32}$$

In the formulation of multi-material NS-FEM, the smoothing domain is also separated two parts along the interface of different material as shown in Fig. 3(b) as material properties are not continuous. That means the associated smoothing domain is not allowed to cross the boundary of each material. It is noted that there are still some smoothing effect at both sides of interface, but smoothing effect is weak compared with internal nodes.

3.3 Cell-based smoothed finite element method

In the cell-based smoothed finite element method (CS-FEM) for the homogenization, the quadrilateral elements are considered. The formulation of stiffness is computed based on the smoothing cells (SC) located inside the quadrilateral elements as shown in Fig. 5. In CS-FEM, the elements are subdivided into several smoothing cells, such as $\Omega^e = \Omega_1^e \cup \Omega_1^e \cup ... \Omega_{nc}^e$. If the number of SC of the elements equals 1, the CS-FEM solution has the same properties with those of standard FEM using the reduced integration [24]. When the number of smoothing cells is approaching infinity, the CS-FEM solution approaches to the solution of standard FEM. Based on our research experience, the numerical solution is always stable and accurate if the number of smoothing cells is equal to 4 [24].

Based on the smoothing theory, the smoothed strain in CS-FEM can be expressed in the following equation:

$$\bar{\varepsilon}_{CS} = \sum_I^n \bar{\mathbf{B}}_{CS}\left(\mathbf{x}_C\right)\mathbf{u}_I \tag{33}$$

where $\bar{\mathbf{B}}_{CS}$ is the smoothed strain matrix. For 2D case

$$\bar{\mathbf{B}}_{CS} = \begin{bmatrix} \bar{b}_{CS} & 0 \\ 0 & \bar{b}_{CS} \\ \bar{b}_{CS} & \bar{b}_{CS} \end{bmatrix} \tag{34}$$

where

$$\bar{b}_{CS} = \frac{1}{A_C}\int_{\Gamma_c} N_I\left(\mathbf{x}\right)n_k\left(\mathbf{x}\right)d\Gamma \tag{35}$$

If one Gaussian point is used for line integration along each segment of boundary Γ_i^C of Ω_C, the above integration equation can be transformed to its algebraic form

$$\bar{b}_{CS} = \sum_{i=1}^{M} N_I\left(\mathbf{x}_i^{\mathrm{GP}}\right) n_{ik}^{C} l_{ik}^{C} \tag{36}$$

where $\mathbf{x}_i^{\mathrm{GP}}$ is the midpoint (Gaussian point) of boundary segment of Γ_i^{C}, whose length and outward unit normal are denoted as l_i^{C} and n_i^{C}, respectively.

The smoothed element stiffness matrix can be obtained by assembly of those all of the smoothing cells of the element, i.e.

$$\overline{\mathbf{K}}_{\mathrm{CS}} = \sum_{C} \overline{\mathbf{B}}_{\mathrm{CS}}^{\mathrm{T}} \mathbf{D} \mathbf{B}_{\mathrm{CS}} A_{\mathrm{CS}} \tag{37}$$

$$\overline{\mathbf{F}}_{\mathrm{CS}} = \sum_{C} \overline{\mathbf{B}}_{\mathrm{CS}}^{\mathrm{T}} \mathbf{D} A_{\mathrm{CS}} \tag{38}$$

The smoothed $\overline{\mathbf{B}}_{\mathrm{CS}}$ matrices are constructed with integration over the boundary of the cell of the element.

3.4. Algorithm for S-FEMs based homogenization

Numerical procedures for computing the effective mechanical properties of composite materials using smoothed finite element method are summarized as follows:

1. Design Composite material

2. Determine the unit cell

3. Divide the domain into a set of elements and obtain information on nodes coordinates and element connectivity

4. Create the smoothing domain for each smoothed finite element method.

5. Loop over all the elements

 (a) Compute the compatible strain $\mathbf{B}$ of the element by Equation using standard finite element formulation and save it to process the smoothed strain.

(b) Evaluate the smoothed strain. In ES-FEM, apply Eq. (22) and (23) to compute the edge-based smoothed strain. In NS-FEM, use Eq. (27) and (29) to calculate the node-based smoothed strain. In CS-FEM, Eq. (34) is adopted to determine the cell-based smoothed strain.

6. Calculation of smoothed stiffness and smoothed force matrix

- For ES-FEM, smoothed stiffness and force in numerical homogenization use Eq. (25) and (26)

- For NS-FEM, smoothed stiffness and force in numerical homogenization use Eq. (31) and (32)

- For CS-FEM, smoothed stiffness and force in numerical homogenization use Eq. (37) and (38)

7. Implement symmetrical boundary conditions by referring to Table 1 and 2.

8. Solve homogenization equation $\overline{\mathbf{K}}\mathbf{d}^n = \overline{\mathbf{F}}^n$

9. Evaluate the homogenized (effective) mechanical properties based on Eq. (13).

10. For 2D elasticity problem, loop step 7 and 8 three times for different boundary conditions. For 3D elasticity problem, repeat the step 7 and 8 six times for different characterized cases.

4. Numerical examples

4.1 Benchmark example

In order to verify the S-FEMs formulation for homogenization, one benchmark example [9] is first studied. As shown in Fig. 6, unit square cell containing a 0.4×0.6 void and the solid phase material properties are $D_{11}=D_{22}=30$ and $D_{12}=D_{33}=10$ (all units are assumed to be consistent) [9]. This problem was solved by Bendsoe and Kikuchi [16].

The numerical solutions obtained from S-FEMs and FEM using the triangular (T3) and quadrilateral (Q4) elements (Fig. 7) are tabulated in Tables 3 and 4, respectively. For the purpose of comparison, the published results are presented in Table 5. Compared these three tables, it is seen the homogenization results obtained from S-FEMs agree very well with the published data.

4.2 Void material

Another example is material 1 with void as shown in Fig. 8. The Young modulus and Poisson's ratio of solid are E=5MPa, v=0.3. A plane stress problem is considered here. The discretized models using triangular (T3) and quadrilateral (Q4) elements are shown in Fig. 9.

Figure 10 shows the convergence of effective (homogenized) bulk modulus using different FEM methods. The effective bulk modulus is defined by [40]:

$$Bu = \frac{E_{eff}}{2(1-v_{eff})} = \frac{E_{eff}}{4(1-v_{eff}^2)}\left(2+2v_{eff}\right) = \frac{1}{4}\left(D_{11}^{eff} + 2D_{12}^{eff} + D_{22}^{eff}\right) \tag{39}$$

The effective elasticity tensor $\mathbf{D}^{eff}$ can be defined as follows:

2D plan stress problem

$$\mathbf{D}^{eff} = \frac{E_{eff}}{1-v_{eff}^2}\begin{vmatrix} 1 & v_{eff} & 0 \\ v_{eff} & 1 & 0 \\ 0 & 0 & \left(1-v_{eff}\right)/2 \end{vmatrix} = \begin{vmatrix} D_{1111}^{eff} & D_{1122}^{eff} & 0 \\ D_{1122}^{eff} & D_{2222}^{eff} & 0 \\ 0 & 0 & D_{1212}^{eff} \end{vmatrix} \tag{40}$$

2D plan stain problem

$$\mathbf{D}^{eff} = \frac{E_{eff}}{\left(1+v_{eff}\right)\times\left(1-2v_{eff}\right)}\begin{vmatrix} 1-v_{eff} & v_{eff} & 0 \\ v_{eff} & 1-v_{eff} & 0 \\ 0 & 0 & \left(1-2v_{eff}\right)/2 \end{vmatrix} = \begin{bmatrix} D_{1111}^{eff} & D_{1122}^{eff} & 0 \\ D_{1122}^{eff} & D_{2222}^{eff} & 0 \\ 0 & 0 & D_{1212}^{eff} \end{bmatrix} \tag{41}$$

where E_{eff} is effective Young Modulus and v_{eff} is effective Poisson's ratio. The short notation $\mathbf{D}_{ijkl}^{eff}$ ($11 \leftrightarrow 1, 22 \leftrightarrow 2, 33 \leftrightarrow 3, 32 \leftrightarrow 4$, 13 or $31 \leftrightarrow 5$, 12 or $21 \leftrightarrow 6$) is used for all the entries of the homogenized effective elasticity tensor [40].

In order to make a comparison, the reference solution is computed using standard FEM with very fine mesh (35621 nodes). As shown in Fig. 10, it is seen that all numerical results approach the reference solution with increased number of degree of freedom (DOF). The FEM, ES-FEM and CS-FEM approach the reference solution from the upper bound, whereas NS-FEM with T3 and Q4 elements approaches the reference solution from the lower side. Among all numerical methods, ES-FEM gives the most accurate solution even much better than quadrilateral (Q4) elements using FEM, which is due to stronger softening effect provided by the ES-FEM.

4.2 Multiple material composites

In this section, two different materials with void are bonded together as shown in Fig. 11. The Young moduli for materials 1 and 2 re $E_1 = 0.1\text{GPa}$ and $E_2 = 2\text{GPa}$, Poisson's ratios for materials 1 and 2 are $v_1 = 0.4$ and $v_2 = 0.3$, respectively. Plane strain problem is considered here.

By using the same set of T3 meshes as shown in Fig. 12, Fig. 13 plots the convergence in the effective (homogenized) elasticity components using the different finite element methods. The solutions of all these methods converge to the reference solution (36260 nodes using FEM) with reducing nodal spacing. In terms of the accuracy, the NS-FEM gives similar results to FEM. However, NS-FEM converges the reference solution from the lower bound, whilst FEM converges the reference solution from the upper bound. Among all numerical methods, again ES-FEM provides the best solution in all elasticity components.

Figure 14 outlines the converegcne of effective (homogenized) bulk modulus. It is observed that numerical solution obtained from ES-FEM is again the closest to the reference solution when the same set of mesh is used.

As computational efficiency is an important criterion to assess the performance of numerical methods, the comparison of different analysis methods is shown in Fig. 15. It is clear that the computational time for the ES-FEM and NS-FEM is longer than FEM when the same set of mesh is used. This is because more nodes are used to form the shape function in ES-FEM and NS-FEM. Nevertheless, in terms of computational efficiency, the ES-FEM performs much better than FEM and NS-FEM. This is due to right softened effect in the ES-FEM model.

Another example of numerical homogenization for multi-phase material is shown in Fig. 16. The Young modulus for material 1 and 2 are $E_1 = 200\text{MPa}$ and $E_2 = 30\text{MPa}$, Poisson's ratio for material 1 and 2 are $v_1 = 0.3$ and $v_2 = 0.35$ respectively. The discretization model is shown in Fig. 17.

The convergence rates of effective mechanical properties are presented with different numerical methods shown in Fig. 18. From Fig. 18, again we found that the ES-FEM gives much better solution than FEM and NS-FEM. Fig. 19 plots the result of effective bulk modulus converging to the reference solution using different methods. As we expected, FEM and ES-FEM give upper bound solution and NS-FEM provides lower bound solution. The ES-FEM is able to achieve a close to exact stiffness, and it gives the best solution in the prediction of effective bulk modulus.

4.3 Tissue Scaffold Example

As the rapid development of additive fabrication technology, scaffold tissue engineering is growing fast. Materials with periodic cellular micro-architectures are becoming particularly advantageous due to high manufacturability and tailored effective properties. Tissue scaffold

involves two important criteria concerned in the design stage: one is overall stiffness, which is able to provide similar load-bearing capacities to surrounding tissues; the other one is permeability, which offers sufficient porosity for mass transfer and vascularization. In this example, it is assumed that the Young modulus E=100Mpa, $v = 0.3$ of scaffold materials the permeability coefficient $\kappa = 0.5$ in the 3D tissue scaffold as shown in Fig. 20.

The discretization of 3D base cell is presented in Fig. 21. Fig. 22 shows the convergence of effective bulk modulus using 3D ES-FEM, NS-FEM and FEM. In order to make a comparison, the reference solution with very fine mesh (216,000 nodes) is also plotted together. It is clearly shown that 3D ES-FEM still gives the best solution of these different numerical methods when the same set of mesh is used. The 3D NS-FEM still approaches the reference solution from the lower bound.

The convergence of effective permeability is presented in Fig. 23 with reference solution using 216,000 nodes. Again, it is found the NS-FEM provides the lower bound solution of effective permeability. The ES-FEM and FEM provides the upper bound solution of effective permeability. In terms of accuracy, the ES-FEM using tetrahedral (T4) elements performs the best.

5. Conclusion

In this paper, smoothed finite element methods (S-FEMs) were formulated to solve the numerical homogenization problems. Various 2D and 3D examples were presented to demonstrate the accuracy and convergence of S-FEMs in the evaluation of effective (homogenized) mechanical properties of periodic microstructural composites. In summary, some conclusions are drawn as follows:

1. The implementation of smoothed finite element method in numerical homogenization of composite material is fairly straightforward. No additional parameters are involved in the formulation.

2. For the first time, the NS-FEM was found to be able to give the lower bound solution in the computation of effective (homogenized) material properties of composites.

3. The ES-FEM was found to stand out from all different forms of finite element method in 2D and 3D, which provided the best solution to characterization of the effective mechanical properties of composites.

Reference

[1] Boso DP, Lefik M, Schrefler BA, Recent developments in numerical homogenization. Computer Assisted Mechanics and Engineering Sciences, 2009; 16: 161–183.

[2] Zhou SW, Li Q, A microstructure diagram for known bounds in conductivity. Journal of Materials Research, 2008; 23(3) 798-811.

[3] Hashin, Z, Shtrikman S, A variational approach to the theory of composite elastic materials. Journal of the mechanics and physics of solids, 1963; 11(2): 127-140.

[4] Berryman JG, Milton GW, Microgeometry of random composites and porous-media. Journal of Physics D-Applied Physics, 1988; 21(1) : p87-97.

[5] Cherkaev AV, Gibiansky LV, coupled estimates for the bulk and shear moduli of a 2-dimensional isotropic elastic composite. Journal of the mechanics and physics of solids, 1993; 41(5): 937-980.

[6]Ye Z, Yu W, A New Approach to Bounding Effective Properties of Random Heterogeneous Materials. 52nd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference.

[7] Hollister SJ, Kikuchi N, A Comparison of Homogenization and Standard Mechanics Analyses for Periodic Porous Composites. Computational Mechanics, 1992; 10: 73－95.

[8] Hassani B, Hinton E, A Review of Homogenization and Topology Optimization I－Homogenization Theory for Media with Periodic Structure. Computer & Structures, 1998: 69: 707－717.

[9] Hassani B, Hinton E, A Review of Homogenization and Topology Optimization II – Analytical and Numerical Solution of Homogenization Equations. Computer & Structures, 1998; 69: 719–738.

[10] Sun H, Di S, Zhang N, Wu C, Micromechanics of Composite Materials using Multivariable Finite Element Method and Homogenization Theory. International Journal of Solids and Structures, 2001; 38: 3007–3020.

[11] Alexander A, Tzeng JT, Three Dimensional Effective Properties of Composite Materials for Finite Element Applications, Journal of Composite Materials, 1997; 31(5): 466-485.

[12] Guedes JM, Kikuchi N, Preprocessing and postprocessing for materials based on the homogenization method with adaptive finite element method. Computer methods in applied mechanics and engineering, 1990; 83: 143-198.

[13] Yi Y, Park S, Youn S, Asymptotic Homogenization of Viscoelastic Composites with Periodic Microstructures. International Journal of Solids and Structures, 1998; 35: 2039–2055.

[14] Fujii D, Chen BC, Kikuchi N, Composite material design of two-dimensional structures using the homogenization design method. International Journal for Numerical Methods in Engineering, 2001; 50: 2031-2051.

[15] Fang Z, Sun W, Tzeng JT, Asymptotic Homogenization and Numerical implementation to Predict the Effective Mechanical Properties for Electromagnetic Composite Conductor. Journal of Composite Materials, 2004; 38: 1371.

[16] Bendsoe MP, Kikuchi N, Generating optimal topologies in structural design using a homogenization method. Computer methods in applied mechanics and engineering, 1988; 71: 197-224.

[17] Michel JC, Moulinec H, Suquet P, Effective properties of composite materials with periodic microstructure: a computational approach. Computer methods in applied mechanics and engineering, 1999; 172: 109-143.

[18] Andreassen E, Andreasen CS, How to determine composite material properties using numerical homogenization. Computational Materials Science, 2014; 83: 488–495.

[19] Zienkiewicz OC, Taylor RL. The Finite Element Method (5th edn). Butterworth Heinemann: Oxford, U.K., 2000.

[20] Liu GR, Meshfree methods: Moving beyond the Finite Element Method. CRC Press: Boca Raton, U.S.A, 2002.

[21] Liu GR, A G space theory and a weakened weak (W2) form for a unified formulation of compatible and incompatible methods: Part I theory. International Journal for Numerical Methods in Engineering, 2010; 81: 1093-1126.

[22] Liu GR, Zhang GY. Smoothed Point Interpolation Methods. Singapore: World Scientific, 2013.

[23] Chen JS, Wu CT, Yoon S, You Y, A stabilized conforming nodal integration for Galerkin mesh-free methods, International Journal for Numerical Methods in Engineering, 2001; 50: 435–466.

[24] Liu GR, Dai KY, Nguyen TT, A smoothed finite element method for mechanics problems, Computational Mechanics, 2007; 39: 859－77.

[25] Liu GR, Nguyen TT, Dai KY, Lam KY, Theoretical aspects of the smoothed finite element method (SFEM). International Journal for Numerical Methods in Engineering, 2007; 71: 902－30.

[26] Liu GR, Nguyen-Thoi T, Dai KY, Lam KY, Theoretical aspects of the smoothed finite element method (SFEM). International Journal for Numerical Methods in Engineering, 2007; 71: 902–930.

[27] Hung NX, Bordas S, Hung N-D, Smooth finite element methods: convergence, accuracy and properties. International Journal for Numerical Methods in Engineering, 2008; 74: 175–208.

[28] Liu GR, Nguyen TT, Nguyen HX, Lam KY, A node-based smoothed finite element method (NS-FEM) for upper bound solutions to solid mechanics problems. Computers and Structures, 2009; 87: 14-26.

[29] Liu GR, Nguyen-Thoi T, Lam KY, An edge-based smoothed finite element method (ES-FEM) for static, free and forced vibration analyses of solids. Journal of Sound and Vibration. 2009; 320: 1100–1130.

[30] He ZC, Li GY, Zhong ZH, Cheng AG, Zhang GY, Li E, Liu GR, An ES-FEM for accurate analysis of 3D mid-frequency acoustics using tetrahedron mesh. Computers and Structures, 2012; 106-107: 125-134.

[31] He ZC, Li GY, Zhong ZH, Cheng AG, Zhang GY, Liu GR, Li E, Zhou Z, An edge-based smoothed tetrahedron finite element method (ES-T-FEM) for 3D static and dynamic problems. Computational mechanics, 2013; 2:221-236.

[32] Li E, Liu GR, Tan V, and He ZC, Modeling and simulation of bioheat transfer in the human eye using the 3D alpha finite element method (αFEM). Intentional Journal for Numerical Methods in Biomedical Engineering, 2010; 26: 955–976.

[33] Li E, Liu GR, Tan V, Simulation of Hyperthermia Treatment Using the Edge-Based Smoothed Finite-Element Method. Numerical Heat Transfer, Part A: Applications, 2010; 57(11): 822 -847.

[34] Li E, Liu GR, Tan V, and He ZC, An efficient algorithm for phase change problem in tumor treatment using αFEM. International Journal of Thermal Sciences, 2010; 49 (10): 1954-1967.

[35] Li E, He ZC, Xu X, A novel edge-based smoothed tetrahedron finite element method (ES-T-FEM) for thermomechanical problems. International Journal of Heat and Mass Transfer, 2013; 66: 723–732.

[36] He ZC, Cheng AG, Zhong ZH, Zhang GY, Li GY, Li E, An improved eigenfrequencies prediction for three-dimensional problems using face-based smoothed finite element method. Journal of Acta Mechanica Solida Sinica, 2013; 26: 140-150.

[37] He ZC, Li GY, Li E, Zhong ZH, Liu GR, Mid-frequency acoustics analysis using edge-based smoothed tetrahedron radial point interpolation method (ES-T-RPIM). Submitted to Computational methods. 2013 (Accepted).

[38] Kuznetsov S. Homogenization Methods for Problems with Multiphysics, Temporal and Spatial Coupling. PhD thesis. Columbia University 2012.

[39] Liu GR. A generalized gradient smoothing technique and the smoothed bilinear form for Galerkin formulation of a wide class of computational methods. International Journal of Computational Methods, 2008; 5:199–236.

[40] Kruijf ND, Zhou SH, Li Q, Mai YW, Topological design of structures and composite materials with multiobjectives. International Journal of Solids and Structures, 2007; 44: 7092-7109.

Figure

(a) Tissue Scaffold
(b) Cuttlebone
Figure 1: 3D printing of composite material

F

(b) Microstructure   (c) Unit cell
$(j = 1, 2, ...N)$

x_i
$(i = 1, 2, ...M)$
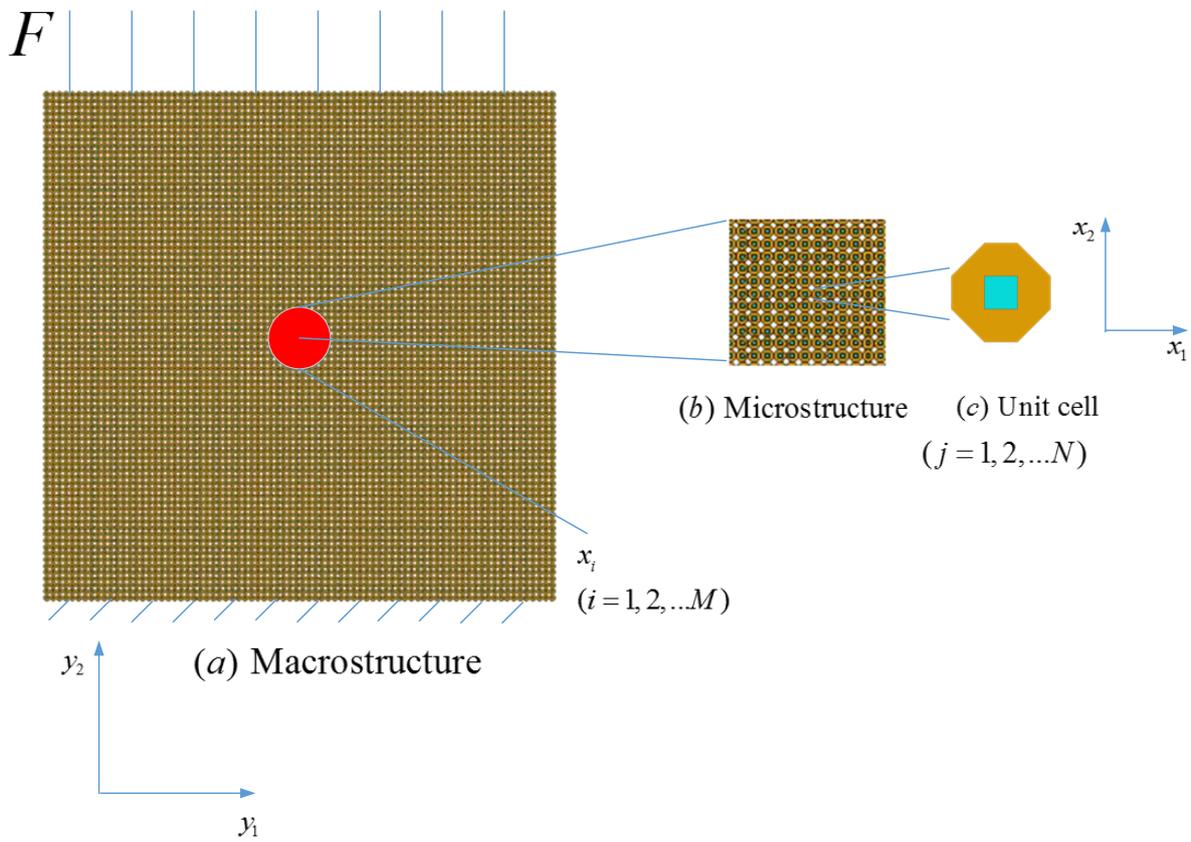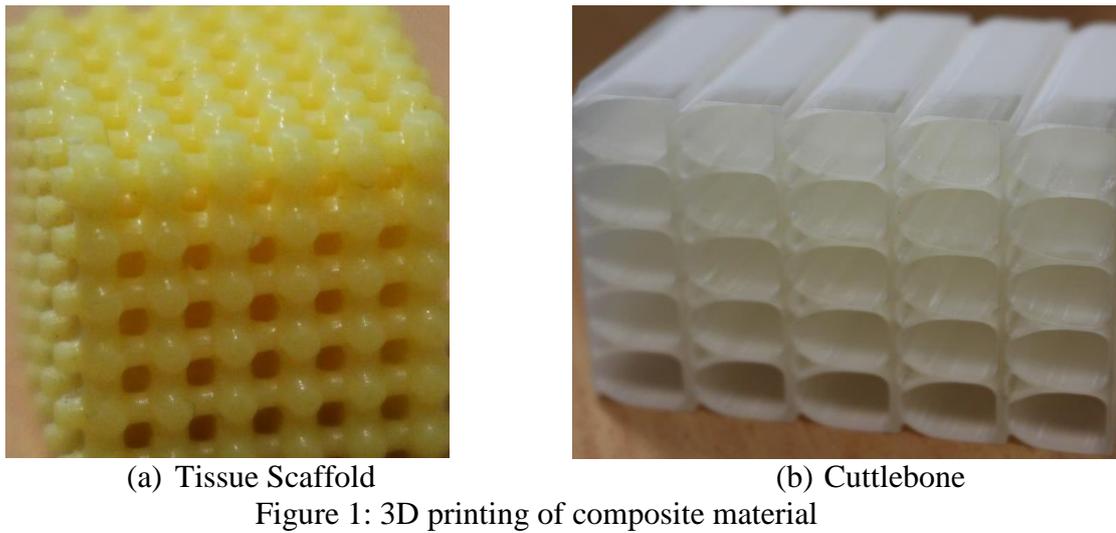
(a) Macrostructure

Figure 2: Composite materials with periodic microstructure
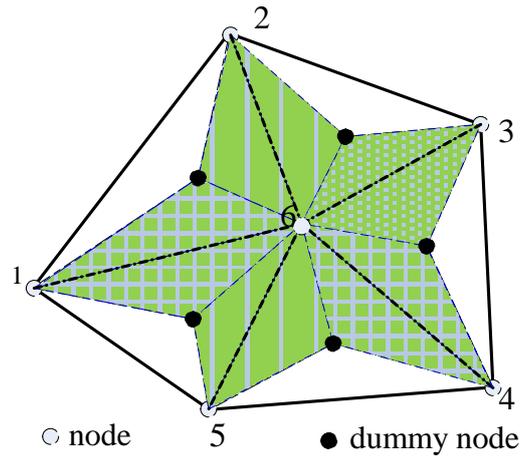
a) Single material
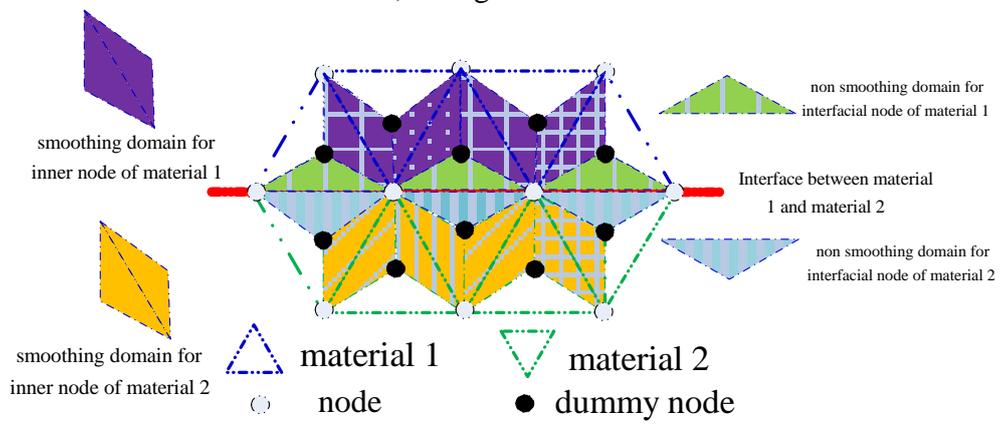
b) Multi-material
Figure 3: Illustration of ES-FEM

○ node   ● dummy node

a) Single material

smoothing domain for
inner node of material 1

smoothing domain for
interfacial node of material 1

Interface between material
1 and material 2

smoothing domain for
interfacial node of material 2

smoothing domain for
inner node of material 2

△ material 1   ▽ material 2

◐ node   ● dummy node
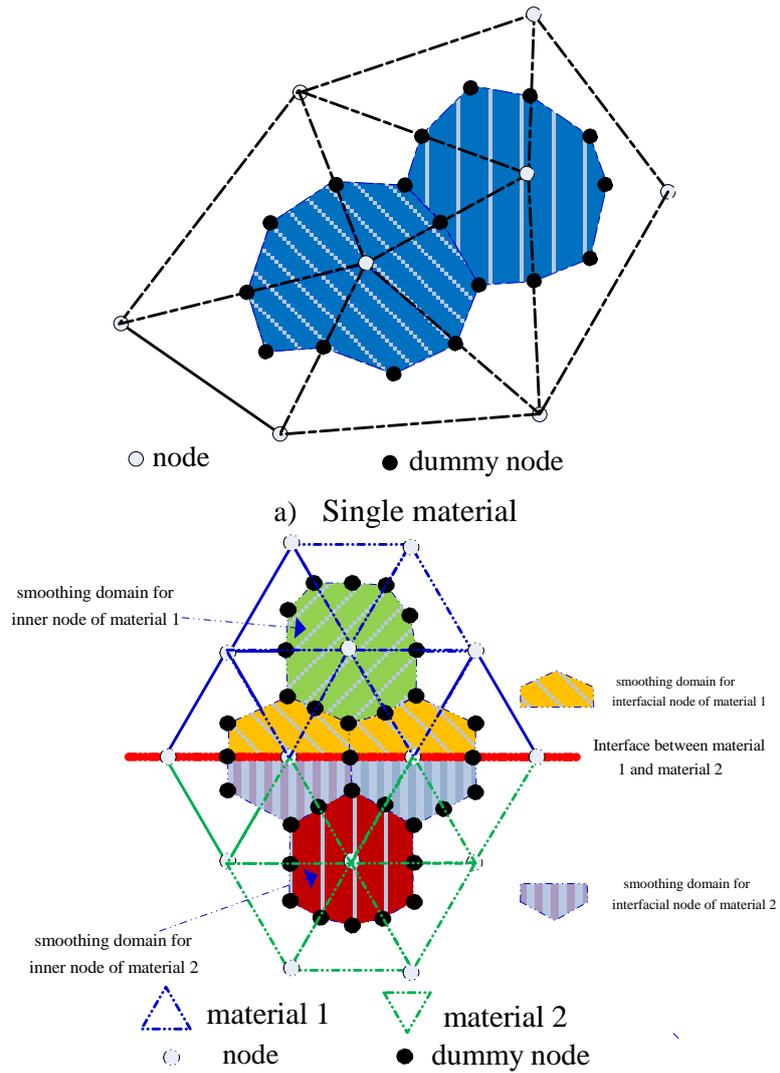
b) Multi-material

Figure 4: Illustration of NS-FEM

a) Two cell

b) Four cell
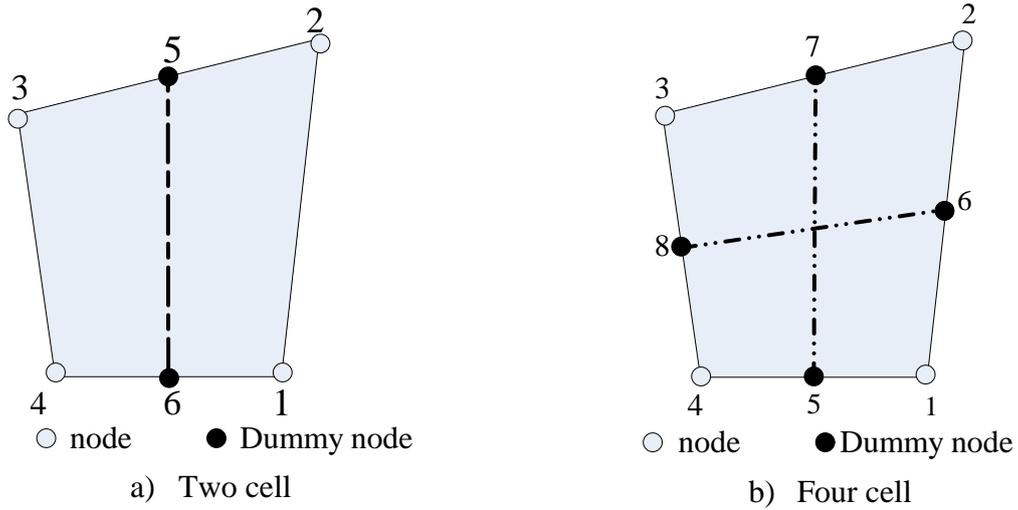
Figure 5: Illustration of CS-FEM
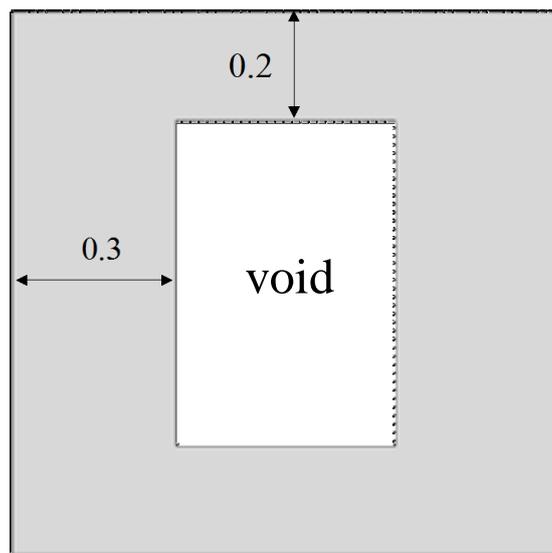
Figure 6: Unit cell of a periodic composite

455
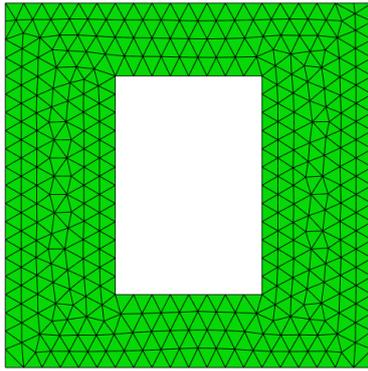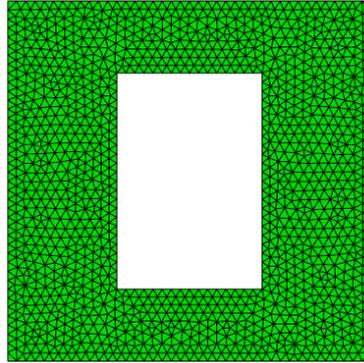
a) 400 nodes with T3   b) 1387 nodes with T3   c) 8321 nodes with T3
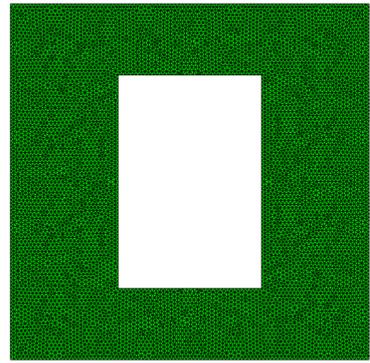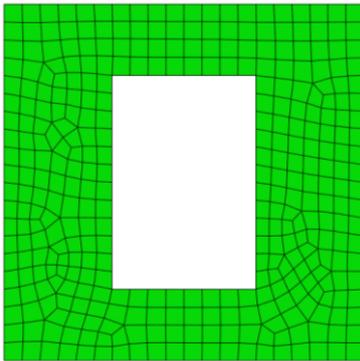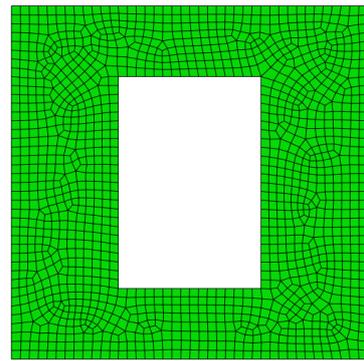
d) 406 nodes with Q4   e) 2341 nodes with Q4   f) 9517 nodes with Q4
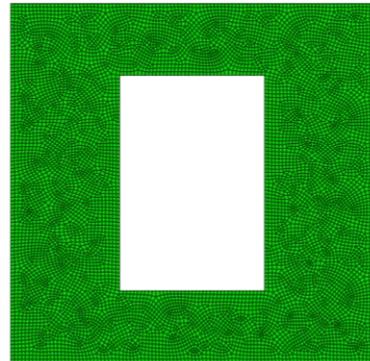
Figure 7: Discretization of base cell using triangular (T3) and quadrilateral (Q4) elements

a) Periodic structure   b) RVE
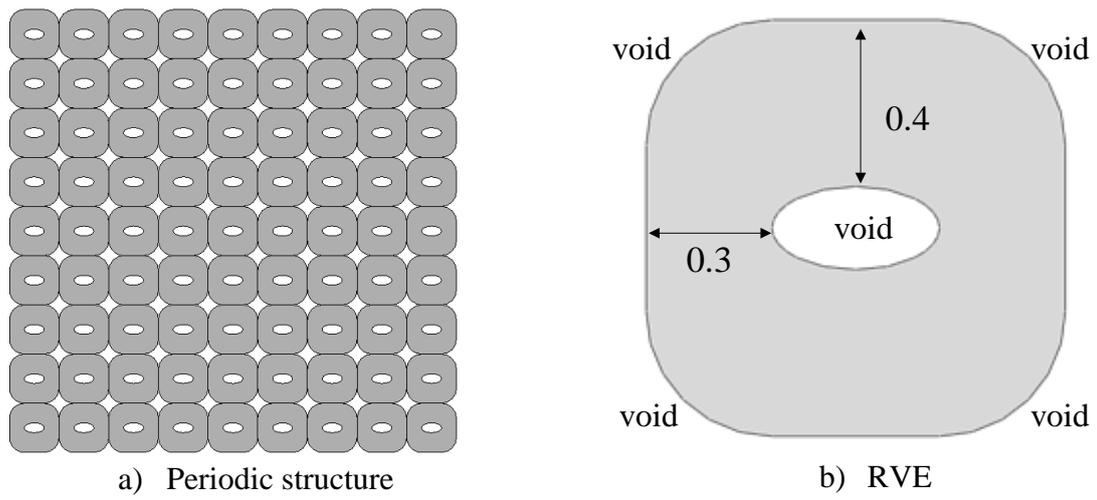
Figure 8: Base cell of a composite

a) 133 nodes   b) 463 nodes   c) 1621 nodes

d) 162 nodes   e) 521 nodes   f) 1837 nodes
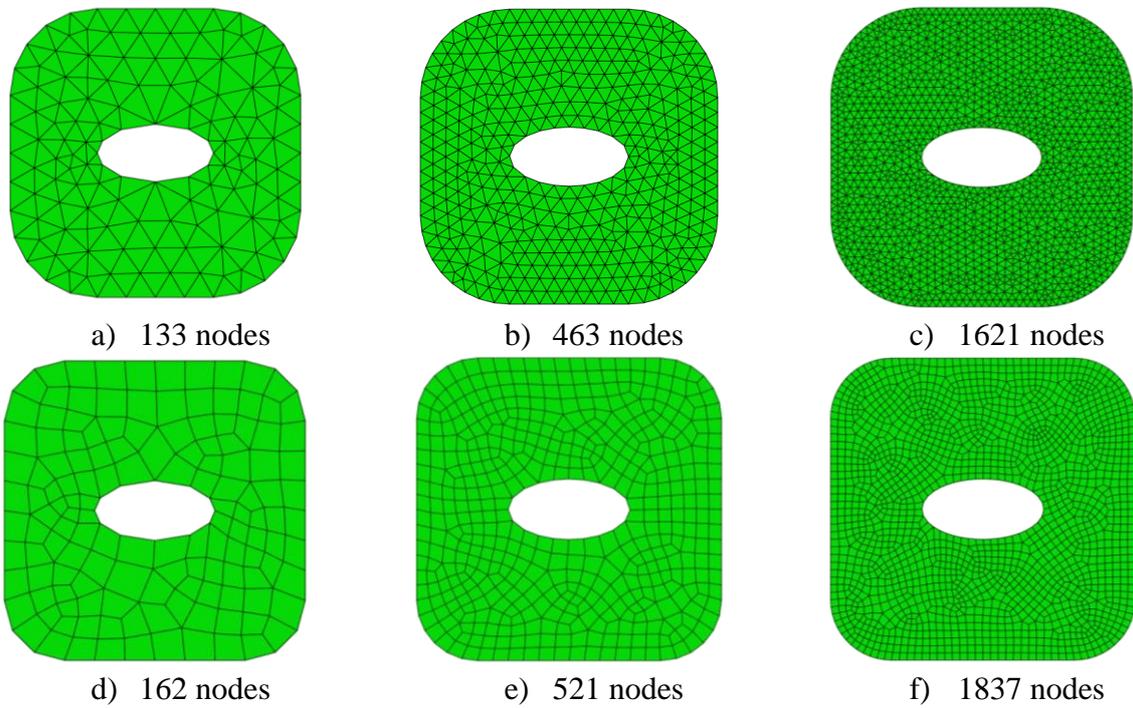
Figure 9:Domain discretization using triangular and quadrilateral element
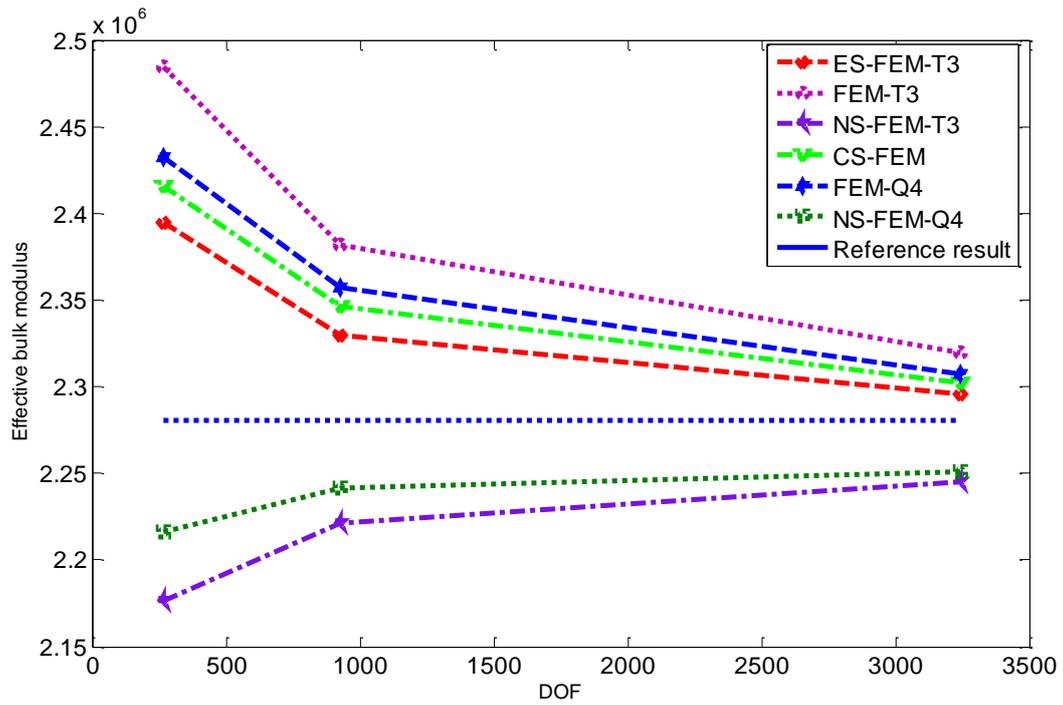
Figure 10: Convergence of bulk modulus

a) Periodic structure

b) RVE
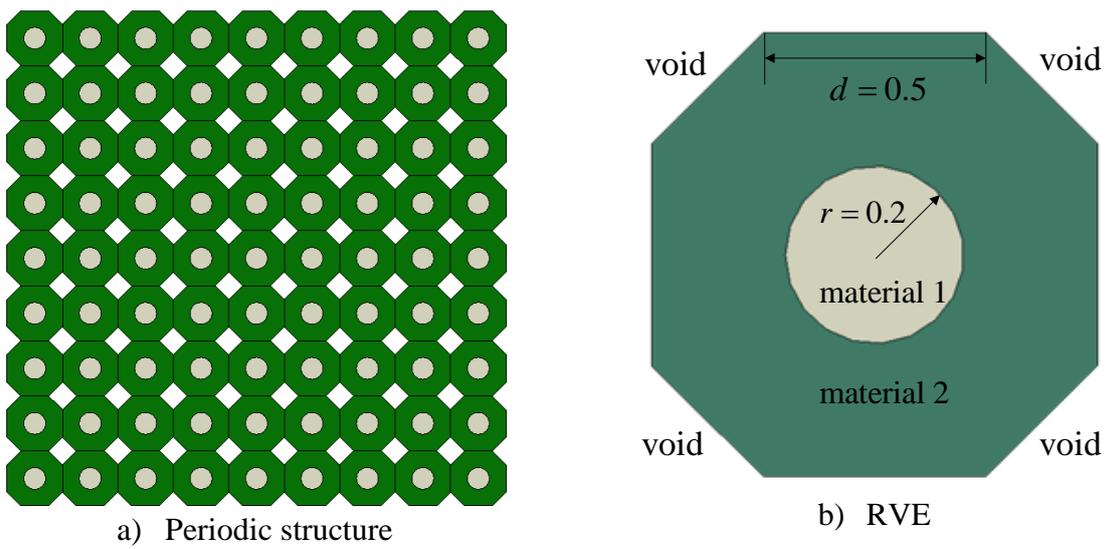
Figure 11: Base cell of a composite
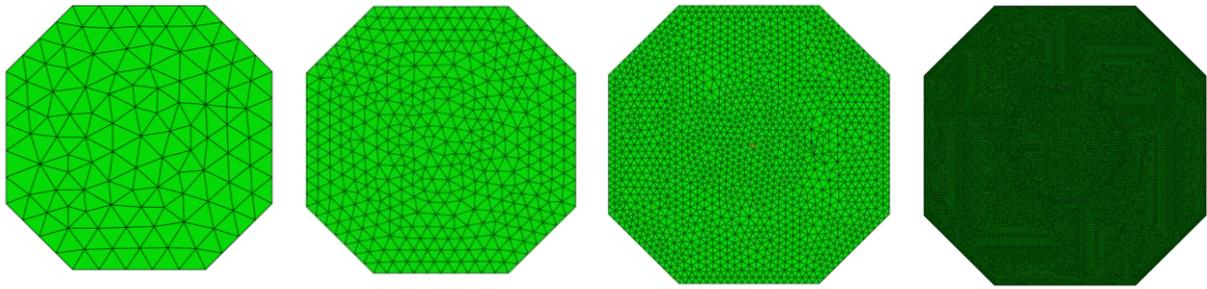
(a)122 nodes   (b) 417 nodes   (b) 1445 nodes   (d) 36260 nodes

Figure 12: Domain discretization using triangular (T3) element
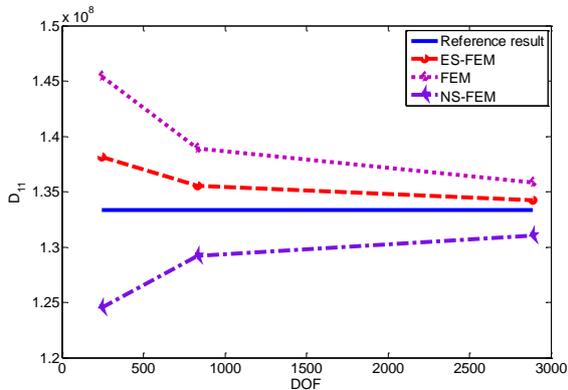
(a)
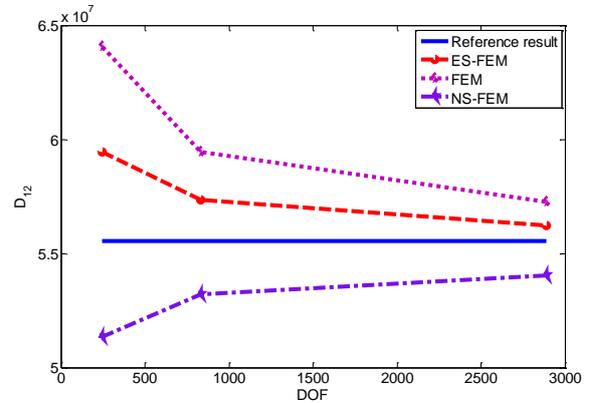
(b)
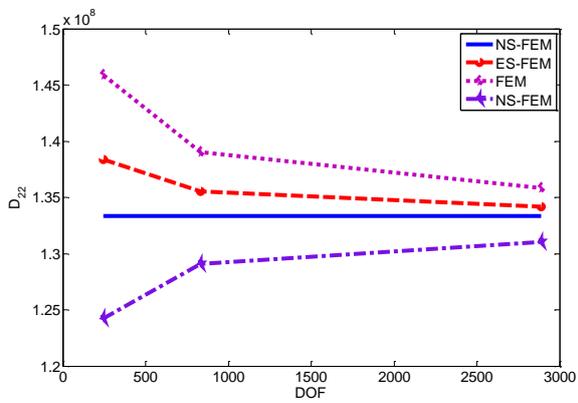
(c)
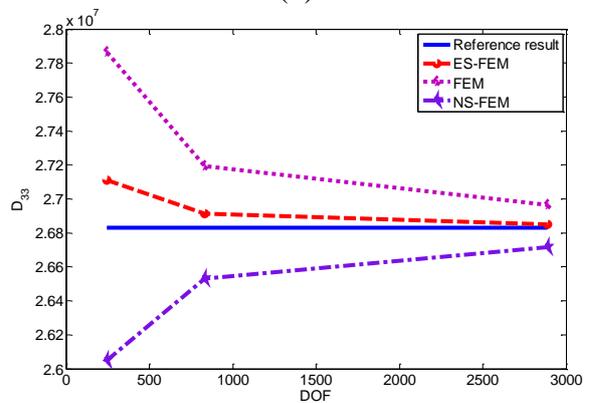
(d)

Figure 13:Convergence of elasticity components
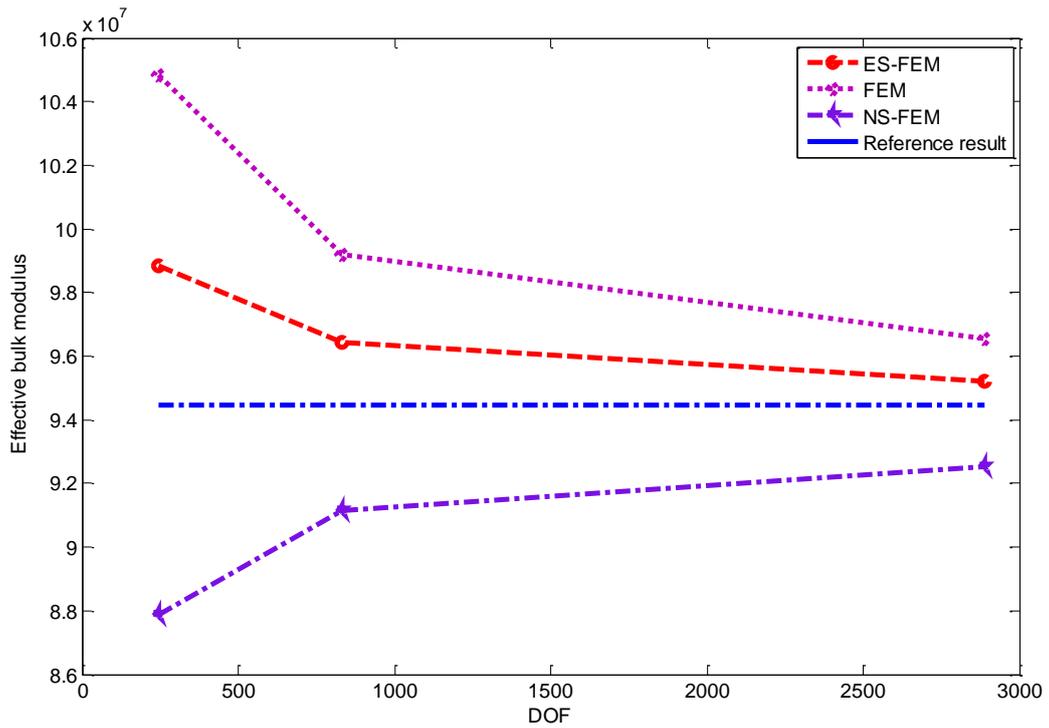
Figure 14: Convergence of effective bulk modulus
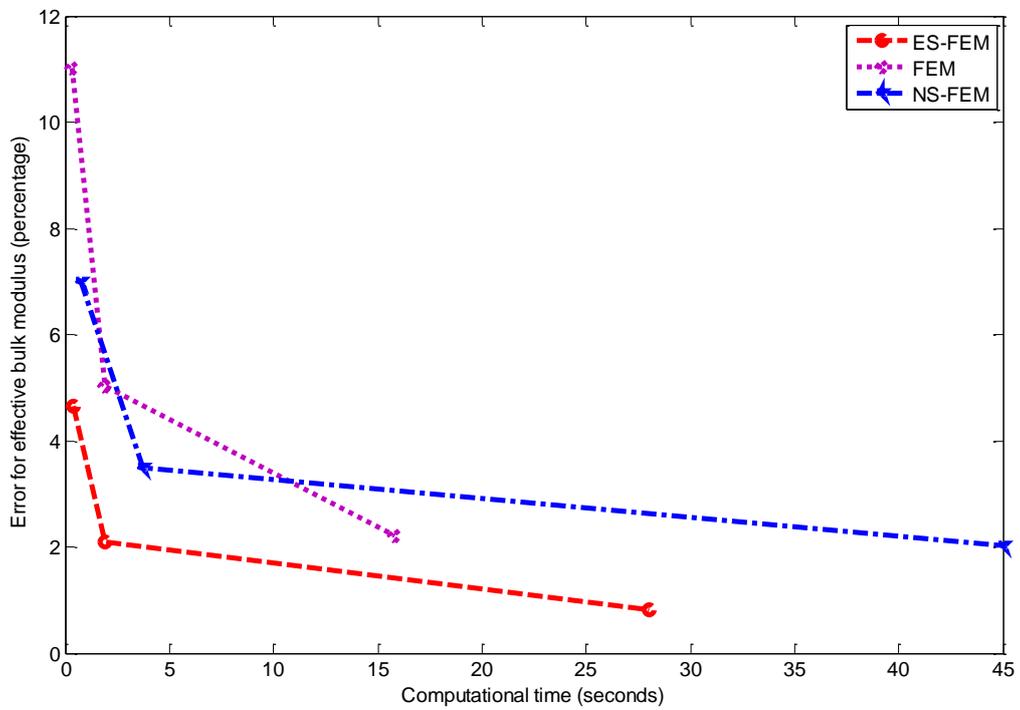
Figure 15: Convergence of effective bulk modulus

a) Periodic structure   b) RVE
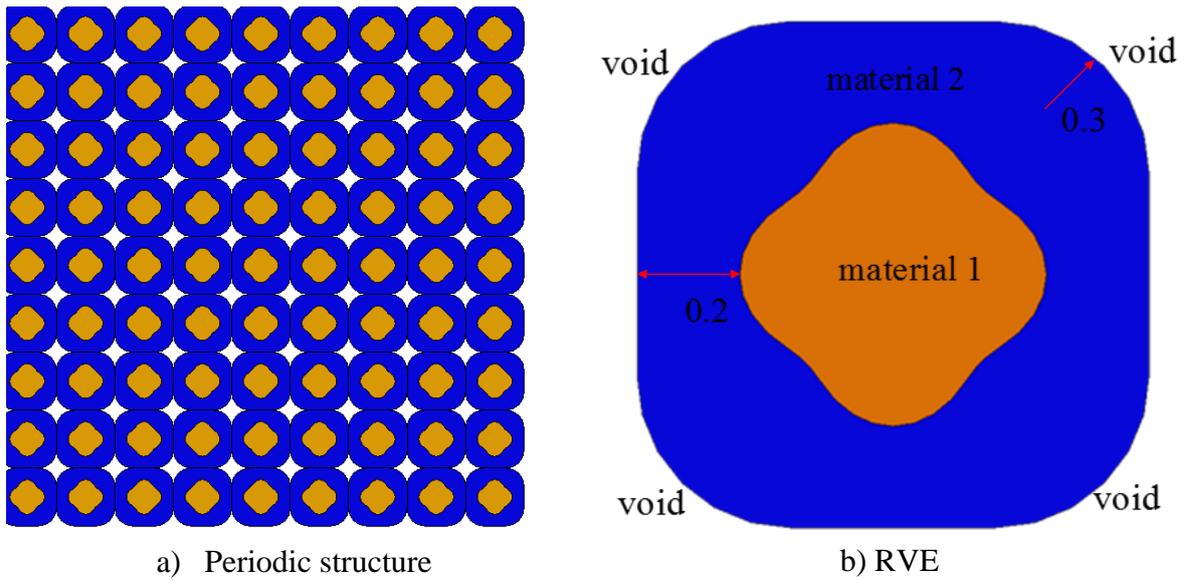
Figure 16: Base cell of a composite

(a)145 nodes   (b) 501 nodes   (b) 1756 nodes   (d) 38685 nodes
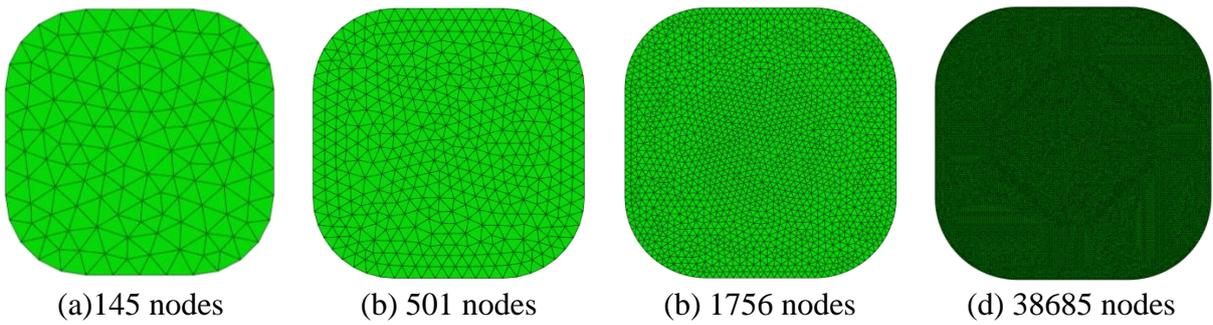
Figure 17: Domain discretization using triangular element
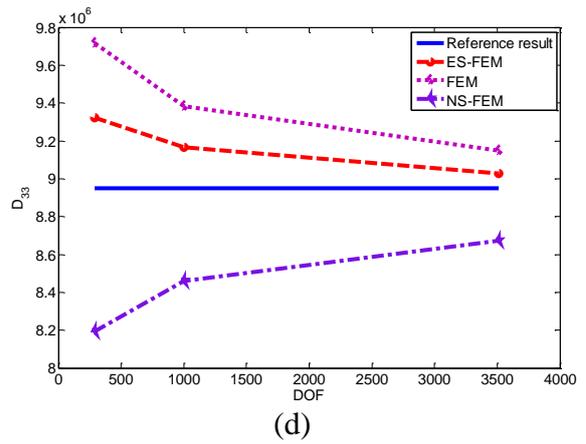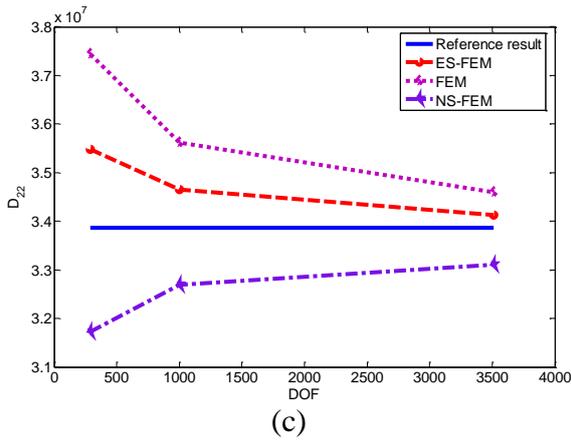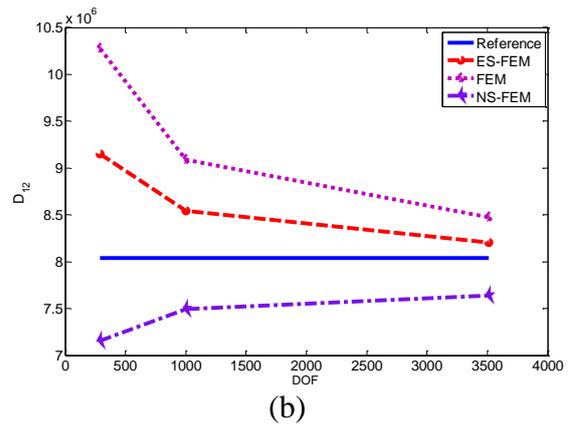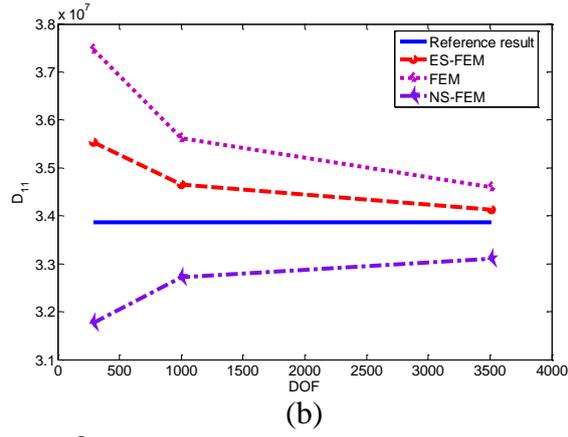
Figure 18:Convergence of bulk modulus
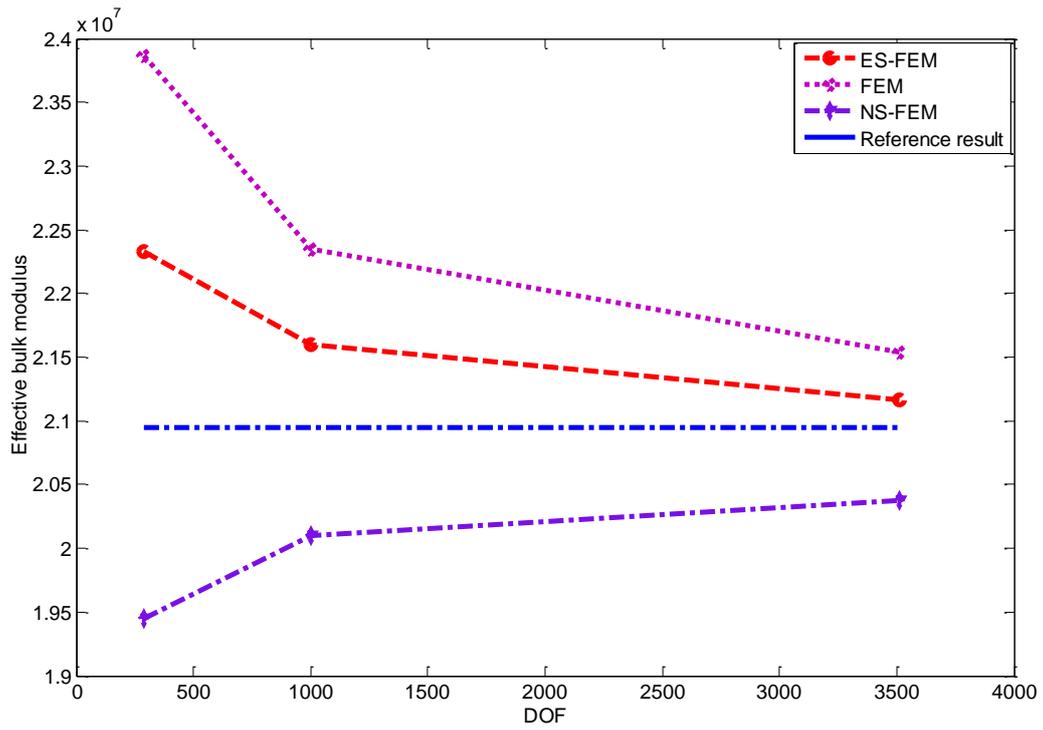
Figure 19: Convergence of bulk modulus
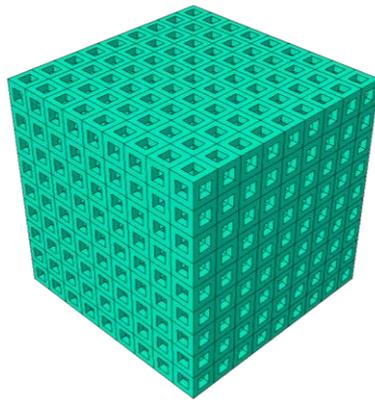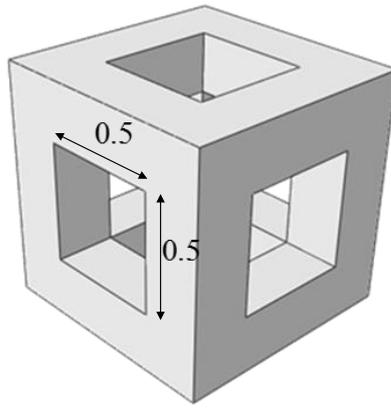
a) 3D tissue Scaffold   b) RVE for stiffness   c) RVE for
                                 model   Permeability model
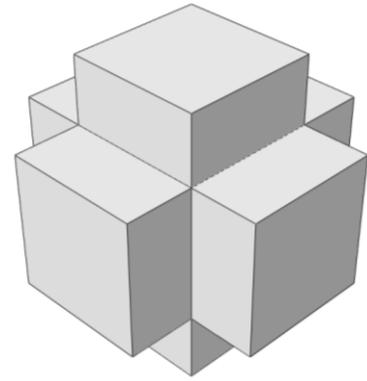
Figure 20: 3D tissue scaffold structure
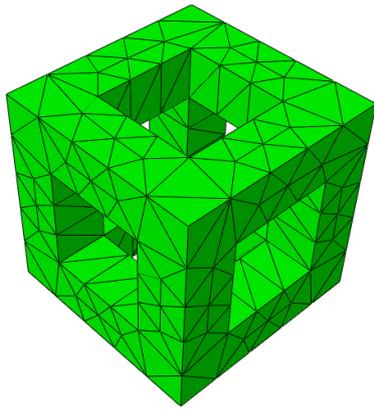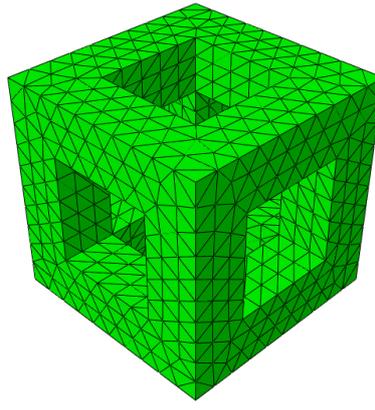
281 nodes   1255 nodes   2346 nodes

a) Stiffness model
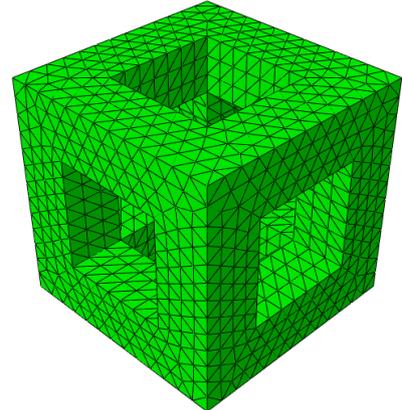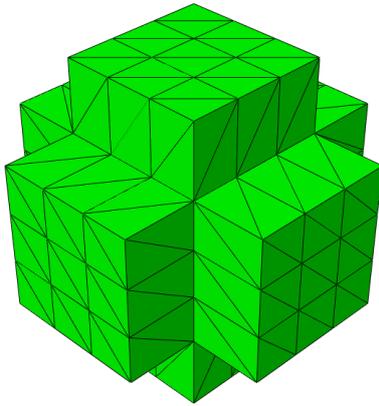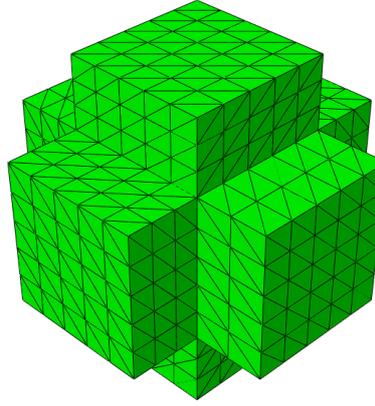
181 nodes   929 nodes   4691 nodes

b) Permeability model
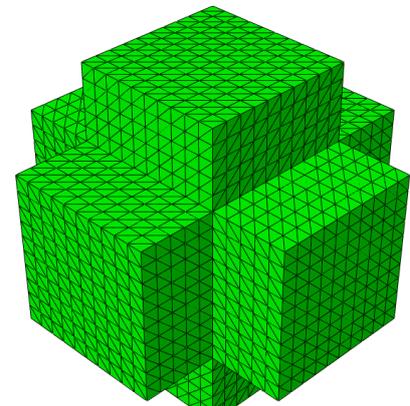
Figure 21: Domain discretization using tetrahedral element
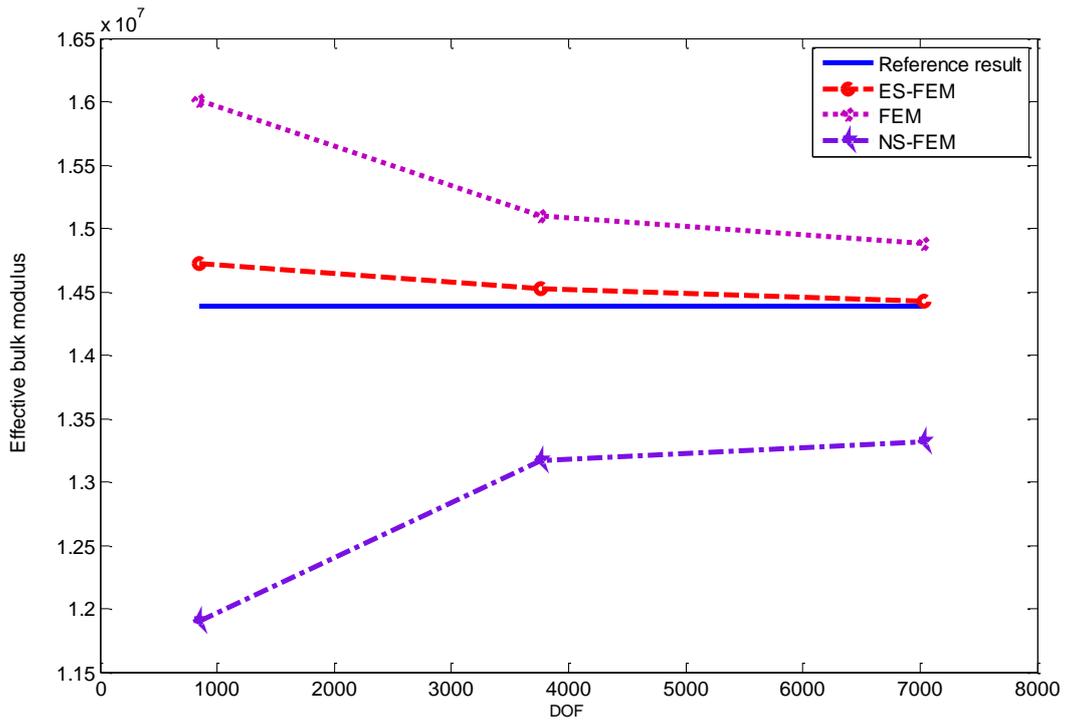
Figure 22: Convergence of effective bulk modulus
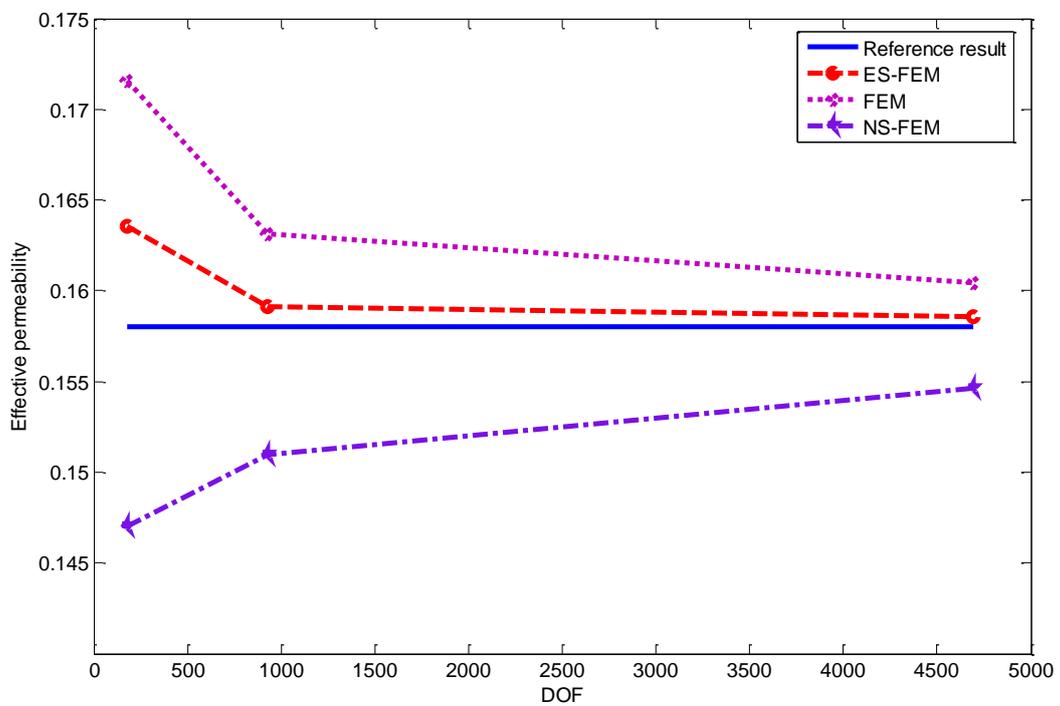
Figure 23: Convergence of effective permeability

Table

Table 1:Symmetry conditionsforthedifferentteststrains in 2D elasticity

Test strains	$x = 0,1$	$y = 0,1$
ε_{12}	$u_y^e = 0$	$u_x^e = 0$
$\varepsilon_{ij}, i = j$	$u_x^e = 0$	$u_y^e = 0$

Table 2:Symmetry conditionsforthedifferentteststrains in 3D elasticity [1]

Test strains	$x = 0,1$	$y = 0,1$	$z = 0,1$
ε_{12}	$u_y^e = u_z^e = 0$	$u_x^e = u_z^e = 0$	$u_z^e = 0$
ε_{23}	$u_x^e = 0$	$u_x^e = u_z^e = 0$	$u_x^e = u_y^e = 0$
ε_{13}	$u_y^e = u_z^e = 0$	$u_y^e = 0$	$u_x^e = u_y^e = 0$
$\varepsilon_{ij}, i = j$	$u_x^e = 0$	$u_y^e = 0$	$u_z^e = 0$

Table 3: Numerical results using different methods (Triangular element)

mesh	D_{11}^H	D_{12}^H	D_{22}^H	D_{33}^H	Methods
400 nodes T3	12.9527	3.2141	17.5111	2.7214	ES-FEM
400 nodes T3	13.1122	3.2948	17.6227	2.8669	FEM
400 nodes T3	12.4404	2.9412	17.1006	2.3718	NS-FEM
1387 nodes T3	12.8557	3.1472	17.4342	2.6556	ES-FEM
1387 nodes T3	12.9305	3.1926	17.4895	2.7142	FEM
1387 nodes T3	12.6398	3.0153	17.2643	2.5082	NS-FEM
8321 nodes T3	12.8181	3.1209	17.4057	2.6318	ES-FEM
8321 nodes T3	12.8447	3.1402	17.4269	2.6523	FEM
8321 nodes T3	12.7399	3.0640	17.3446	2.5765	NS-FEM

Table 4: Numerical results using different methods (Quadrilateral element)

mesh	D_{11}^H	D_{12}^H	D_{22}^H	D_{33}^H	Remarks
406 nodes T4	13.0335	3.2257	17.5535	2.7713	CS-FEM
406 nodes T4	13.1122	3.2948	17.6227	2.8032	FEM
406 nodes T4	12.5140	2.9605	17.1780	2.3793	NS-FEM
2341 nodes T4	12.8833	3.1375	17.4446	2.6901	CS-FEM
2341 nodes T4	12.9109	3.1388	17.4621	2.7129	FEM
2341 nodes T4	12.6777	3.0256	17.2966	2.5293	NS-FEM
9517 nodes T4	12.8599	3.1123	17.4168	2.6715	CS-FEM
9517 nodes T4	12.8876	3.1095	17.4310	2.6934	FEM
9517 nodes T4	12.7400	3.0631	17.3470	2.5768	NS-FEM

Table 5: Published results reported by other researches [9]

mesh	D_{11}^H	D_{12}^H	D_{22}^H	D_{33}^H	Remarks
20x20 4-node	13.015	3.241	17.552	2.785	Ref. [17]
1st adapt	12.910	3.178	17.473	2.714	Ref. [17]
2nd adapt	12.865	3.146	17.437	2.683	Ref. [17]
3rd adapt	12.844	3.131	17.421	2.668	Ref. [17]
436 8-node	12.839	3.139	17.422	2.648	HOMOG case (a)
305 8-node	12.820	3.124	17.407	2.634	HOMOG case (a)

Cohesive zone modeling of cracking along the Cu/Si interface in nanoscale components

*Y.B. Yan[1], T. Sumigawa[2], and T. Kitamura[2]

[1]Institute of Systems Engineering, China Academy of Engineering Physics, Miayang 621900, China.
[2]Department of Mechanical Engineering and Science, Kyoto University, Kyoto 606-8501, Japan.

*Corresponding author: yanyabin@gmail.com

Abstract

In this study, crack initiation and propagation along the Cu/Si interface in multilayered films (Si/Cu/SiN) with different thicknesses of the Cu layer (20 and 200 nm) are experimentally investigated using a nano-cantilever and millimeter-sized four-point bending specimens. To examine the cohesive zone model (CZM) criterion for interfacial delamination along the Cu/Si interface in nanoscale stress concentration, an exponential type of CZM is utilized to simulate the observed delamination processes using the finite element method. After the CZM parameters for the Cu/Si interface are calibrated by an experiment, interface cracking in other experiments is predicted. This indicates that the CZM criterion is universally applicable for describing cracking along the interface regardless of specimen dimensions and film thickness which include the differences in plastic behavior and residual stress. The CZM criterion can also predict interfacial cracking along Cu/Si interfaces with different stress singularities.

Keywords: Interface, Cohesive zone model, Delamination, Nanoscale, Thin films

Introduction

Many micro-electronic and mechanical devices include bi-material interfaces, which, however, are liable to crack due to stress concentration originated from deformation mismatch [Rice (1998); Kitamura et al. (2002)]. So, for the assurance of reliability, it is important to evaluate the interface strength under nanoscale stress concentration.

Cohesive zone model (CZM) for cracking in bulk components has been successfully used for interfacial delamination in many material systems [Elices et al. (2002)]. However, the application for delamination induced by nanoscale stress concentration in small components has not been fully investigated. So, it is necessary to examine the applicability of CZM for fracture in nano-components. In this study, an exponential type of CZM is used to simulate the crack initiation at interface edge and propagation along the Cu/Si interface with nanoscale stress concentration in different experiments. The reliability of CZM concept for nano- components is examined on the basis of experiments and analyses.

1. Experiments on crack initiation and propagation

1.1 Tested materials

The tested materials are multilayer Si/Cu/SiN (silicon /copper/silicon nitride) with different Cu thicknesses. After a Si (100) wafer surface is cleaned by inverse sputtering, a Cu layer with a thickness of 20 or 200 nm is deposited by radio-frequency (RF) magnetron sputtering. A SiN thin layer of about 500 nm thickness is then formed on the Cu layer.

Since the Cu layer shows elasto-plastic behavior during experiments, the corresponding plastic properties of 20 nm [Sumigawa et al. (2010a)] and 200 nm [Takahashi et al. (2007)] thick Cu layers determined in previous experiments are used in following numerical calculations.

1.2 Specimens and experiment set-ups

1.2.1 Nano-cantilever experiment

Fig. 1 shows the specimen used for the nano-cantilever experiments of the Cu thin films of 20 and 200 nm thick (denoted as nano-cantilever (20 nm Cu) [Sumigawa et al. (2010b)] and nano-cantilever (200 nm Cu) [Hirakata et al. (2007)]). Crack initiation at interface edge is investigated here. A minute mechanical loading apparatus is used to apply a force and the behavior of interface fracture is observed in situ by transmission electron microscopy (TEM). The load is applied to the SiN layer with a diamond loading tip to apply stress to Cu/Si interface by a bending moment as shown in Fig. 1. No damage or defect is observed near the Cu/Si interface edge before the experiments.
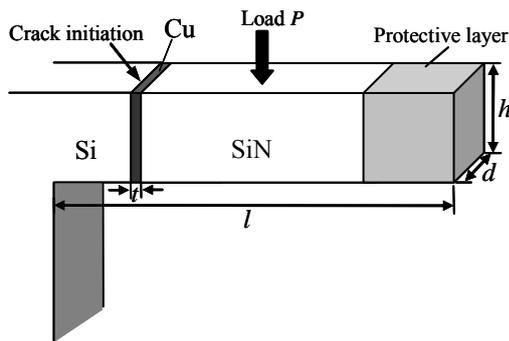
No.	t	l	d	h
A1		955	250	434
A2	20	840	308	233
A3		871	415	642
I1		1270	340	300
I2	200	1295	260	320
I3		1260	390	290
I4		762	320	980

Fig. 1 Schematic illustration of nano-cantilever (20 nm Cu) and nano-cantilever (200 nm Cu) specimens (length unit: nm)

1.2.2 Modified four-point bend experiment [Hirakata et al. (2007)]

A rectangular coupon with millimeter-scale width cut from the material with the 200 nm-thick Cu thin film is glued to a plate of stainless steel, as shown in Fig. 2. After a pre-crack is introduced, the load P is applied at a constant displacement rate. The whole specimen size is millimeter-scale, which is almost a thousand times larger than those of the nano-cantilever specimens. The study focus is the crack propagation along the Cu/Si interface from the pre-crack.
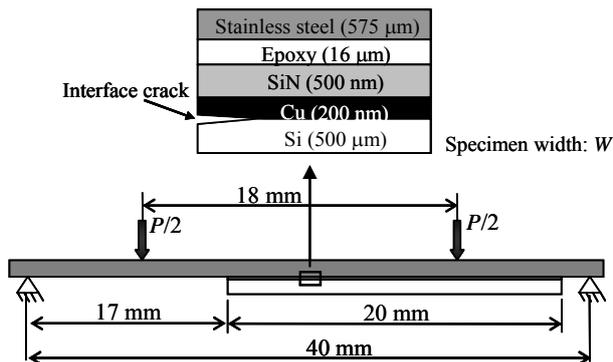
Specimen No.	W, mm
P1	5.58
P2	4.79
P3	4.54

Fig. 2 Schematic illustration of the modified four-point bend specimen (200 nm Cu) and the loading system

1.3 Experimental results

1.3.1 Nano-cantilever experimental results

Fig. 3 shows the relationships between the applied load and deflection at the end of the cantilever arm under monotonic loading. With increasing the applied load, the experimental curves become nonlinear by plastic deformation of Cu layer. At the critical load, the crack initiates at interface edge

and immediately leads to the complete delamination of the entire interface. Similar fracture behavior is observed in the nano-cantilever (20 nm Cu) experiments. The critical loads, P_C, for crack initiation are listed in the table of Fig. 3, which shows great dependency on specimen geometry.

Specimen	Critical load, P_C, μN
A1	16.9
A2	10.3
A3	85.3
I1	13.0
I2	11.7
I3	27.5
I4	94.6
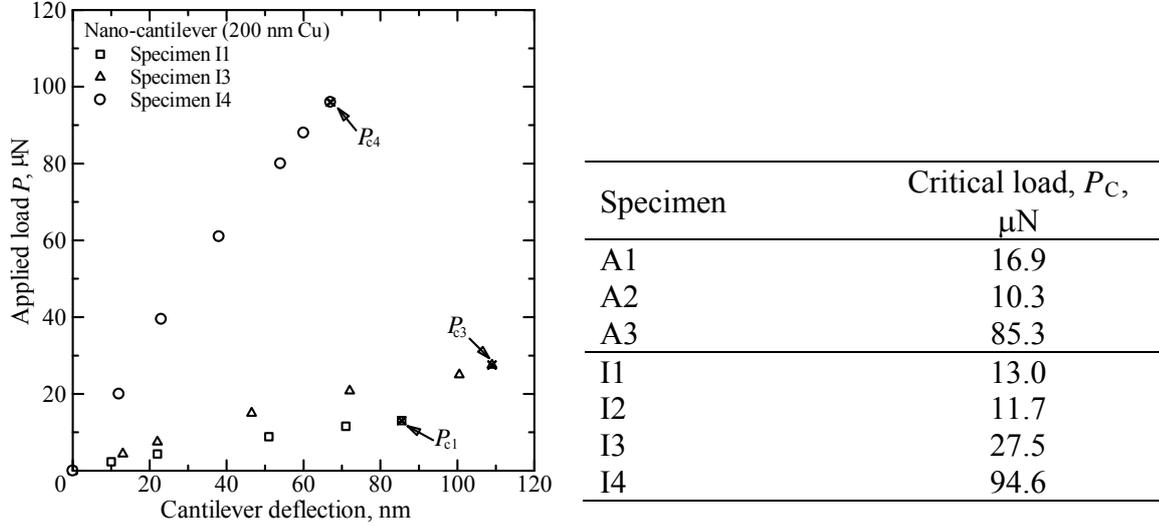
Fig. 3 Relationship between applied load and cantilever deflection for specimens I1, I3, and I4, and the critical loads of all tested nano-cantilever specimens

1.3.2 Modified four-point bend experimental results

Table 1 shows critical loads, P_C, for crack propagation in all the tested specimens. The obtained critical loads are nearly a million times larger than those in the nano-cantilever tests due to the huge difference in specimen dimensions.

Table 1 Critical loads of four-point bend (200 nm Cu) specimens

Specimen No.	Critical load P_C, μN
P1	5.40×10^6
P2	4.86×10^6
P3	4.65×10^6

2. Cohesive zone model

In the exponential CZM [Xu and Needleman (1993)], with increasing interfacial separation, the tractions across the interface increase to reach a maximum, and then decrease, eventually vanishing with complete decohesion.

The interfacial potential is defined as

$$\phi(\Delta_n, \Delta_t) = \phi_n + \phi_n \exp(-\frac{\Delta_n}{\delta_n})\{[1 - r + \frac{\Delta_n}{\delta_n}]\frac{1-q}{r-1} - [q + (\frac{r-q}{r-1})\frac{\Delta_n}{\delta_n}]\exp(-\frac{\Delta_t^2}{\delta_t^2})\}$$ (1)

where $\Delta_t = \sqrt{\Delta_{t1}^2 + \Delta_{t2}^2}$, $q = \phi_t/\phi_n$, and $r = \Delta_n^*/\delta_n$. ϕ_n and ϕ_t are the work of the normal and shear separations; Δ_n and Δ_t are the normal and shear displacement jumps, respectively; δ_n and δ_t are the normal and shear interface characteristic length parameters, respectively. Δ_n^* is the critical magnitude of Δ_n at complete shear separation, where normal traction is zero.

The relations between the interfacial tractions and the potential are given by

$$T_n = \frac{\partial \phi}{\partial \Delta_n}, \text{ and } T_t = \frac{\partial \phi}{\partial \Delta_t}. \tag{2}$$

Substituting Eq. (1) into Eq. (2), we obtain the interfacial tractions as follows

$$T_n = \left(\frac{\phi_n}{\delta_n}\right)\exp\left(-\frac{\Delta_n}{\delta_n}\right)\left\{\frac{\Delta_n}{\delta_n}\exp\left(-\frac{\Delta_t^2}{\delta_t^2}\right) + \frac{1-q}{r-1}\left[1-\exp\left(-\frac{\Delta_t^2}{\delta_t^2}\right)\right]\left(r - \frac{\Delta_n}{\delta_n}\right)\right\} \tag{3}$$

$$T_t = \frac{\phi_n}{\delta_n}\left(\frac{2\delta_n}{\delta_t}\right)\frac{\Delta_t}{\delta_t}\left(q + \frac{r-q}{r-1}\frac{\Delta_n}{\delta_n}\right)\exp\left(-\frac{\Delta_n}{\delta_n}\right)\exp\left(-\frac{\Delta_t^2}{\delta_t^2}\right) \tag{4}$$

The normal and shear cohesive energy (works of normal and shear separations) are related to σ_{max} and τ_{max} by

$$\phi_n = \sigma_{max}\delta_n \exp(1), \quad \phi_t = \sqrt{\exp(1)/2}\,\tau_{max}\delta_t \tag{5}$$

Fig. 4 shows the traction-separation relationships for (a) normal and (b) shear separation.
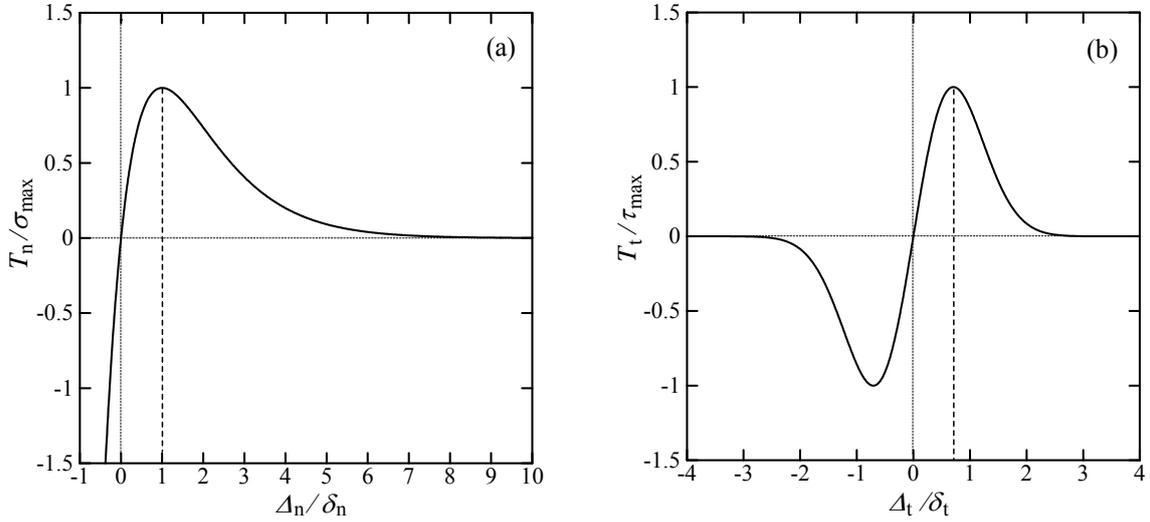
Fig. 4 (a) Normal and (b) shear traction-separation curves for the exponential CZM

3. Determination of CZM parameters

The CZM parameters are determined by calibrating the experimental results for specimen I1, which is one of the nano-cantilever (200 nm Cu) tests. Since the normal stress dominates the crack initiation and propagation in the nano-cantilever and four-point bend specimens, only the CZM parameters for normal separation need to be calibrated, *i.e.* σ_{max} and δ_n.
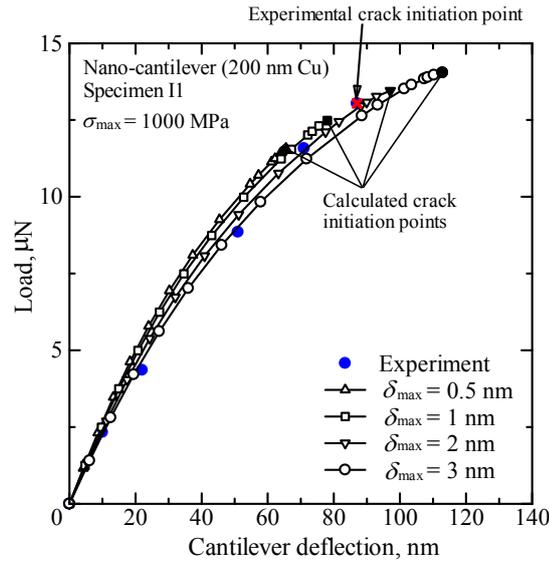
Fig. 5 Effect of interface characteristic length parameter, δ_n, on the calculated load-deflection curves

With σ_{max} = 1000 MPa, the effect of the interface characteristic length parameter, δ_n, is investigated using the experimental data of specimen I1. As shown in Fig. 5, the slope of the calculated load-deflection curve is sensitive to the value of δ_n. With decreasing δ_n, the slope of the calculated curve becomes steep. The simulation for δ_n = 1 nm results in the best correspondence with the experimental data. Then, δ_n = 1 nm is used in subsequent calculations.

Fig. 6 shows the calculated load-deflection curves with different cohesive strengths under δ_n =1 nm. The right figure gives an enlarged view of the square region in the left near the critical load. When the cohesive strength increases, the critical lateral forces become larger. The simulation with σ_{max} = 1060 MPa gives good prediction for the critical lateral force with the experimental results.
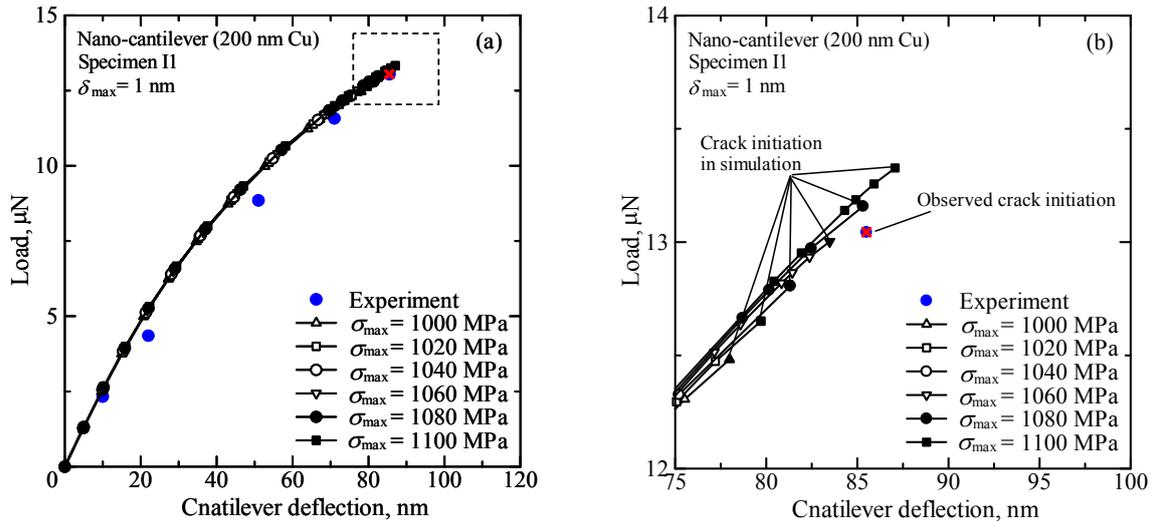
Fig. 6 Effect of the cohesive strength, σ_{max}, on the calculated critical loads for crack initiation

From the calibration, the CZM parameters of the Cu/Si interface are determined to be σ_{max} = 1060 MPa, and δ_n = 1 nm, which give normal cohesive energy ϕ_n = 2.85 J/m^2.

4. Prediction of delamination with CZM parameters

All the experimental and predicted critical loads are compared in Fig. 7. Although the critical load magnitude has a difference of nearly 6 orders, the CZM parameters solely determined by specimen I1 can still universally predict interface cracking along Cu/Si interface cracking regardless of specimen dimensions and film thickness which include the differences in plastic behavior and residual stress. The good prediction both on the crack initiation at interface edge and propagation from a pre-crack tip indicates that the CZM criterion can describe cracking along Cu/Si interfaces with different stress singularities. These prove the versatility of the CZM criterion for the design of micro/nano devices.
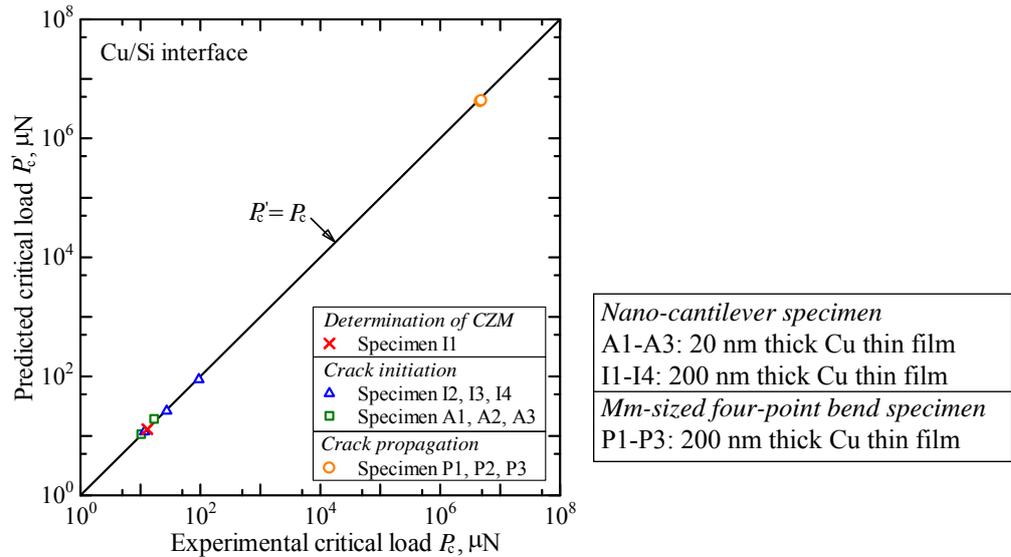
Fig. 7 Comparison of experimental and predicted critical loads of all tested specimens

5. Conclusions

We have investigated the universal applicability of CZM to the initiation and the propagation of interface cracking in nano-cantilever (20 nm and 200 nm Cu) tests and modified four-point bending (200 nm Cu) tests. The results obtained can be summarized as follows:

(1) By calibrating with the experimental results of the nano-cantilever (200 nm Cu) test, the CZM parameters of the Cu/Si interface were determined as follows: cohesive strength σ_{max} = 1060 MPa and interface characteristic length parameter δ_n = 1 nm.

(2) The obtained CZM parameters give excellent prediction of crack initiation at the Cu/Si interface edge in nano-cantilever (20 nm Cu) and (200 nm Cu) experiments regardless of the specimen geometry, plastic behavior and residual stress.

(3) The CZM predicts the crack propagation along the Cu/Si interface in the mm-sized modified four-point bending (200 nm Cu) specimen very well, though the specimen size has a difference of thousands of times. Moreover, this also shows the validity of the CZM parameters for prescribing the interface toughness under different stress singularities.

References

Elices, M., Guinea, G.V., Gomez, J., Planas, J. (2002) The cohesive zone model: advantages, limitations and challenges, *Engineering Fracture Mechanics*, **69**, 137–163.

Hirakata, H., Takahashi, Y., Truong, D.V., Kitamura, T. (2007) Role of plasticity on interface crack initiation from a free edge and propagation in nano-component, *International Journal of Fracture*, **145**, 261–271.

Kitamura, T., Shibutani, T., and Ueno, T. (2002) Crack initiation at free edge of interface between thin films in advanced LSI, *Engineering Fracture Mechanics*, **69**, 1289–1299.

Rice, J.R. (1988) Elastic fracture mechanics concepts for interfacial cracks, Journal of Applied Mechanics **55**, 98–103.

Sumigawa, T., Shishido, T., Murakami, T., Iwasaki, T., Kitamura T. (2010a) Evaluation on plastic deformation property of copper nano-film by nano-scale cantilever specimen, *Thin solid films*, **518**, 6040–6047.

Sumigawa, T., Shishido, T., Murakami, T., Kitamura, T. (2010b) Interface crack initiation due to nano-scale stress concentration, *Materials Science and Engineering A*, **527**, 4796–4803.

Takahashi, Y., Hirakata, H., Kitamura, T., (2007) Quantitative evaluation of plasticity of a ductile nano-component, *Thin Solid Films*, **516**, 1925–1930.

Xu, X.P., and Needleman, A. (1993) Void nucleation by inclusion debonding in a crystal matrix, *Modelling and Simulation in Materials Science and Engineering*, **1**, 111–132.

Nonlinear finite element analysis of Concrete Filled Steel Tube (CFST) columns under projectile impact loading

A.S. Shakir[1, 2*], Z.W. Guan[1], S.W. Jones[1]

[1]School of Engineering, University of Liverpool, Liverpool, L69 3QG, UK

[2]Civil Engineering Department, College of Engineering, Al Muthanna University, Al Muthanna Province, 72001, Iraq

*Corresponding author: alaa1979@liverpool.ac.uk

Abstract

This paper presents a three dimensional nonlinear finite element analysis for Concrete Filled Steel Tube (CFST) columns subjected to lateral impact load. The finite element models were developed using commercial code ABAQUS/Explicit, which were validated against the experimental results. The study has been carried out to examine the influence of several parameters such as the length of the tube and projectile configurations. The Laser Doppler Velocimeter (LDV) and High Speed Camera (HSC) were used to capture the impact velocity and the deflection of the specimens, respectively, which were then implemented in the numerical modelling. It was found that the finite element simulations are in a good agreement with the experimental results, in terms of load-displacement traces and deformation modes. It was also demonstrated that with the increasing the length of the tube the lateral displacement was increased and the impact force was decreased.

Keywords: Concrete filled steel tube, Impact, Failure, Finite element analysis.

Introduction

Concrete Filled Steel Tube (CFST) is increasingly used in many structural applications such as seismic-resistance constructions, high buildings, bridges' piers, decks of railways, and offshore structures [Shanmugam and Lakshmi (2001); He et al. (2011); Sundarraja and Prabhu (2011)]. CFST have more advantages than the conventional reinforced concrete and steel structures, namely high speed of construction work resulted from the omission of framework and the reinforcing bars, low structural costs and conservation the environment [Morino et al. (2001); Morino and Tsuda (2002); Starossek et al. (2008)]. CFST offered a good damping merits and excellent seismically resistant [Kang et al. (2007)].

In recent years, a number of researches had been conducted to study the impact behavior of the CFST members through experimental and theoretical works and finite element analysis. [Bambach et al. (2008); Deng et al. (2012); and Al-Thairy and Wang (2013)]. [Wang et al. (2013)] investigated the impact performance of the concrete filled steel tube members. It has been demonstrated that the lateral deflection and the impact force have been affected by the axial load.

At present with the development of digital computers and numerical techniques, the finite element method (FEM) has emerged as a powerful analytical tool for structures analysis. This is opened spacious world for engineers to model rationally many

aspects of the phenomenological behaviour encountered in CFST [Starossek et al. (2008)]. These aspects include the confinement effect, modelling of cracking and crushing, the behaviour of the materials properties after cracking and crushing and many other properties.

Due to the ability of finite element technique to simulate the behaviour and the failure mode of CFST, many numerical studies were conducted to study the behaviour of CFST with various parameters and different type of loadings.

[Schneider (1998)] presented experimental and a numerical study to investigate the behaviour of CFT under axial compression load. ABAQUS software programme was used in this study. The concrete core was modelled with 20-node brick element while the steel tube modelled with 8-node shell element, a gap element was used to simulate the contact behaviour between steel and concrete core. Uniaxial stress-strain curve with the available model in ABAQUS was used to simulate the behaviour of concrete with confinement effect. A comparison between the numerical and experimental results was made and it showed a good agreement for the columns strength and the failure mode.

[Yu et al. (2010a)] presented a nonlinear finite element analysis study on concrete filled steel tubular frame structures undergo dynamic load in fire conditions. Steel I-section beams and circular CFST columns composed the frame structures. The numerical results were in good agreement with the test results. It was concluded that the bending moment of the CFST subjected to fire was highly affected and reduced due to the internal forces redistribution in the frame.

[Deng et al. (2012)] carried out a theoretical and numerical research to study the CFSTs and post-tensioned CFSTs under flexural load. The concrete properties were modelled using Drucker-Prager plasticity model, while the behaviour of steel tube was modelled as elastic-perfectly plastic. The results predicted from the theoretical sectional analysis and finite element model were compared with the test results. It was found that both of the theoretical sectional and finite element analyses were efficient to predict the ultimate moment capacity.

A study of three test series had been conducted by [Yousuf et al (2012); (2013); and (2014)]. They investigated the transverse impact resistance of the hollow and concrete filled mild and stainless steel square tube columns. This study included both static and dynamic tests. The performance of the hollow and concrete filled mild and stainless steel tube columns was studied through finite element analysis using the software ABAQUS. The numerical results showed a good agreement with the experimental findings. They found that the impact energy was improved by using the stainless steel columns and the axial compressive load affects the static and impact strength especially for the stainless steel tubes.

The main aim of this study is to perform a numerical simulation of CFST columns subjected to lateral impact loading using the nonlinear finite element programme ABAQUS.

Experimental work

Twenty two specimens of CFST column with outer diameter 114.3 mm and 3.6 mm wall thickness have been tested under lateral impact loading with height 2.6 m and

maximum mass 107.5 kg as shown in Fig. 1. The specimens were in three different lengths (686, 1029, and 1543) mm short, medium and long tubes respectively. To examine the indenter shape effect on the behavior of the CFST, two types of indenter were used in this test; the first one is spherical indenter with three different diameters 60 mm (BI), 40 mm (MI) and 20 mm (SI). The second type is Flat Indenter (FI) with 40 mm square section. A high strength steel clamp was used to provide a fixed ends for the specimens to simulate the real case of the fixed ends column. A Laser Doppler Velocimeter (LDV) with the Dantec Flowlite LDV system was used to obtain the impact force and local indentation in this study. A High Speed Camera (HSC) was used to capture the local and overall displacement and the mode of failure of the specimen. It has been concluded that with increasing the specimen's length, the peak force decreased and the local displacement increased and the increasing of the spherical diameter leads to increase the peak force and reducing the local indentation.
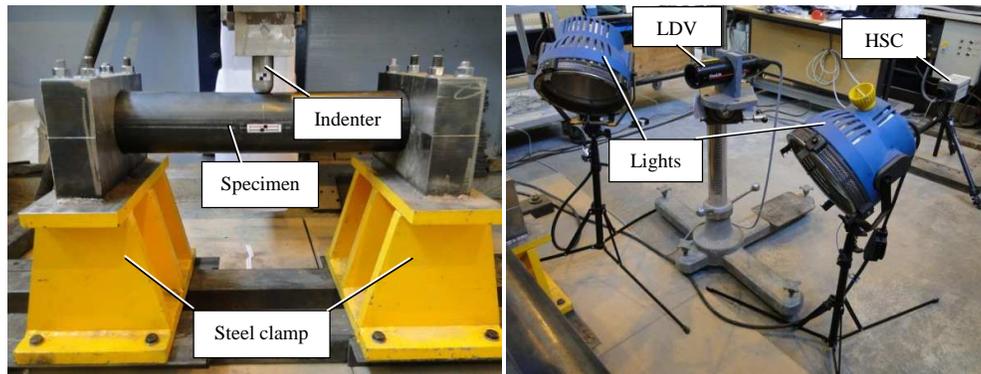
Figure1. The experimental work setup with data recording system

Finite element modeling

General

The commercial finite element program ABAQUS was used to simulate the behaviour of the CFST columns under lateral impact loading. There are two main constituent materials considered to model the impact behaviour of CFST column. The materials are the concrete core and the steel tube. In addition, the type and the properties of the contact between the steel tube and the concrete core are very important to simulate the CFST.

Finite element type and boundary conditions

Due to the symmetry, three dimensional models are used to model a quarter of CFST columns as shown in Fig. 2. The three dimensional eight nodes solid element with reduced integration C3D8R is used to model the concrete core and the steel tube while the R3D4 is used to model the indenter and the steel clamp. To provide accurate results, many models with different mesh sizes have been tried to find out the moderate element size with reasonable computational time. To simulate the experimental test conditions, the clamp was restrained in all direction to provide a fixed case while the tube end was free in the axial direction of the columns and fixed with the other directions. The indenter was restrained against all degree of freedom

except for the vertical displacement. Symmetric boundary conditions were applied on the symmetric plans.
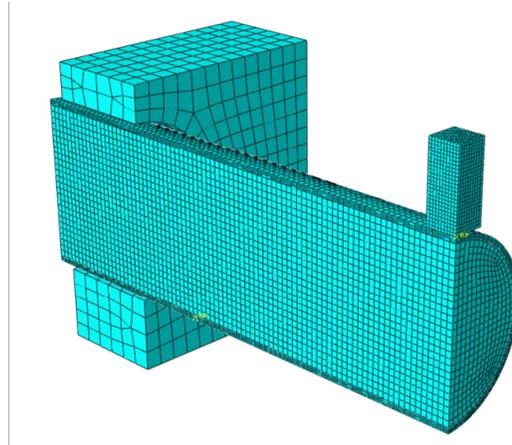
Figure2. Finite element mesh of CFST

Materials modelling

The stress-strain curve of the steel tube is assumed to be elastic perfectly plastic, the elastic modulus and Poisson's ratio were 200 GPa and 0.3 respectively. The (*Plastic option) in ABAQUS is used in steel material model

The Concrete Damage Plasticity available model in ABAQUS 6.13 is used in this study to describe the behaviour of the confined concrete in CFST columns under impact loading with both tension stiffening and compression hardening definition. Both of the elastic and plastic parts are included for the concrete model.

The confined concrete stress-strain models described by [Hu et al. (2003)] is used in this study. [Ellobody et al. (2006) and Dai and Lam (2010)] were adopted this model in their numerical analysis and the finite element results showed a good agreement with the experimental work data. The difference between the stress strain relationship for confined and unconfined concrete is showed in Fig. 3
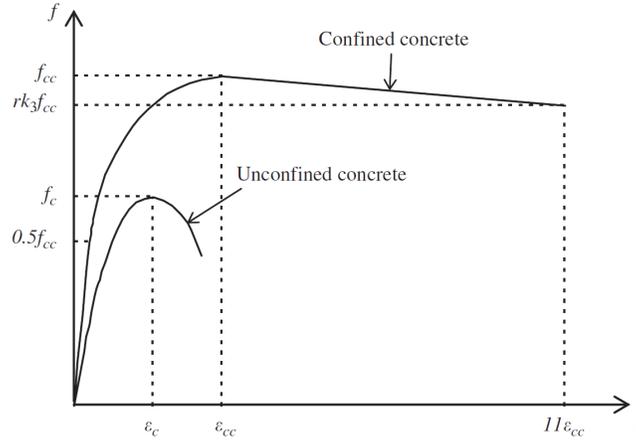
Figure 3. Stress-Strain relationship for confined and unconfined concrete (Dai and Lam, 2010)

Where f_{ck} is the cylinder compressive strength of unconfined concrete which is equal to eighty percent of the unconfined cube strength $0.8f_{ck, cube}$. While f_{cc} is the cylinder compressive strength of confined concreter. ε_{ck} and ε_{cc} are the corresponding strain for f_{ck} and f_{cc} respectively. The axial compressive strength for confined concrete structural hollow section can be predicted by the proposed relationship between f_{cc} and f_{ck} [Mander and Priestley (1988)].

$$f_{cc} = f_{ck} + k_1 f_l \tag{1}$$

$$\varepsilon_{cc} = \varepsilon_{ck}\left(1 + k_2 \frac{f_l}{f_{ck}}\right) \tag{2}$$

[Richart et al. (1928)] suggested the value of the constant k_1 and k_2 to be 4.1 and 20.5 respectively and the strain of unconfined concrete ε_{ck} can be taken 0.003 [ACI (1999)]. Based on the formulas proposed by [Hu et al. (2003)], the lateral force f_l value can be predicted:

$$f_l/f_y = 0.043646 - 0.000832(D/t) \qquad for\ 21.7 \le D/t \le 47 \tag{3}$$

$$f_l/f_y = 0.006241 - 0.0000357(D/t) \qquad for\ 47 \le D/t \le 150 \tag{4}$$

Where D is the outer tube diameter, t is wall thickness of the tube and f_y is the yield strength of the tube. [Hu et al. (2003) and Ellobody et al. (2006)] suggested the proportional limit stress for the linear part in the confined concrete stress-strain curve (elastic) to be $0.5f_{cc}$. The modulus of elasticity for the confined concrete E_{cc} can be predicted by the [ACI (1999)] empirical formula which expressed as $E_{cc} = 4700\sqrt{f_{cc}}$ MPa [Dai and Lam (2010)]. The nonlinear portion of the curve represents the compressive strength f between the elastic limit $0.5f_{cc}$ and the maximum compressive strength f_{cc}. [Saenz (1964)] suggested formula to predict f:

$$f = \frac{E_{cc}\varepsilon}{1 + (R + R_E - 2)\left(\frac{\varepsilon}{\varepsilon_{cc}}\right) - (2R - 1)\left(\frac{\varepsilon}{\varepsilon_{cc}}\right)^2 + R\left(\frac{\varepsilon}{\varepsilon_{cc}}\right)^3} \tag{5}$$

Where

$$R_E = \frac{E_{cc}\varepsilon_{cc}}{f_{cc}} \quad ,R = \frac{R_E(R_\sigma-1)}{(R_\varepsilon-1)^2} - \frac{1}{R_\varepsilon} \quad ,R_\sigma = R_\varepsilon=4$$

The start of the third portion of the strain-stress curve is f_{cc} while its end is f_u. ε_u $=11\varepsilon_{cc}$ is the correspondence strain at $f_u = rk_3f_{cc}$ [Hu et al. (2003); Ellobody and Young (2006b); Ellobody et al. (2006); Dai and Lam (2010)]. For the circular steel tube section with $21.7 \leq \left(D/t\right) \leq 150$, the value of the parameter k_3 can be obtained from the proposed by [Hu et al. (2003)].

$$k_3 = 1 \qquad for\ 21.7 \leq \left(D/t\right) \leq 40 \tag{6}$$

$$k_3 = 0.0000339\left(D/t\right)^2 - 0.0100085\left(D/t\right) + 1.3491 \quad for\ 40 \leq \left(D/t\right) \leq 150 \tag{7}$$

According to [Giakoumelis and Lam (2004); Ellobody and Young (2006b); Ellobody et al. (2006); Dai and Lam (2010)] the value of r with compressive cube strength of 30 MPa can be taken as 1.0 while with compressive strength of 100 MPa as 0.5, a linear interpolation used with compressive strength between 30 MPa and 100 MPa.

The tension stiffening for the concrete is defined as a displacement using the model proposed by [Li et al. (2002)] as shown in Fig. 4

$$\sigma = f_t \left\{1 - exp\left[-\left(\frac{0.05}{w_c/w_{ccr}}\right)^{1.3}\right]\right\} \tag{8}$$

Where f_t is the concrete tensile stress and it is obtained from the equation:

$$f_t = 0.34\sqrt{fc'} \tag{9}$$

Where and fc' is the compressive strength of concrete.

The w_c and w_{ccr} are the crack width and the critical crack width respectively. 1.5 mm is taken for the critical crack width in this study. The adopted concrete Poisson's ratio is 0.2.
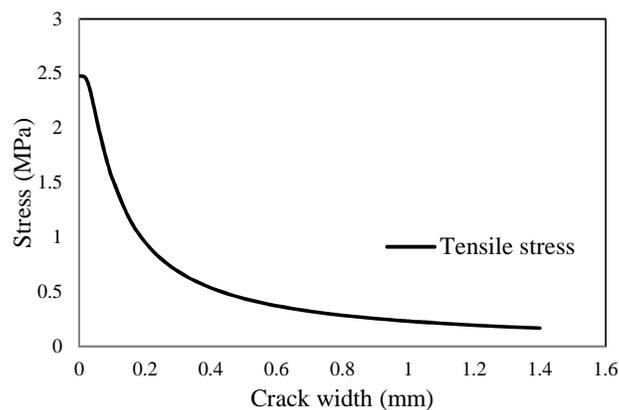
Figure 4. Concrete material behavior model in tension

The surface to surface contact was defined between the indenter and the steel tube and between the steel tube and the clamp with hard contact option and coefficient of friction 0.01 and 0.1 for the interaction between the indenter and the tube and

between the clamp and steel tube respectively. To model to interaction between the concrete core and the steel tube a pressure contact was defined. The coefficient of friction adopted in this study was 0.2.

Results and discussion

The results of the numerical analysis for CFST columns with different indenter are compared with those obtained from the experimental tests and the curves of the impact force versus local indentation for these specimens are plotted against the experimental results in Fig. 5, 6 and 7. The deformed shapes for these columns are also examined. In General it can be seen that the numerical results show good agreement with the experimental data. For the short columns tested with (BI) in Fig 5 (a) the difference between the results was only 2.2 % and 8.4% for the peak force and local displacement respectively while the difference for the columns tested with medium indenter was 0.81% for the peak load and 5.58% for the local displacement as shown in Fig. 5 (b). The impact force for short CFST column tested with SI using the finite element analysis was 232.31 kN at a local displacement 16.67 mm compared with the experimental values of 220.13 kN and 16.36mm as shown in Fig 6 (a). In terms of the local displacement, stiffer results were obtained from the numerical analysis compared with the experimental data for the short columns tested with FI which were 9.89 mm and 11.1 mm for the numerical and experimental results respectively. However the maximum difference between the peak forces obtained from the finite element analysis and the one evaluated from the experimental test was 7.37% as shown in Fig. 6 (b).

(a)   (b)
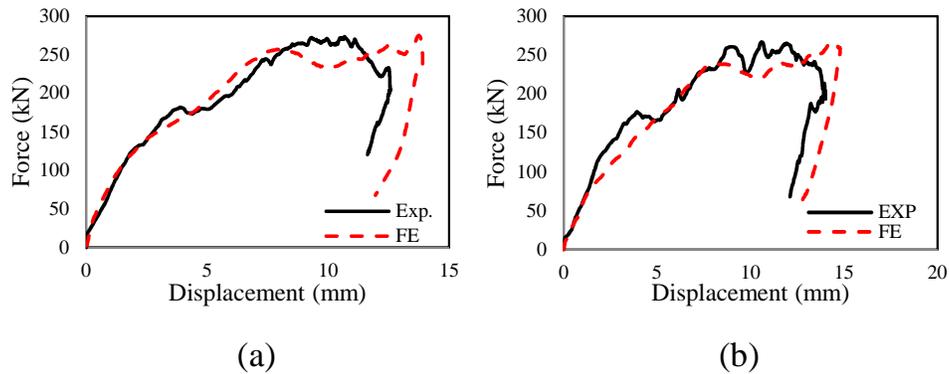
Figure (5) Load-deflection curve for the short CFST column with (a) BI (b) MI

(a)   (b)
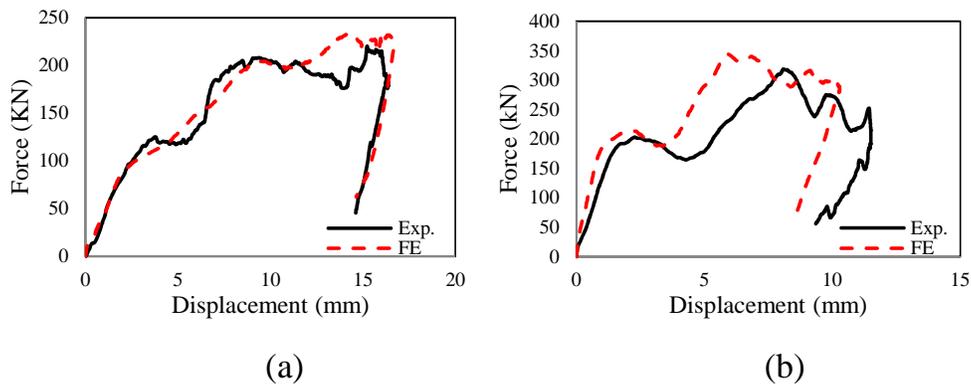
Figure (6) Load-deflection curve for the short CFST column with (a) SI (b) FI

Fig. 7 (a) and (b) show the good agreement between the experimental and numerical results for the medium and long CFST columns respectively. The experimental and numerical peak force and the local displacement for medium CFST columns were 193.68 kN, 18.78mm and 187.87 kN and 18.54 respectively. The numerical simulation for the long CFST columns was able to capture the beak force reduction which resulted from the vibration of the tested columns during the test with same natural frequency of the whole system tested as shown in Fig. 7 (b).

(a)   **(b)**
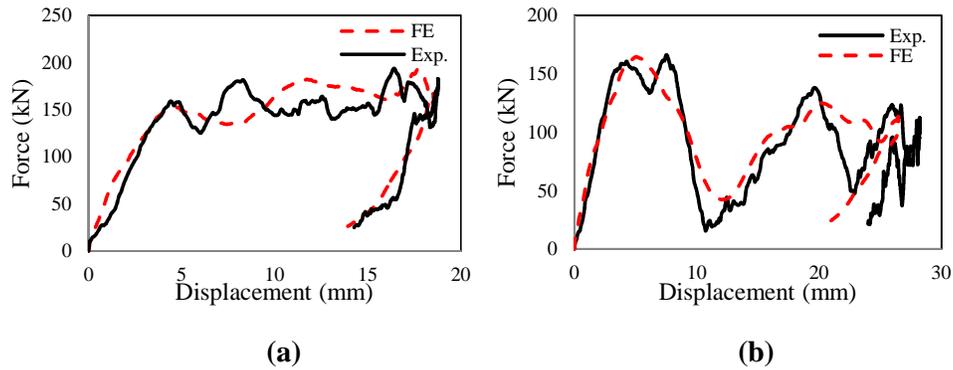
Figure (7) Load-deflection curve for the short CFST column with (a) medium tube (b) long tube

From the comparison between the failure mode of the experimental test and numerical model for the short, medium and long CFST columns, the numerical failure modes were in good agreement with those obtained from the experimental tests as can be seen from Figs. 8 (a), (b) and (c).

(a)   (b)   (c)
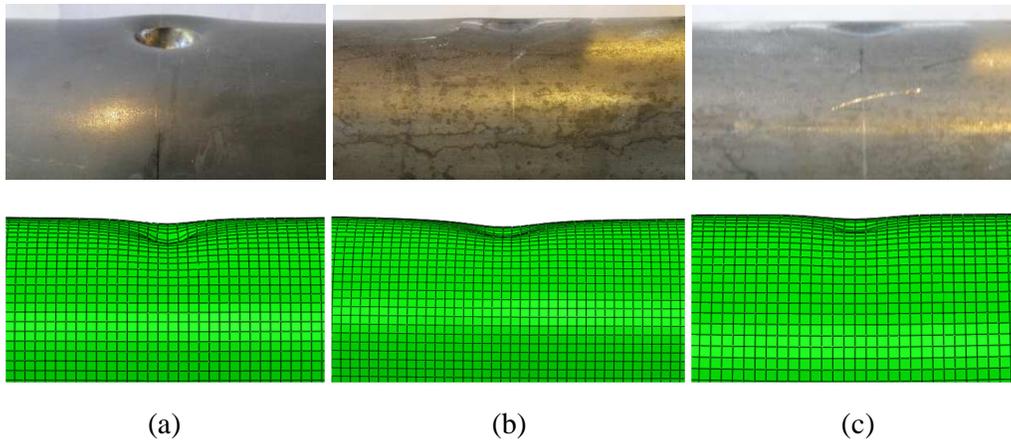
Figure (8) Experimental and numerical failure mode (a) Short tube (b) Medium tube (c) Long tube

5 Conclusions

The paper presents the numerical analysis of CFST columns under lateral impact loading. The concrete was molded using an equivalent stress-strain curve for the confined concrete while the steel tube was molded as elastic-perfectly plastic material. The predicted impact force, local displacement and the failure mode using the finite element analysis were compared with those evaluated from the experimental tests of the CFST columns. The comparison between the numerical experimental and

finite element results showed a very good agreement for both the force-displacement curves and the deformed shape.

Acknowledgment

References

ACI 1999. Building Code Requirements for Structural Concrete and Commentary. *ACI 318-95.* Detroit (USA): American Concrete Institute.

BAMBACH, M. R., JAMA, H., ZHAO, X. L. & GRZEBIETA, R. H. 2008. Hollow and concrete filled steel hollow sections under transverse impact loads. *Engineering Structures,* 30**,** 2859-2870.

DAI, X. & LAM, D. 2010. Numerical modelling of the axial compressive behaviour of short concrete-filled elliptical steel columns. *Journal of Constructional Steel Research,* 66**,** 931-942.

DENG, Y., TUAN, C. Y. & XIAO, Y. 2012. Flexural Behavior of Concrete-Filled Circular Steel Tubes under High-Strain Rate Impact Loading. *Journal of Structural Engineering,* 138**,** 449-456.

ELLOBODY, E. & YOUNG, B. 2006b. Nonlinear analysis of concrete-filled steel SHS and RHS columns. *Thin-Walled Structures,* 44**,** 919-930.

ELLOBODY, E., YOUNG, B. & LAM, D. 2006. Behaviour of normal and high strength concrete-filled compact steel tube circular stub columns. *Journal of Constructional Steel Research,* 62**,** 706-715.

GIAKOUMELIS, G. & LAM, D. 2004. Axial capacity of circular concrete-filled tube columns. *Journal of Constructional Steel Research,* 60**,** 1049-1068.

HE, D., DONG, J., WANG, Q. & CHEN, X. Mechanical behaviour of recycled concrete filled steel tube conlumns strengthened by CFRP. Multimedia Technology (ICMT), 2011 International Conference on, 26-28 July 2011 2011. 1110-1113.

HU, H. T., HUANG, C. S., WU, M. H. & WU, Y. M. 2003. Nonlinear analysis of axially loaded concrete-filled tube columns with confinement effect. *Journal of Structural Engineering,* 129**,** 1322-1329.

KANG, J. Y., CHOI, E. S., CHIN, W. J. & LEE, J. W. 2007. Flexural Behavior of Concrete-Filled Steel Tube Members and Its Application. *International Journal of Steel Structures,* 7**,** 319-324.

LI, Q., DUAN, Y. & WANG, G. 2002. Behaviour of large concrete specimens in uniaxial tension. *Magazine of Concrete Research,* 54**,** 385-391.

MANDER, J. B. & PRIESTLEY, M. J. N. 1988. Theoretical stress-strain model for confined concrete. *Journal of Structural Engineering,* 114**,** 1804-1826.

MORINO, S. & TSUDA, K. 2002. Design and construction of concrete-filled steel tube column system in Japan. *Earthquake Engineering and Engineering Seismology,* 4**,** 51-73.

MORINO, S., UCHIKOSHI, M. & YAMAGUCHI, I. 2001. Concrete-filled steel tube column system-its advantages. *International Journal of Steel Structures,* 1**,** 33-44.

RICHART, F. E., BRANDTZAEG, A. & BROWN, R. L. 1928. A study of the failure of concrete under combined compressive stresses. *Bulletin ; no. 185* University of Illinois. Engineering Experiment Station.

SAENZ, L. P. 1964. Discussion of "Equation for the stress-strain curve of concrete," by Desayi and Krishnan. *ACI journal,* 61**,** 1229-1235.

SCHNEIDER, S. P. 1998. Axially loaded concrete-filled steel tubes. *Journal of Structural Engineering,* 124**,** 1125-1138.

SHANMUGAM, N. E. & LAKSHMI, B. 2001. State of the art report on steel–concrete composite columns. *Journal of Constructional Steel Research,* 57**,** 1041-1080.

STAROSSEK, U., FALAH, N. & LOHNING, T. 2008. Numerical Analyses of the Force Transfer in Concrete-Filled Steel Tube Columns. *The 4th International Conference on Advances in Structural Engineering and Mechanics(ASEM08)***,** 2651-2666.

SUNDARRAJA, M. C. & PRABHU, G. G. 2011. Investigation on strengthening of CFST members under compression using CFRP composites. *Journal of Reinforced Plastics and Composites,* 30**,** 1251-1264.

WANG, R., HAN, L. & HOU, C. 2013. Behavior of concrete filled steel tubular (CFST) members under lateral impact: Experimental and FEA model. *Journal of Constructional Steel Research,* 80**,** 188-201.

Yousuf, M., Uy, B., Tao, Z., Remennikov, A., Liew, R. 2012. Behavior and resistance of hollow and concrete filled mild steel columns due to transvers impact loading. *Australian journal of stractural engineering*, 13, 65-80.

Yousuf, M., Uy, B., Tao, Z., Remennikov, A., Liew, R. 2013. Transverse impact resistance of hollow and concrete filled stainless steel columns. *Journal of construction steel research*, 82, 177-189.

Yousuf, M., Uy, B., Tao, Z., Remennikov, A., Liew, R. 2014. Impact behavior of precompressed hollow and concrete filled mild and stainless steel columns. *Journal of construction steel research*, 96, 54-68.

YU, M., ZHA, X., YE, J. & LI, Y. I. 2010a. Fire responses and resistance of concrete-filled steel tubular frame structures. *International Journal of Structural Stability and Dynamics,* 10**,** 253-271.

Modelling of Energy-absorbing Behaviour of Metallic Tubes Reinforced Polymer Foams

*R.A Alia[1,2], Z. Guan[1], W.J. Cantwell[1,3], A.K. Haldar[1]

[1]School of Engineering, University of Liverpool, Liverpool, L69 3GH.U.K.
[2]Universiti Kuala Lumpur Institute of Product Design & Manufacturing, 56100. Malaysia.
[3]Aerospace Research and Innovation Center (ARIC), Khalifa University of Science, Technology and Research (KUSTAR), Abu Dhabi, UAE.

*Corresponding author: hsaalia@liv.ac.uk

Abstract

This paper presents the findings of a research study investigating the energy-absorbing characteristics of polymer foams reinforced with steel and aluminium tubes. Initial attention focused on establishing the influence of tube diameter on the specific energy absorption (SEA) and the failure characteristics of the tubes. In the next stage of the investigation, the tubes were embedded in a range of polymer foams in order to establish the influence of foam density on the crush behaviour of these lightweight structures. The specimens were tested under quasi-static tests at a loading rate of 1mm/minute and the numerical analysis has been performed by using Abaqus software that was validated against the experimental results. The numerical metal tubes and foams were subjected to axial quasi-static loading similar to the experimental condition. The simulation results indicated that failure response of polymer foams reinforced by metallic tubes under quasi-static axial compression were in good agreement with the experimental results.

Keywords: Energy absorption, Numerical analysis, Quasi-static, Aluminium, Steel, Foams, Failure mechanisms.

Introduction

In recent years there has been an increased occurrence to develop high speed, energy-efficient transport systems. One of the important aspects to be improved in crashworthiness is the ability to absorb the impact energy during a crash. A considerable amount of research has thus been undertaken to develop effective impact energy absorption systems for use in high speed vehicles. The key approach of strengthening a structure is to dissipate the crush energy in order to prevent injuries and reduce damages. The metallic tube structures subject to buckling under axial loading are excellent for this purpose due to their low cost, easy fabrication, and high efficiency in absorbing energy. Extensive investigations previously carried out have revealed that various parameters, such as material properties, cross-sectional geometries, diameter to thickness ratio and the loading condition can affect the energy absorption capability of metallic tubes [Alexander (1960); Jones (2010) and McGregor et al. (1993)]. During the crush of metallic tubes, this energy is mainly absorbed due to the irreversible plastic deformation mechanisms that dominate the buckling process.

Several experiments on the impact of thin-walled tubes of circular and square cross section with different sizes have been performed by Jones [Jones (2010)]. Jones has observed that a circular shaped tube is the most efficient geometry in absorbing energy. Alexander developed theoretical study on the calculation of mean collapse load for circular tubes in concertina collapse mode [Alexander (1960)]. In another study on the effect of strain rate, McGregor et al. [McGregor et al. (1993)] found that strain rate effect on the mechanical properties is insignificant in the range of velocity 7–9 m/s. A number of recent researchers have investigated the axial crushing of circular aluminum tubes by conducting experimental and numerical modelling [Pled et al. (2012); Al Galib and Limam (2004) and Karagiozova et al. (2005)]. Pled et. al [Pled et al. (2012)] has reported that the boundary conditions will influence the crushing mode of circular aluminium tubes. Although a considerable amount of studies has been dedicated to study the crushing of thin walled tubes with different materials or geometries, little attention has been focused on the research of the using metallic tubes as reinforcement for foams in sandwich structures. This paper studies the numerical simulation of the axial crushing of metal circular tubes for use in lightweight energy-absorbing structures. The final section of this paper verifies the results between experimental data and numerical analysis.

Experimental Procedures

The primary aim of this research study was to investigate the energy-absorbing characteristics of tube-reinforced foams, similar to that shown in Fig. 1b. However, prior to testing the reinforced foams, attention focused on establishing the influence of the geometry of the individual metal tubes on their resulting energy-absorbing characteristics. The effect of varying the ratio of the inner diameter of the tube to its thickness, D/t, on energy absorption was then investigated by conducting compression tests on a range of aluminium and steel tubes, details of which are given in Table 2. Here, three different sizes of tubing were considered, with outer diameters ranging from approximately 12.6 mm to 25.4 mm. The values of D/t for the tubes ranged from 5.2 to 13.1 for the aluminium alloy tubes and 5.5 to 13.1 for the steel tubes. Prior to testing, the tubes were cut to a length of 20 mm and ground at both ends to ensure that they were parallel.

Crushing tests were conducted on individual tubes using an Instron 4505 universal test machine. Each test was undertaken at a crosshead displacement rate of 1 mm/minute and interrupted when the tube was fully crushed, i.e. to a point beyond which the force started to increase rapidly. The specific energy absorption (SEA) of the tubes was then determined by dividing the energy under the load-displacement trace up to densification (bottoming-out displacement) by the mass of the sample.

Individual tubes were then embedded into crosslinked PVC foams with densities ranging from 38.3 to 224 kg/m^3. Details of the tubes embedded into foams samples are given in Table 3. In preparation for these tests, a 12.6 mm diameter hole was drilled into a 50 mm square block of thickness 20 mm and either a steel or aluminium tube with an outer diameter of 12.6 mm and a length of 20 mm, was inserted into the hole. The tube/foam combinations were subsequently loaded in compression at a crosshead displacement rate of 1 mm/minute.

Finite Element Modelling

Finite Element models were developed using the commercial finite element Abaqus software package to simulate the crushing of metal tubes individually and embedded in foams. Here, quarter model of the structures were constructed as the structures are symmetrical in x-axis and z-axis as shown in Fig. 1a. A series of three aluminium and three steel circular tubes with diameters ranging from 12.62 mm to 25.40 mm and 20 mm height were simulated. The details of tubes are shown in Table 2. The tubes were modelled by using 8-node 3-D deformable solid and were extruded to a length of 20 mm in y-direction. Two square plates of 30 mm x 30 mm x 1mm were created to represent the upper and lower platen. The square plates were initially defined as 3-D discrete rigid and converted to shell elements. A reasonable mesh size of 1 mm and three elements through the thickness were defined for the tube models. In the latter stage of this research program, one metal tube of 12.62 mm was inserted into three foam materials with densities of 38.3, 90.4 and 224 kg/m^3. The foam materials were modelled as 8-node 3-D deformable solid and the mesh size used was the same as tubes which is 1 mm.

(a) (b)
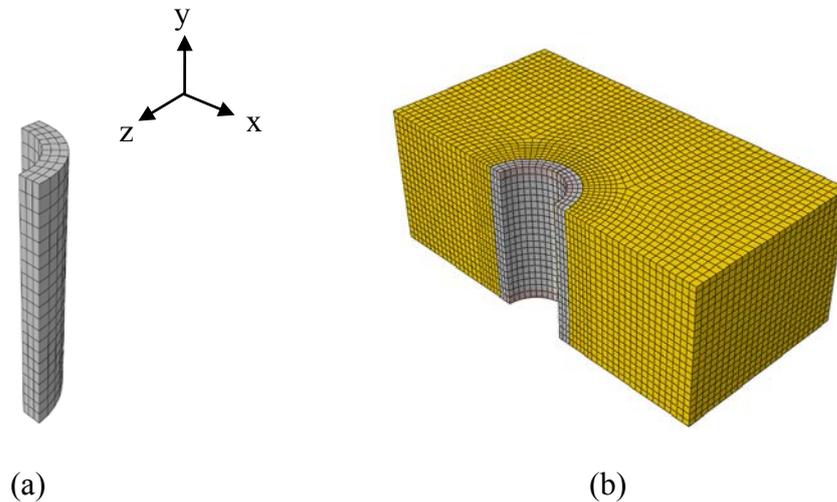
Figure 1. A quarter model of a tube and the cross-section view of a 12.62 mm metal tube in a foam block.

The isotropic elastic-plastic material model has been assumed as the mechanical properties of the metallic tubes. The mechanical properties of these metals are presented in Table 1 and their plastic stress-plastic strain diagrams are illustrated Fig. 2. The plastic yield stress and strain values were obtained from the true stress-true plastic strain curves of the metals. The foams were modelled by using crushable foam plasticity definition that was developed by Deshpande and Fleck (Deshpande and Fleck 2000). The mechanical material properties of foams were obtained from stress-strain curve of the foams. To yield accurate numerical results, it is important to define the required interactions of the structure to prevent interpenetration during the crushing process. Firstly, the top and bottom surfaces of the structure were selected to interact by Surface-to-surface definition. Next, the self-interaction between tube surfaces to itself was defined. This was done by selecting the General Contact option and allowing the outer tube surface to interact with itself.

For the next stage of simulation, a metal tube was embedded in foam, the interaction between tube and foam has to be considered. In the same General Contact option, the contact interaction of tube and foam surfaces was defined. The friction coefficient was set to 0.1 for all contact surfaces between tube and foam [Pled et al. (2012)]. In all simulations of axial crushing, the bottom plate was fixed to be stationary. The axial crushing of the structure was done by moving the top rigid platen in downward direction but constrained in all other degrees of freedom. The top and bottom end of the sample were set to be free to deform in all directions.

Table 1. Mechanical properties of metals.

Property	Aluminium Alloy 6063-T6	Mild Steel
Density, ρ [kg/m^3]	2543	7966
Young's modulus, E [GPa]	70.4	200
Yield stress, σ [MPa]	218	277
Tensile strength [MPa]	237	399
Poisson's Ratio, υ	0.33	0.33

(a)   (b)
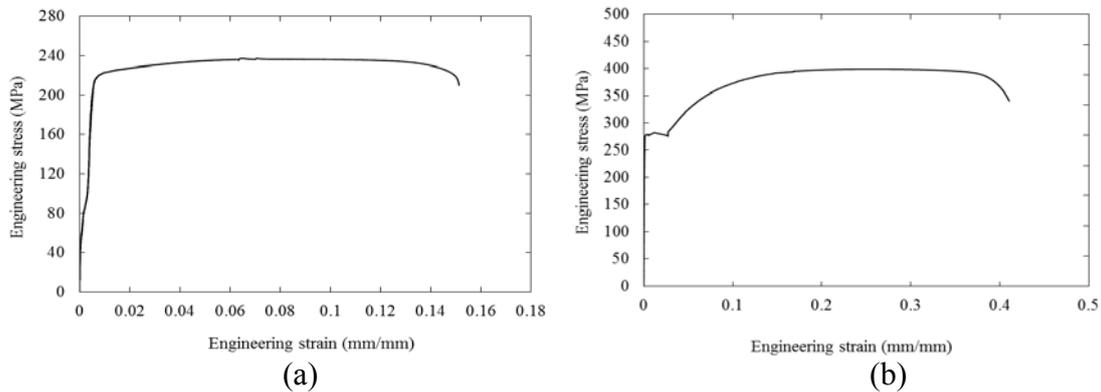
Figure 2. Engineering stress-strain curve following a tensile test on (a) 12.62 mm diameter (D/t = 5.21) aluminium tube and (b) 12.62 mm diameter (D/t = 5.51) steel tube.

Results and Discussion

The initial part of this investigation focused on understanding the influence of tube diameter on the energy-absorbing characteristics of the individual metal tubes. The experimental axial crush of tubes results was compared to verify the presented finite element Abaqus modelling. Fig. 3 shows the typical experimental and Abaqus simulation load-displacement traces following compression tests on 20 mm long aluminium tubes having different D/t ratios. From the two graphs, it is clear that the simulation exhibits very close results of load-displacement to the experimental curves. It is evident from the compressed tubes in both experimental and simulation that the tubes deformed by forming two rings which is known as concertina failure mode. This asymmetric failure mode was also observed by Al Galib and Limam [Al Galib and Limam (2004)] for crushing of A6060 Aluminum tubes.

Table 2 shows the variation of SEA with the tube D/t ratio for both the aluminium and steel tubes. The energy-absorbing characteristics of the aluminium tubes have been shown to be superior to those of their steel counterparts. Here, it is also clear that the energy-absorbing capability of the tubes decreases rapidly with increasing D/t. The specific energy absorption obtained from Abaqus simulation results are very close to the experimental results with percentage of difference ranging between 1.2 to 9.7%.

(a) (b)
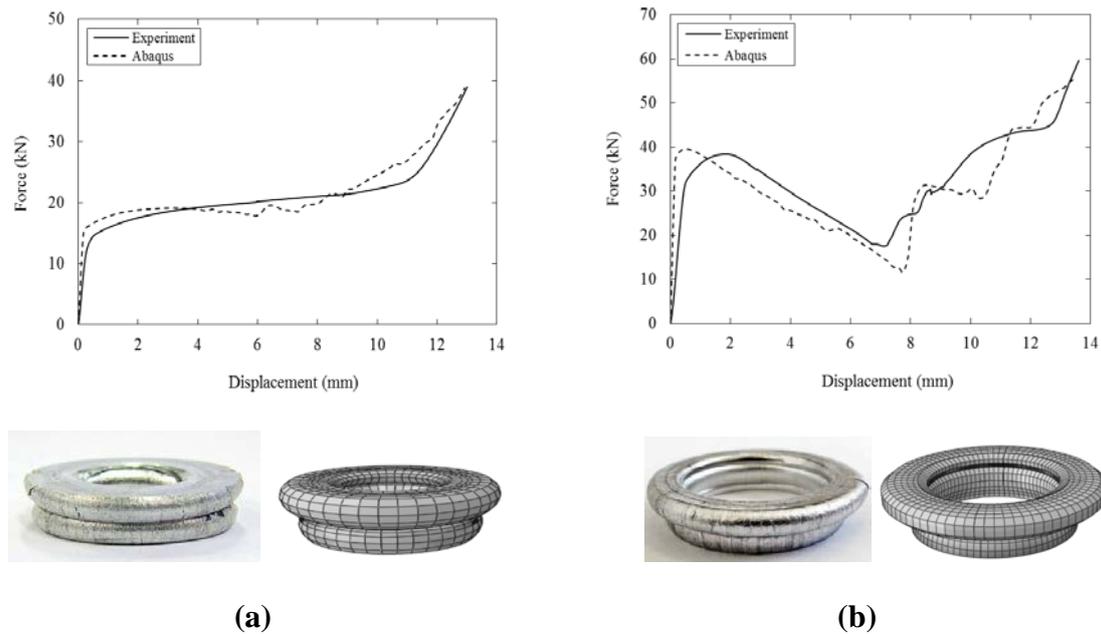
Figure 3. Load-displacement traces and the deformed aluminium tubes from experimental and Abaqus simulation results of (a) 12.62 mm diameter (D/t = 5.21) and (b) 25.40 mm diameter (D/t = 13.12).

Table 2. Summary of the geometrical and specific energy absorbing characteristics of the 20 mm long aluminium and steel tubes.

Specimen ID	Outer diameter, D_o (mm)	Inside diameter, D (mm)	Thickness, t (mm)	D/t	Experiment SEA (kJ/kg)	Abaqus SEA (kJ/kg)	Experiment to Abaqus (%)
Alu12	12.62	9.12	1.75	5.21	70.07	70.89	1.2
Alu16	16.00	12.36	1.82	6.79	63.47	57.96	8.7
Alu25	25.40	22.04	1.68	13.12	52.96	48.08	9.2
Ste12	12.62	9.26	1.68	5.51	41.46	40.75	1.7
Ste16	15.78	12.42	1.68	7.39	36.94	33.50	9.3
Ste25	25.40	22.04	1.68	13.12	24.12	21.78	9.7

The next stage of this study discusses the energy-absorbing characteristics of foams reinforced with relatively thick metal tubes tested at quasi-static loading. Tubes with low values of D/t (and therefore higher values of SEA) were embedded in a range of polymer foams with a view to developing lightweight energy-absorbing structures.

Fig. 4 shows the load-displacement traces and the compressed tubes embedded in PVC C70.200 foam for aluminium and steel tubes. As before, the aluminium-based systems offered superior properties to the steel-based materials. It has been shown that the foam does not modify the energy-absorbing capability of the embedded tubes. The traces from the Abaqus simulation agreed well with the experimental results. It shows that the load-displacement of tubes embedded in foams is resulted from a direct combination of forces from tubes and foams compressed individually.

| (a) | (b) |
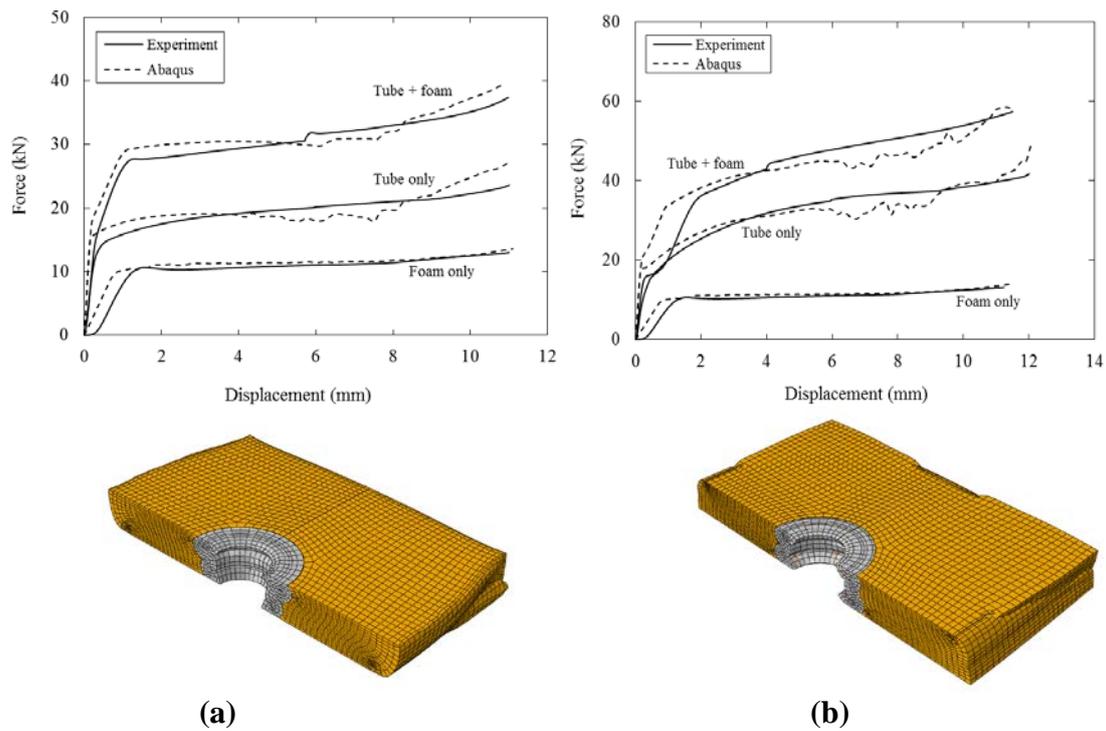
Figure 4. Load-displacement traces and the deformed 12.62 mm diameter tubes embedded in C70.200 foams from experimental and Abaqus simulation results for (a) aluminium and (b) steel.

Table 3. Summary of the SEA following tests on the 20 mm long (diameter = 12.62 mm) of aluminium and steel tubes with foam densities.

Specimen ID	PVC Foam	Foam Density (kg/m³)	Experiment SEA (kJ/kg)	Abaqus SEA (kJ/kg)	Experiment to Abaqus (%)
Aluf40	C70.40	38.3	68.43	71.49	4.5
Aluf55	C70.55	56.0	68.56	70.74	3.2
Aluf200	C70.200	224.0	69.50	70.40	1.3
Stef40	C70.40	38.3	41.2	40.3	2.1
Stef55	C70.55	56.0	43.1	42.5	1.4
Stef200	C70.200	224.0	40.2	43.1	7.3

Conclusions

Finite element simulations of the crushing of individual metal tubes and metal tubes embedded into foam materials have been carried out. Initially, the influence of tube diameter (D/t) parameter on the specific energy absorption of empty tubes has been numerically investigated. Similar post-buckling deformation shapes to experimental observations have been successfully predicted by using Abaqus software. Abaqus simulation predictions were observed to be consistent with the experimental results. Given that the metal tubes absorb much greater levels of energy than the foams in which they are embedded, the density of the latter should be set as low as possible, ensuring that the metal reinforcements are held in place during the loading process. The experimental and numerical evidence suggests that it should be possible to predict the energy-absorbing capacity of multi-tube systems using a simple rule of mixtures approach based on the mass fractions of the tubes and the foam. The finite element simulations, once they are validated against experimental results, provide excellent tools to further study crushing characteristics of foams embedded with metal tubes.

References

Alexander, J. M. (1960) An Approximate Analysis of the Collapse of Thin Cylindrical Shells under Axial Loading, *The Quarterly Journal of Mechanics and Applied* **XIII**, 10–15.

Deshpande, V. S. and Fleck, N. A. (2000) Isotropic Constitutive Models for Metallic Foams, *Journal of the Mechanics and Physics of Solids* **48(6-7)**, 1253–1283.

Al Galib, D. and Limam, A. (2004) Experimental and Numerical Investigation of Static and Dynamic Axial Crushing of Circular Aluminum Tubes, *Thin-Walled Structures* **42(8)**, 1103–1137.

Jones, N.(2010) Energy-Absorbing Effectiveness Factor, *International Journal of Impact Engineering* **37(6)**, 754–65.

Karagiozova, D., Nurick, G. N. and Yuen, S. C. K. (2005) Energy Absorption of Aluminium Alloy Circular and Square Tubes under an Axial Explosive Load, *Thin-Walled Structures* **43(6)**, 956–982.

McGregor, I. J., Meadows, D. J., Scott, C. E. and Seeds, A. D. (1993) Structural Crashworthiness and Failure Impact Performance of Aluminum Structure, *Applied Science*. Elsevier Applied Science.

Pled, F., Yan, W. and Wen, C. (2012) Crushing Modes of Aluminium Tubes under Axial Compression, *5th Australasian Congress on Applied Mechanics*.

Porohyperelastic finite element model for the kangaroo humeral head cartilage based on experimental study and the consolidation theory

N. Thibbotuwawa[1], T. Li[1], *Y.T. Gu[1]

[1]School of Chemistry, Physics and Mechanical Engineering
Queensland University of Technology, Brisbane, QLD, 4001, Australia.
*Corresponding author: yuantong.gu@qut.edu.au

Abstract

Solid-extracellular fluid interaction is believed to play an important role in the strain-rate dependent mechanical behaviors of shoulder articular cartilages. It is believed that the kangaroo shoulder joint is anatomically and biomechanically similar to human shoulder joint and it is easy to get in Australia. Therefore, the kangaroo humeral head cartilage was used as the suitable tissue for the study in this paper. Indentation tests from quasi-static (10^{-4}/sec) to moderately high strain-rate (10^{-2}/sec) on kangaroo humeral head cartilage tissues were conduced to investigate the strain-rate dependent behaviors. A finite element (FE) model was then developed, in which cartilage was conceptualized as a porous solid matrix filled with incompressible fluids. In this model, the solid matrix was modeled as an isotropic hyperelastic material and the percolating fluid follows Darcy's law. Using inverse FE procedure, the constitutive parameters related to stiffness, compressibility of the solid matrix and permeability were obtained from the experimental results. The effect of solid-extracellular fluid interaction and drag force (the resistance to fluid movement) on strain-rate dependent behavior was investigated by comparing the influence of constant, strain dependent and strain-rate dependent permeability on FE model prediction. The newly developed porohyperelastic cartilage model with the inclusion of strain-rate dependent permeability was found to be able to predict the strain-rate dependent behaviors of cartilages.

Keywords: Solid-extracellular fluid interaction, Drag force, Strain-rate dependent behavior, Porohyperelasticity, Finite element method

Introduction

In mundane activities such as lifting, throwing etc., shoulder cartilages are subjected to physiologically different strain-rates. It is essential that the shoulder cartilage have the ability to undergo controlled deformation in response to these different external loading conditions. Solid-extracellular fluid interaction is considered to play a significant role in facilitating this behavior of shoulder cartilage tissues by rendering its ability to perform as a mechanically efficient tissue. It is crucial to understand the extent to which solid-extracellular fluid interaction facilitates the strain-rate dependent behavior of shoulder cartilage tissues, in order to identify its implications for initiation of shoulder osteoarthritis and development of artificial shoulder cartilage tissues etc. Therefore, the main objective of the current study is to understand how solid-extracellular fluid interaction facilitates strain-rate dependent behavior of shoulder cartilage tissues.

Evidences from literatures indicated that the mechanical behavior of articular cartilages is strain-rate dependent (Finlay and Repo 1979; Lai et al. 1981; Oloyede and Broom 1992; Oloyede et al. 1992; Radin et al. 1970; Woo et al. 1980). According to experimental findings, with increasing strain-rate, the stiffness quickly increases at the beginning and then obtains an

asymptotic value (Oloyede et al. 1992). The interplay between solid and fluid contributes significantly to this behavior with 70%-80% of the load being supported by the matrix at low strain-rates (Oloyede and Broom 1992), while fluid contributes to a similar percentage at large strain-rates (Li and Herzog 2004; Oloyede and Broom 1992). Researchers have different opinions regarding the mechanisms underlying the strain-rate dependent behavior of cartilage tissues. Back in 1980's, McCutchen (1982) argued that this apparent time-dependent behavior is solely related to the volume loss due to water exudation from the cartilage matrix, rather than the intrinsic viscoelasticity of the matrix. Oloyede et al. (1992) substantiated the argument of McCutchen (1982) by loading the cartilage from quasi-static to impact strain-rates and observed a poroelastic behavior at low strain-rates and elastic behavior at impact strain-rate. Therefore, they claimed that the apparent viscoelastic behavior is due to the drag forces introduced by reduction of permeability with strain and solid-extracellular fluid frictional interactions (Oloyede and Broom 1996). DiSilvestro et al.(2001) proposed that the flow-independent viscoelastic mechanism is the main contributor to the strain-rate dependent behaviors of cartilage. Based on the close confirmation of Biphasic-poroviscoelastic FE model to experimental results, they concluded that the mechanism underlying the strain-rate dependency is dominated by the viscoelasticity of matrix material. Considering the cartilage as a fiber reinforced composite material, Li et al. (2003) claimed that the fluid flow can induce fibril stiffening, which is the main mechanism governing the strain-rate dependent behaviors.

The main argument of Li et al.'s (2003) study is that, the models that do not distinguish between fibrils and proteoglycans are not able to capture the nonlinear transient response of cartilage tissue. They have further argued that the ratio of maximum transient stiffness to equilibrium stiffness has a limitation (<1.5) in models that do not distinguish between fibrils and proteoglycans, hence are not adequate to study the tissue response at high strain-rates. Although biphasic theory (Mow et al. 1980) comprises of those limitations, the porohyperelastic model -based on Biot's (1941) theory which is used in this paper does not have these limitations. Further, hyperelastic material model used in this study to represent solid skeleton is capable of improving the model deviations such as decrease in stiffness observed in fibril reinforced model (Li et al. 2003) at large strains. Compared to the poroviscoelastic model (DiSilvestro et al. 2001), the current model is able to explain the experimental observation without adopting viscoelastic theories.

In this study, we follow the same arguments as of McCutchen (1982) and Oloyede and Broom (1996). Hence we believe that the drag forces and solid-extracellular fluid interaction is the main reason of apparent strain-rate dependent behavior of shoulder cartilage tissues. In order to test this hypothesis, a FE model was developed under the poromechanics framework of Biot (1941) and the modeling results are compared to the experimental results obtained through indentation testing under different strain-rates.

Experimental animal model for shoulder cartilage

Choice of animal model for shoulder cartilage tissue requires the shoulder joint of the animal model to be anatomically and biomechanically similar to that of human joint. Also sufficient tissue thickness is required for the macro scale testing to be carried out. Quadruped animal use forelimbs for weight bearing activities. However, humans do not use forelimbs (shoulder joints) much for weight bearing activities. Therefore, macropods, rat and certain species of mice (kangaroo mice, dipodids, springhare and hopping mice) are the animal models that can be considered to have anatomically and biomechanically similar shoulders to humans. Tissue

thickness of rat and mice are not sufficient for macroscopic testing. In macropod family, tree kangaroo, which is a rare species, is known to have a similar anatomy and biomechanics to that of a human shoulder (Sonnabend and Young 2009). Considering these facts kangaroo was chosen as the suitable animal model for the current study.

Experimental methodology

Tissue harvesting, preparation and thickness measurements

Visually normal, kangaroo cartilage samples of 8mm diameter with 2-3mm of subchondral bone intact were harvested using a specially designed stainless steel puncher. The samples were obtained from central load bearing area of the humeral heads (Fig. 1a). The specimens were harvested from five adult kangaroos (approximately 5 years old) within 24 hours of slaughtered, from an abattoir (Fig. 1b). Until testing, all samples were wrapped around a 0.15M saline wetted towel and stored at -20°C. The sides of cartilage on bone samples were visualized through a microscope (Olympus SZ40, Tokyo, Japan) for magnification purposes and uncalcified cartilage thickness was measured using high precision vernier caliper. Measurements were taken from eight locations of approximately equal gaps around the perimeter and final thickness of cartilage was taken as the mean of measurements.

Indentation testing

Subchondral bone underneath the cartilage sample was properly constrained using stainless steel holder (Fig. 1c), to ensure that the deformation data obtained is only due to the deformation of the cartilage. Prior to testing, all the samples were thawed in 0.15M saline for 1.5 hours. The indentation testing was carried out at 10^{-4} s^{-1}, 5×10^{-4} s^{-1}, 5×10^{-3} s^{-1} and 10^{-2} s^{-1} strain-rates (Fig. 1d). The samples were indented up to 30% engineering strain. A limit of 3MPa was imposed on the amount of stress that samples were subjected to, in order to minimize the damage to the tissues (Morel and Quinn 2004; Quinn et al. 2001). However, none of the cartilage samples attained this limit of stress during the strain-rates tested. The testing was done on Instron testing machine (Model 5944, Instron, Canton, MA) using a plane ended polished indenter of 3mm diameter.. Depending on the thickness of the samples, the speed of the Instron machine was adjusted to obtain the required strain-rate. After each test, prior to the next one, the cartilage was unloaded and allowed to recover for 2 hours in 0.15M saline.
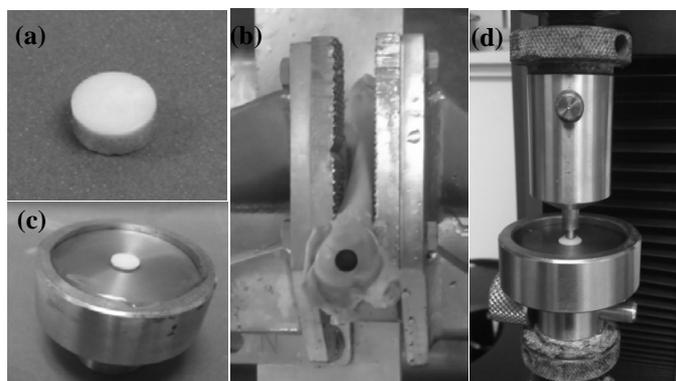
Figure 1 (a) 8mm diameter cartilage sample. (b) Specimen harvested region: Central humeral head. (c) Bone is constrained using a stainless steel holder arrangement. (d) Indentation testing on sample

Pertinent porohyperelastic theory

The governing equations of Biot's (1941) theory for cartilage, considering solid skeleton as a hyperelastic material, can be found in Oloyede and Broom (1994) and Oloyede and Broom (1996) and is summarized below. The fluid flow is taken to follow Darcy's law (equation 1).

$$V_{ij} = -K_{ij}\left(\frac{du}{dx}\right) \qquad (1)$$

Herein, V_{ij}, K_{ij} and u are fluid velocity, direction dependent permeability and fluid excess pore pressure, respectively. Based on the equation (1), the continuity equation of fluid flow:

$$\psi\left[\frac{\partial}{\partial x}\left(K_x\left(\frac{\partial u_x}{\partial x}\right)\right) + \frac{\partial}{\partial y}\left(K_y\left(\frac{\partial u_y}{\partial y}\right)\right) + \frac{\partial}{\partial z}\left(K_z\left(\frac{\partial u_z}{\partial z}\right)\right)\right] = \frac{\partial(\sigma_{ij} - u)}{\partial t} \qquad (2)$$

ψ, K_i and σ_{ij} are matrix bulk moduli, permeability in i[th] direction and applied stress, respectively. The constitutive law for solid skeleton is as follows.

$$\sigma_{ij} = \sigma_{sij} + u\delta_{ij} \qquad\qquad \sigma_{sij} = \frac{\partial W}{\partial \lambda_{ij}} \qquad (3)$$

σ_{sij}, δ_{ij}, λ_{ij} and W are effective solid skeleton stress, Kronecker delta, stretch ratio and isotropic strain energy potential, respectively.

Porohyperelastic FEA model

Cartilage was modeled as a porous media saturated with fluid based on the generalized framework of Biot's (1941) theory. The model was developed in commercial software (ABAQUS 6.12 version). Axisymmetric element (8-node axisymmetric quadrilateral, bilinear displacement and bilinear pore pressure) are adopted to reduce the computational cost based on the characteristics of test sample and loadings. The large deformations and geometric nonlinearity were considered in the calculation. The 'pore pressure (*p*)' (*p*=0) boundary condition was enforced on the upper surface of portion where the indenter is not touching the surface and the right side of the cartilage to enable the fluid flow through these boundaries.. The lower boundary of the model was fixed in vertical and lateral direction, to represent the physiological conditions in cartilage-bone interface. As the stiffness of indenter and bone are higher than the cartilage, the indenter and the bone were modeled as rigid bodies for the ease of modeling. The preliminary studies indicated that the material model used for the indenter and bone does not significantly affect the prediction result.

To account for the non-linear large deformation, the solid skeleton was modeled as an isotropic hyperelastic material. The highly nonlinear stress-strain behavior observed during the current study is unable to be represented by lower order hyperelastic material model such as neo-Hookean or Mooney-Rivlin. Higher order hyperelastic material models such as Yeoh model are believed to be more suitable in explaining the nonlinearity of cartilage tissues (Oloyede et al. 2009). However, 2nd order polynomial hyperelastic function gave an accurate description of the material behavior observed during this study, for humeral head cartilage tissues. Due to lesser number of parameters in 2nd order polynomial hyperelastic function, Inverse FE procedure can be implemented to obtain unique set of material parameters. The general functional form of hyperelastic materials is:

$$W = \sum_{i+j=1}^{N} C_{ij} (\bar{I}_1 - 3)^i (\bar{I}_2 - 3)^j + \sum_{i=1}^{N} \frac{1}{D_i} (J - 1)^{2i} \tag{4}$$

Here, W, $\bar{I}_1$ and $\bar{I}_2$ are isotropic strain energy potential, first and second deviatoric strain invariants, respectively. The J is the volume change during the deformation. C_{ij} and D_i are material parameters related to stiffness and the volumetric change of the cartilage. Setting $N=1$, the above equation reduces to Mooney-Rivilin model. If for all C_{ij} with $j \neq 0$ are set to zero, then $N=1$, $N=2$ and $N=3$ would represent neo-Hookean, the 2nd order polynomial hyperelastic function and Yeoh models, respectively.

The strain dependent permeability function used in this study is shown in equation 5 where intrinsic permeability (K_o) is related to permeability (K) as follows.

$$K = K_o \left(\frac{e}{e_o} \right)^m exp \left\{ \frac{M}{2} \left[\left(\frac{1+e}{1+e_o} \right)^2 - 1 \right] \right\} \tag{5}$$

Here e is the void ratio (ratio of volume of pores to volume of solid), a quantity representing 'dilatation' in ABAQUS. The e_0 is the initial void ratio, which is taken to be 4, based on the assumption that on average 80% of cartilage is filled with pores (Holmes and Mow 1990). The M and m are dimensionless material parameters which were taken to be 4.638 and 0.0848 respectively (Holmes and Mow 1990).

Material parameter identification

Material parameters for the 2nd order polynomial hyperelastic function were obtained using an inverse-FE procedure. Following the approach developed by Simon et al.(1998), stiffness parameters of the 2nd order hyperelastic function, C_{10} and C_{20} were obtained by curve fitting the force-indentation experimental data at highest strain-rate, 10^{-2}/sec, to FE model prediction considering the material as incompressible. On the other hand, considering the material as compressible, parameters related to volumetric change of the 2nd order hyperelastic function D_1 and D_2 were obtained by curve fitting the force-indentation experimental data at lowest strain-rate, 10^{-4}/sec to FE model prediction. Assuming the fluid flow is negligible, the intrinsic permeability of the cartilage was obtained by curve fitting the porohyperelastic FE model prediction to experimental data at lowest strain-rate.

Results and discussion

Biomechanical parameters: Comparison with the literature values

The measured average thickness of kangaroo humeral head cartilage samples was 0.75 ± 0.123mm. The reported thickness value for human shoulder cartilage is 1.44mm (Soslowsky et al. 1992), which is approximately two times higher than that of kangaroo cartilages. Given larger size of humans, these values are reasonable compared to average kangaroos. The 2nd order polynomial hyperelastic function fitted well to the both low and high strain-rate data with R-squared (R^2) vales greater than 0.98. The functional form of the R^2, an error indicator is shown in equation (6).

$$R^2 = 1 - \frac{\sum_i (e_i - f_i)^2}{\sum_i (e_i - \bar{e}_i)^2} \tag{6}$$

Here, e_i is the experimental data and f_i is the FE model prediction. The $\bar{e}_i$ is the mean value of experimental results. The average stiffness parameters, i.e. C_{10} and C_{20}, identified from the inverse finite element analysis were 0.1174 ± 0.0884MPa and 0.1367 ± 0.0767MPa, respectively. The average compressibility parameters, i.e. D_1 and D_2 were 0.0982 ± 0.0588MPa and 0.0636 ± 0.0407MPa. The permeability value identified from the average data of the lowest strain-rate is 7.62×10^{-8}mm/sec. The value of permeability obtained through inverse-FE procedure is at the same order with that reported in the literature for the central region of the humeral head cartilage, which is $1.82 \pm 1.27 \times 10^{-8}$ mm/sec (Huang et al. 2005). The hyperelastic material parameters for shoulder cartilage tissues have not been reported elsewhere. Nevertheless, since μ (shear modulus) $= 2C_{10}$ and assuming Poisson's ratio to be 0.15 (Demarteau et al. 2006; Korhonen et al. 2002), Young's modulus (E) for the average data of this study is obtained to be 0.485MPa. This value is within the range (0.28-0.8MPa) reported for bovine humeral head cartilage (Demarteau et al. 2006; Korhonen et al. 2002). The calculated E for human shoulder cartilages from reported compressive modulus (H_A) (Huang et al. 2005) using equation (7) is 0.142MPa. The Poisson's ratio (υ) was also taken to be 0.15 in this calculation. This value is 3 times smaller than the calculated values in our study. Considering the possible differences in thickness and compositions of cartilages in different species we would consider the value obtained for E in this study is acceptable.

$$H_A = \frac{E(1-\upsilon)}{(1+\upsilon)(1-2\upsilon)} \tag{7}$$

Porohyperelastic FE model: Effect of solid-extracellular fluid interaction and drag force

The comparison of average experimental stress-strain response to the porohyperelastic FE model prediction is shown in Figure 2a and 2b for the cases of constant and strain dependent permeability. In general, both models showed the strain-rate dependent nature, indicating the ability of the poromechanics framework (Biot 1941) to capture the stain-rate dependency. However, compared to the model with constant permeability, the model with strain dependent permeability is able to capture the experimental results at the three lowest strain-rates (10^{-4}/sec, 5×10^{-4}/sec and 5×10^{-3}/sec) well. This is mostly evident at 5×10^{-4}/sec where significant improvement in R^2 value is observed (Fig. 2d). Both constant and strain dependent models were not able to adequately capture the stress-strain variation at 10^{-2}/sec strain-rate. However, the model with strain dependent permeability ($R^2 = 0.8571$) was still able to better capture the variation at 10^{-2}/sec in comparison to the model with constant permeability ($R^2 = 0.7815$) (Fig. 2d). Therefore, the strain-rate dependent tissue response from 10^{-4}/sec to 5×10^{-3}/sec can be attributed to the solid-extracellular fluid interaction and drag forces induce due to shrinkage of pores during tissue deformation, which is represented by strain dependent permeability.

In addition to the strain dependent permeability, we believe that the strain-rate dependent drag forces should be considered when strain-rates reach an order of 10^{-2}. This could be one of the reasons that the strain dependent permeability model cannot adequately predict the tissue response at the highest strain-rate tested, 10^{-2}/sec. One way to take into account the strain-rate dependent drag forces is to consider permeability as a function of strain-rate. According to equation (2) above and has been mentioned by Oloyede and Broom (1996), the fluid exudation from the cartilage will decrease with increasing in strain-rate. This can be attributed to decrease in permeability with the increase of strain-rate. Inverse FE curve fitting to experimental results of current study at 10^{-2}/sec indicated that, a permeability value of 1.62×10^{-8}mm/sec would fit to the experimental results well. This permeability value is approximately 4.7 times smaller than the value (7.62×10^{-8} mm/sec) at the smallest strain-rate. The permeability values obtained using Inverse FE procedure for 5×10^{-3}/sec and 5×10^{-4}/sec

were almost the same, which are closer to 3.62×10^{-8} mm/sec. This indicates that the effect of strain-rate on permeability at these relatively low strain-rates is negligible. Therefore, for strain-rates larger than 10^{-2}/sec, inclusion of drag forces through the strain-rate dependent permeability is reasonable. Similar phenomena has earlier been postulated by Oloyede and Broom (1992). According to their experimental observations, they have stated, "a comparison
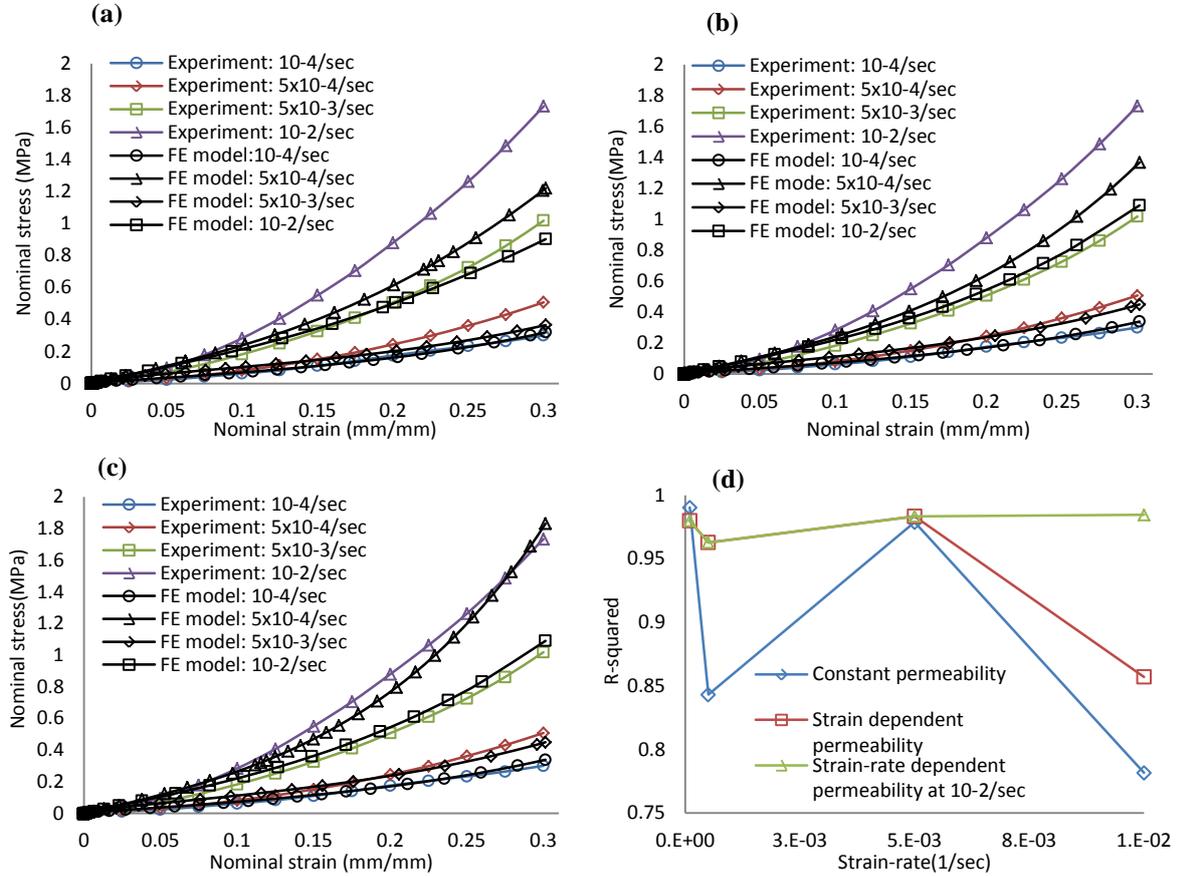
Figure 2. Comparison of FE model prediction to experimental data. (a) Constant permeability (b) Strain dependent permeability (c) With the inclusion of strain-rate dependent permeability at 10^{-2}/sec (d) R-squared values for strain-rates tested

of the effective stress and excess pore pressure curves reveals a distinct dissimilarity in their relationship as the strain-rate is increased from $10^{-3}\mathrm{sec}^{-1}$ to $10^{-2}\mathrm{sec}^{-1}$". The obtained model results after inclusion of strain-rate dependent permeability is shown in Figure 2c. The experimental results are well predicted by the model (R-squared > 0.96) (Fig. 2d).

In finding further evidence for the underlying mechanism of the above mentioned observations, higher pore pressure value is observed for the case of strain-rate dependent permeability with comparison to strain dependent permeability, at 10^{-2}/sec rate (Fig. 3a and 3b). Moreover, the smaller fluid velocities observed in strain-rate dependent case (Fig. 3c) reflects higher drag forces. Therefore, over and above a certain strain-rates (between 5×10^{-3}/sec - 10^{-2}/sec) will significantly reduce the mobility of the fluid leading to literally lock the fluid inside the tissue. We believe that this locking effect will become more prominent at larger stain-rates, causing the tissues to act as purely elastic solid, as evident by the experimental findings of Oloyede et al. (1992).

In summary, the present study has investigated how solid-extracellular fluid interaction facilitates the strain-rate dependent behavior of shoulder cartilage tissues. The porohyperelastic FE model prediction has been compared with indentation tests on kangaroo
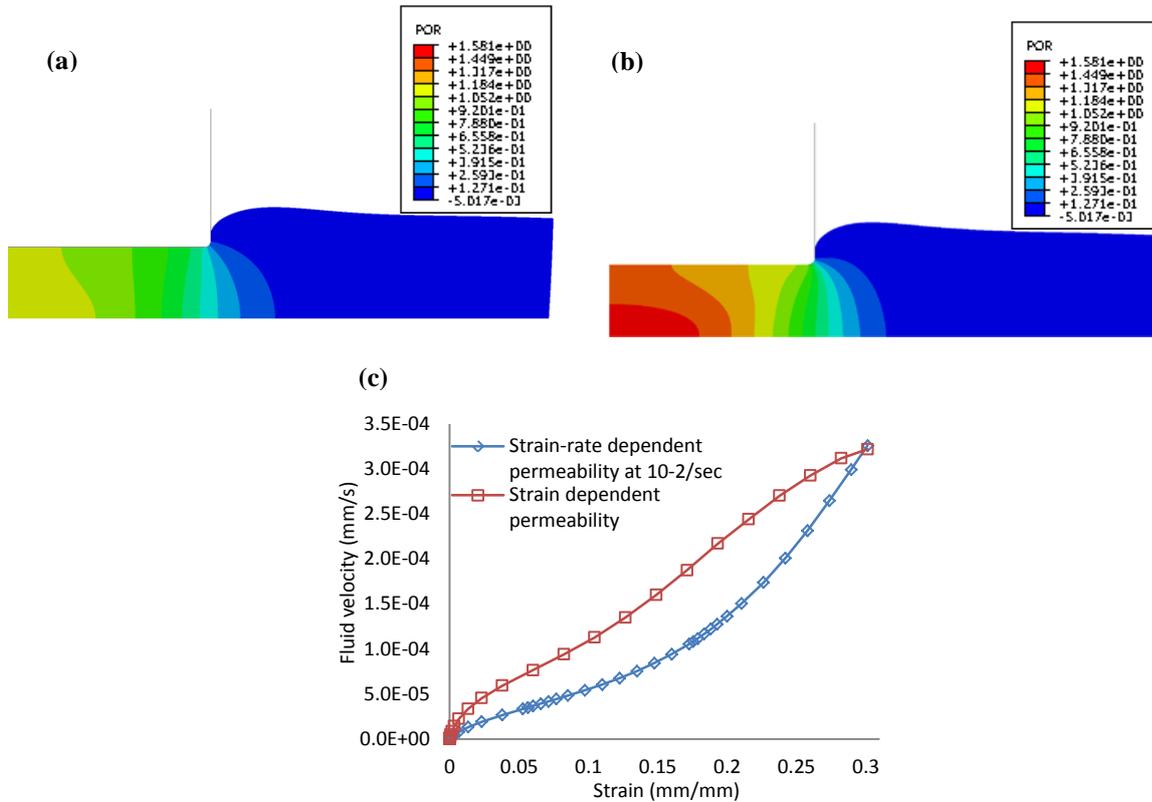
Figure 3. Pore pressure profiles (a) At 10-2/sec: strain-dependent permeabilty. (d) At 10-2/sec: strain-rate dependent (c) fluid velocity at the bottom left of cartilage matrix

humeral head cartilage under quasi-static to high strain-rates. The effect of constant permeability, strain dependent permeability and strain-rate dependent permeability on FE model prediction has been considered with the objective of investigating the effect of solid-extracellular fluid friction forces and drag forces on strain-rate dependent behavior. According to the current investigation when a tissue is deformed under a given strain-rate, shrinking pores will restrict the fluid motion, hence solid-extracellular fluid frictional interaction forces and drag forces will be generated. The magnitude of these forces will depends on the strain, strain-rate, the structure of the pore network and the size of the pores. At higher strain-rates, permeability will reduce significantly due to large drag forces, locking the fluid inside the tissue.

Conclusion

Strain-rate dependent nature of kangaroo humeral head cartilage tissues from 10^{-4}/sec to 10^{-2}/sec is well captured by a newly developed porohyperelastic FE cartilage model with strain-rate dependent permeability. The model with strain dependent permeability was only able to predict the strain-rate dependency from 10^{-4}/sec to 5×10^{-3}/sec and was better than the model with constant permeability. The drag forces are believed to be dominating the tissues response from an intermediate strain-rate from 5×10^{-3}/sec to 10^{-2}/sec. This is the main reason why the strain-rate dependent model is superior at higher strain-rates. Therefore, it is necessary to include the strain-rate dependent permeability in order to predict the tissue response at large strain-rates. Therefore, the strain-rate dependent behavior of shoulder

cartilages can be attributed to solid-extracellular fluid interaction and drag forces. In physiological point of view, reduction of permeability at large strain-rates indicates that the fluid will be locked inside the tissue. This is believed to facilitate the ability of the tissues to function as a protective layer for bone ends injurious loads.

Acknowledgement

References

Biot MA. 1941. General theory of three‐dimensional consolidation. Journal of applied physics 12(2):155-164.

Demarteau O, Pillet L, Inaebnit A, Borens O, and Quinn TM. 2006. Biomechanical characterization and in vitro mechanical injury of elderly human femoral head cartilage: comparison to adult bovine humeral head cartilage. Osteoarthritis Cartilage 14(6):589-596.

DiSilvestro MR, Zhu Q, and Suh J-KF. 2001. Biphasic Poroviscoelastic Simulation of the Unconfined Compression of Articular Cartilage: II—Effect of Variable Strain Rates. Journal of Biomechanical Engineering 123(2):198.

Finlay J, and Repo R. 1979. Energy absorbing ability of articular cartilage during impact. Medical and Biological Engineering and Computing 17(3):397-403.

Holmes M, and Mow V. 1990. The nonlinear characteristics of soft gels and hydrated connective tissues in ultrafiltration. J Biomech 23(11):1145-1156.

Huang CY, Stankiewicz A, Ateshian GA, and Mow VC. 2005. Anisotropy, inhomogeneity, and tension-compression nonlinearity of human glenohumeral cartilage in finite deformation. J Biomech 38(4):799-809.

Korhonen R, Laasanen M, Töyräs J, Rieppo J, Hirvonen J, Helminen H, and Jurvelin J. 2002. Comparison of the equilibrium response of articular cartilage in unconfined compression, confined compression and indentation. J Biomech 35(7):903-909.

Lai W, Mow VC, and Roth V. 1981. Effects of nonlinear strain-dependent permeability and rate of compression on the stress behavior of articular cartilage. Journal of biomechanical engineering 103(2):61-66.

Li LP, Buschmann MD, and Shirazi-Adl A. 2003. Strain-rate Dependent Stiffness of Articular Cartilage in Unconfined Compression. Journal of Biomechanical Engineering 125(2):161.

Li LP, and Herzog W. 2004. Strain-rate dependence of cartilage stiffness in unconfined compression: the role of fibril reinforcement versus tissue volume change in fluid pressurization. J Biomech 37(3):375-382.

McCutchen C. 1982. Cartilage is poroelastic, not viscoelastic (including and exact theorem about strain energy and viscous loss, and an order of magnitude relation for equilibration time). J Biomech 15(4):325-327.

Morel V, and Quinn TM. 2004. Cartilage injury by ramp compression near the gel diffusion rate. Journal of orthopaedic research 22(1):145-151.

Mow V, Kuei S, Lai W, and Armstrong C. 1980. Biphasic creep and stress relaxation of articular cartilage in compression: theory and experiments. Journal of biomechanical engineering 102(1):73-84.

Oloyede A, and Broom N. 1992. Stress-sharing between the fluid and solid components of articular cartilage under varying rates of compression. Connective tissue research 30(2):127-141.

Oloyede A, and Broom N. 1996. The biomechanics of cartilage load-carriage. Connective tissue research 34(2):119-143.

Oloyede A, and Broom ND. 1994. The generalized consolidation of articular cartilage: an investigation of its near-physiological response to static load. Connective tissue research 31(1):75-86.

Oloyede A, Crawford RW, Moody HR, Nguyen TC, and Brown CP. 2009. Assessment of common hyperelastic constitutive equations for describing normal and osteoarthritic articular cartilage. Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine 223(6):643-652.

Oloyede A, Flachsmann R, and Broom ND. 1992. The dramatic influence of loading velocity on the compressive response of articular cartilage. Connective tissue research 27(4):211-224.

Quinn T, Allen R, Schalet B, Perumbuli P, and Hunziker E. 2001. Matrix and cell injury due to sub‐impact loading of adult bovine articular cartilage explants: effects of strain rate and peak stress. Journal of orthopaedic research 19(2):242-249.

Radin EL, Paul IL, and LowY M. 1970. A comparison of the dynamic force transmitting properties of subchondral bone and articular cartilage. The Journal of Bone & Joint Surgery 52(3):444-456.

Simon B, Kaufmann M, McAfee M, Baldwin A, and Wilson L. 1998. Identification and determination of material properties for porohyperelastic analysis of large arteries. Journal of biomechanical engineering 120(2):188-194.

Sonnabend D, and Young A. 2009. Comparative anatomy of the rotator cuff. Journal of Bone & Joint Surgery, British Volume 91(12):1632-1637.

Soslowsky LJ, Flatow EL, Bigliani LU, and Mow VC. 1992. Articular geometry of the glenohumeral joint. Clinical Orthopaedics and Related Research 285:181-190.

Woo S-Y, Simon B, Kuei S, and Akeson W. 1980. Quasi-linear viscoelastic properties of normal articular cartilage. Journal of biomechanical engineering 102(2):85-90.

Topology Optimization of Anisotropic Materials under Harmonic Response Based on ICM Method

*H.L. Ye, Y.M. Li，N. Chen, Y. K. SUI

College of Mechanical Engineering and Applied Electronics Technology, Beijing University of Technology, China

*Corresponding author: yehongl@bjut.edu.cn

Abstract

Topology optimization of anisotropic materials is one of the most challenging research topics in the field of structural optimization and full of innovation. This paper aims at to find the optimal lay-out of anisotropic materials under harmonic response loads within a specified region. The optimization model which subject to response amplitude of the harmonic excitation is established and solved by using Independent, Continuous, Mapping (ICM) method. The filter function of elemental mass matrix, elemental stiffness matrix and elemental weight would be introduced, by which the three matrixes are updated in iteration putted into the dynamic topology optimization of differential equation to analyses the design sensitivity and optimize the structure. An explicit expression of constraint(s) with respect to the topological variables is obtained based on Rayleigh's quotient and sequential approximation method with filter functions. Then, the mathematical formulation of optimal problem of anisotropic materials is established and solved by dual sequence quadratic programming (DSQP). Finally, Numerical examples are provided to demonstrate the validity and effectiveness of the ICM method.

Keywords: Topology optimization；anisotropic materials；harmonic response；ICM method；dynamic topology optimization

Introduction

The vibration Control is a major problem in industrial engineering. How to make materials property achieve optimization and lightweight as far as possible at the same time, is bottleneck of the high-end equipment manufacturing and aerospace industry, which urgently need to be solved. The traditional parameter optimization for structure optimization design cannot meet the requirement of lightweight design in the engineering field. However, through the method of topology optimization, the structure of the topology configuration could be redesign, which is of a new train thought of the structural dynamic design.

A large number of examples have proved that topology optimization is a magic tool to improve the material mechanics performance. Now, many studies have been carried out on dynamic response topology optimization. SIMP method [*Allahdadian S, Boroomand B, Barekatein A R(2012).*] is applied to study the optimal topology of the support structure with minimum compliance design under harmonic force. Kang [*Kang Z, Zhang X, Jiang S, et al(2012)*] aims at damping structure composite board to research minimum amplitude of topology optimization under harmonic excitation by SIMP and GCMMA method. Rong [*Rong J H, Xie Y M，Yang X Y, et al(2000)*] applies the ESO method to study the random dynamic response of minimum weight topology optimization problems. Zhang [*Zhang Qiao, Zhang Weihong, Zhu Jihong. (2010)*] make use of the RAMP methods research the dynamic response topology optimization subject to the random dynamic response of white noise excitation. Recently there are some literatures such as [*Motamarri*

P, Ramani A, Kaushik A(2012)]take advantage of equivalent statics to solve the problem of dynamic response, it avoids to solve the complicated problems such as dynamic equation problems.

In engineering field, engineers often want to get the optimal structure of minimizing weight with satisfying some mechanical constraint at a certain point of interest. As for the minimum weight problem subjected to dynamic response displacement, the objective function and constraints include higher nonlinear implicit function equations. It is more difficult to have it explicit, and difficult to analysis the design sensitivity. However, sensitivity analysis of dynamics is a major problem for topology optimization.

In this paper, ICM method [*Sui Yunkang(1996), Sui Yunkang& Ye Hongling(2013)*]is extended to construct the optimal model of anisotropic materials under harmonic response loads within a specified region. And the optimal model is solved by dual sequence quadratic programming. Numerical examples show that this method is effective and valid for the problem of topology optimization subjected to dynamic response displacement.

The ICM (Independent Continuous Mapping) method for anisotropic materials

ICM method, namely Independent, Continuous and Mapping method, designs a special type of topological variable independent of specific physical quantity to indicates the 'exist-null' of elements, which is proposed by Sui (1996) for skeleton and continuum structures.

The "polish function" and "filter function" are the key points in the ICM method, which are used to map discrete variables and inverse continuous variables. By introduce material's retention ratio v, the discrete variables "0-1" could indicate the optimal lay-out of or topological structure, namely

$$t_i = H(v) = \begin{cases} 1 & v_i / v_i^0 \in (0,1] \\ 0 & v_i / v_i^0 = 0 \end{cases} \tag{1}$$

Where t_i is topological variable, v_i^0 is material's initial value, $H(v)$ is mapping relation, which can be regard as a step function like Fig.1. The "polish function" means using a smoothed curve to approximate the step function like Fig.2. And the "filter function" is inverse function to the "polish function" like Fig.4. The filter function can also be considered as the inverse function of the step function like Fig.3.
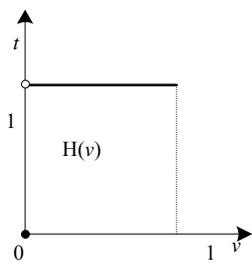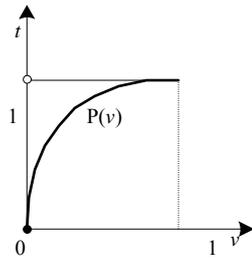
Fig.1 Step function   Fig.2 Polish function   Fig.3 The Hurdle function   Fig.4 Filter function
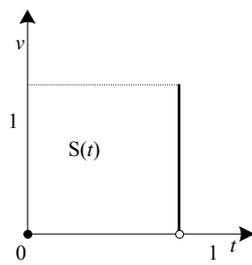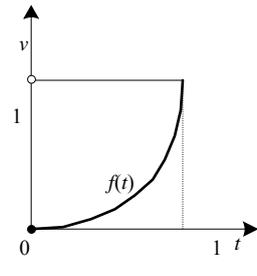
The important role of filter function is use to identify the physical parameters, the element weight, element quality matrix and element stiffness matrix cloud be recognized.

$$w = \bar{w} f_w(t_i), \boldsymbol{M} = \bar{\boldsymbol{M}} f_m(t_i), \boldsymbol{K} = \bar{\boldsymbol{K}} f_k(t_i) \tag{2}$$

Where $\bar{w}, \bar{\boldsymbol{M}}, \bar{\boldsymbol{K}}$ are respectively initial element weight, initial element mass matrix, and initial element stiffness matrix. $w, \boldsymbol{M}, \boldsymbol{K}$ are weight, mass matrix and stiffness matrix of ith element. $f_w(t_i), f_k(t_i), f_m(t_i)$ are respectively filter functions of weight, mass matrix, stiffness matrix of ith element. Thus, the model of continuum topology optimization with dynamic response constraints can be written as follows

$$\begin{cases} find \quad \boldsymbol{t} \in E^N \\ make \ W = \sum_{i=1}^{N} f_w(t_i) w_i^0 \rightarrow \min \\ s.t. \quad u_j(f_k(t_i), f_m(t_i)) \le \overline{u}_j \ (j = 1, \cdots, J) \\ \qquad 0 \le t_i \le 1 \ (i = 1, \cdots, N) \end{cases} \tag{3}$$

Where $\overline{u}_j$ is allowable amplitude constraint, J is the total number of effective displacement constraints. As the equations of continuum problem, the filter function could recognize the elastic tensor and density of structure

$$\begin{cases} C_{ijkl} = f_k C_{ijkl}^0 \\ \rho = f_m \rho^0 \end{cases} \tag{4}$$

Where f_k and f_m are filter function equations of different physical parameters. As we know, there is a relationship between C_{ijkl} and Young's modulus. Assume that there is a matrix A

$$\boldsymbol{C}_{ijkl} = E\boldsymbol{A} = f_k \boldsymbol{C}_{ijkl}^0 = f_k E^0 \boldsymbol{A} \tag{5}$$

For 2d orthotropic material, assuming $\sigma_{33} = \sigma_{32} = \sigma_{31} = 0$, and ignoring the z axis direction of the two modulus, elastic tensor can be expressed as:

$$\boldsymbol{C}_{ijkl} = f_k \boldsymbol{C}_{ijkl}^0 = \begin{bmatrix} f_k \dfrac{E_1^0}{1-v_{12}v_{21}} & f_k \dfrac{E_2^0 v_{21}}{1-v_{12}v_{21}} & 0 \\ f_k \dfrac{E_2^0 v_{21}}{1-v_{12}v_{21}} & f_k \dfrac{E_2^0}{1-v_{12}v_{21}} & 0 \\ 0 & 0 & f_k G_1^0 \end{bmatrix} \tag{6}$$

So, for the anisotropic material, Eq(3) could be expressed

$$E_1 = f_k E_1^0, E_2 = f_k E_2^0, G_{12} = f_k G_1^0 \tag{7}$$

Usually, the composite material's each independent modulus displays a marked difference. The high ratio in low topology variable (null zone) may impact on structure mechanics of exist zone performance, as the local mode in dynamic topology optimization. So ICM identification equations for 2d orthotropic material could be

$$E_1, E_2, G_{12} = \begin{cases} E_1^0 f_k, E_2^0 f_k, G_{12}^0 f_k & 0.1 \le t \le 1.0 \\ \max\left(E_1^0, E_2^0, G_{12}^0\right) f_k & t_{\min} \le t \le 0.1 \end{cases} \tag{8}$$

In ICM method, the polish function is applied to eliminate intermediate variables

$$E = E_0 \bullet f(t) \bullet P_\beta(t) \tag{9}$$

Where $\boldsymbol{E} = \left(E_1, E_2, G_{12}\right), \boldsymbol{E}_0 = \left(E_1^0, E_2^0, G_{12}^0\right)$. In order to prevent the numerical instability, define

$$\mu_{ij} = \frac{d(i,j)}{\sum_{j=1}^{N} d(i,j)} \tag{10}$$

Eq.9 could be

$$E = \sum_{i=1}^{i<j} \mu_{ij} \bullet E_i \bullet f(t) \bullet P_\beta(t) \tag{11}$$

The β of Eq.9 is a variable parameter. The topology variable t_i is approaching to 0-1 with the β increasing. Throughout the iterative process, β is changed in stages.

Dynamic response equations

Dynamic response equation has been extensively researched.
$$\boldsymbol{M\ddot{u}} + \boldsymbol{C\dot{u}} + \boldsymbol{Ku} = \boldsymbol{P} \tag{12}$$
Define the structural damping coefficient is R, the dynamic equation can be turned into
$$\left(-\omega^2[\boldsymbol{M}] + (i\omega R + 1)[\boldsymbol{K}]\right)\{\boldsymbol{u}\} = \{\boldsymbol{P}\} \tag{13}$$
In the process of optimization, the reciprocal transformation to design variables as follows
$$x_i = \frac{1}{f_k(t_i)} \tag{12}$$
Derivation with design variables for both sides of equation, then
$$\left(-\omega^2[\boldsymbol{M}] + (i\omega R + 1)[\boldsymbol{K}]\right)\left\{\frac{\partial\boldsymbol{u}}{\partial x}\right\} = \left\{\frac{\partial\boldsymbol{P}}{\partial x}\right\} - \left(-\omega^2\frac{\partial[\boldsymbol{M}]}{\partial x} + (i\omega R + 1)\frac{\partial[\boldsymbol{K}]}{\partial x}\right)\{\boldsymbol{u}\} \tag{15}$$

$\boldsymbol{P}$ is constant, so $\dfrac{\partial\boldsymbol{P}}{\partial x} = 0$, define

$$\begin{cases} \left(-\omega^2\dfrac{\partial[\boldsymbol{M}]}{\partial x} + (i\omega R + 1)\dfrac{\partial[\boldsymbol{K}]}{\partial x}\right)\{\boldsymbol{u}\} = \{\boldsymbol{Z}\} \\[2mm] \dfrac{\partial\boldsymbol{u}}{\partial x} = \{\boldsymbol{\varphi}\} \end{cases} \tag{16}$$

Through the filter functions, the derivative of stiffness matrix and mass matrix is as followed,
$$\frac{\partial[\boldsymbol{K}]}{\partial x} = \frac{\partial f_k(t_i)}{\partial x}[\boldsymbol{K}_0], \frac{\partial[\boldsymbol{M}]}{\partial x} = \frac{\partial f_M(t_i)}{\partial x}[\boldsymbol{M}_0] \tag{17}$$
So solving the sensitivity is transformed into the problem of solving differential equations
$$\left(-\omega^2[\boldsymbol{M}] + (i\omega R + 1)[\boldsymbol{K}]\right)\{\boldsymbol{\varphi}\} = \{\boldsymbol{Z}\} \tag{18}$$
Solve Eq(12) could get the dynamic response $\boldsymbol{u}$ and Eq(17) could get the sensitivity $\boldsymbol{\varphi}$.

Numerical solution of topology optimization

In order to solve the optimal model of Eq(3), the objective function needs to be modified by second order Tailor expansion. The power function is adopted to recognize structure weight as filter function. The objective function of
$$w = \sum_{i=1}^{N} w^0 f_w(t_i) = \sum_{i=1}^{N}\frac{w^0}{x_i^\alpha}, \alpha = \frac{a}{b} \tag{19}$$
For completely eliminating the intermediate variable, the optimization model is also need to be modified, so polish function is introduced to objective function
$$w = \sum_{i=1}^{N}\frac{w^0}{\left(P_\beta(x_i)\right)^\alpha} \tag{20}$$
Solve the first and second order partial derivative

$$\begin{cases} \dfrac{\partial w}{\partial x} = \sum_{i=1}^{N} \dfrac{-\alpha\left(e^{-\beta} + \beta e^{-\beta x}\right)}{x\left(1 + xe^{-\beta} - \beta e^{-\beta x}\right)^{(\alpha+1)}} w^0 \\[2em] \dfrac{\partial^2 w}{\partial x^2} = \sum_{i=1}^{N} \dfrac{-\alpha(\alpha+1)\left(e^{-\beta} + \beta e^{-\beta x}\right)^2}{\left(1 + xe^{-\beta} - \beta e^{-\beta x}\right)^{(\alpha+2)} + \dfrac{\alpha\beta^2 e^{-\beta x}}{\left(1 + xe^{-\beta} - \beta e^{-\beta x}\right)^{(\alpha+1)}}} w^0 \end{cases} \tag{21}$$

Omit the constant terms, the objective function is approximated as

$$W = \sum_{i=1}^{N}(a_i x_i^2 + b_i x_i) \tag{22}$$

Constraint function $u_j(f_k(t_i), f_m(t_i)) \leq \overline{u_j}$ of Eq(3) can also be explicited by Taylor's approximation. However, as the local approximation, it will take more truncation errors. So we have to find another effective way to make it explicit.
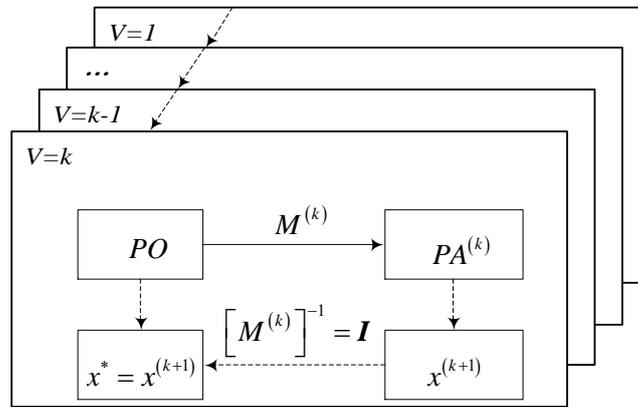
Fig.1 The process card of sequential approximate
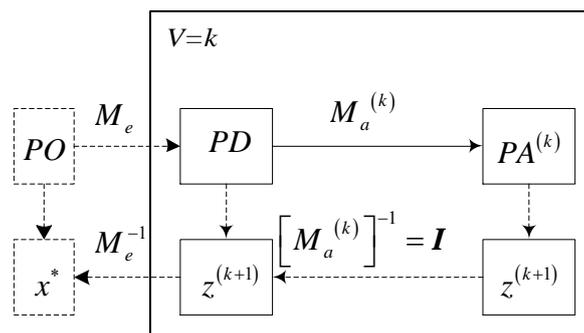
Fig.2 The process card of double mapping

Fig.1 shows the process of sequential approximate. Where v denotes the number of iterations of topology optimization process. PO is the original mathematical programming, PA is mapping mathematical programming. M denotes mapping relationship, M^{-1} denotes inverse relationship $x^{(k+1)}$ is the solution of *PA*. The solution of the original mathematical programming x^* is approximated to $x^{(k+1)}$, namely $x^* = Ex^{(k+1)}$, E is defined as the unit matrix. Sequential

approximation can be understood as the PO to PA when the iterative process start every times. It is turn into a parallel mapping and inversion, as $x^{(0)} \to x^{(1)} \to \cdots \to x^{(k)} \to \cdots \to x^*$, equal to $x^* = \boldsymbol{M}^{-1} x^{(0)}$. According to RMI(Relation Mapping Inverse)(Sui, 1996), there is slight error in the sequence of approximate approach. The inverse relationship $\boldsymbol{M}^{-1}$ is simply putted as $\boldsymbol{E}$, The process of sequential approximate of Fig.1 could be improved as following card.

Fig.2 shows the process of double mapping. PO can be mapped to PD by precise mapping M_e, which is the process sequential approximate. This is a series-parallel connection mapping and inversion method. PO and PA are in series by PD, which is formed a double mapping. The sequential approximate PD to PA$^{(k)}$ formed a multiple parallel, then

$$\left(RMI\right)^2 = \left(RMI_1^{(j)}\right) \times \left(\sum_k RMI_2^{(k)}\right) \tag{23}$$

Where the j is the outer circulation variable, k is the inner circulation variable. The mapping relation M_a is defined linear Taylor expansion.

Now, the response amplitude constraints $u_j(f_k(t_i), f_m(t_i)) \le \overline{u}_j$ could be

$$\boldsymbol{u}_j = \boldsymbol{u}_j^{k-1} + \sum_{i=1}^N \frac{\partial \boldsymbol{u}_j}{\partial t_i}\left(x_i - x_i^{(k-1)}\right) \le \overline{\boldsymbol{u}}_j \tag{24}$$

Define

$$\boldsymbol{\phi}_{ji} = \overline{\boldsymbol{u}}_j - \boldsymbol{u}_j^{(k-1)} + \sum_{i=1}^N \left|\boldsymbol{u}_j'\right| x_i^{(k-1)} \tag{25}$$

Eq(24) can be transformed

$$\sum_{i=1}^N \left|\boldsymbol{u}_j'\right| x_i < \boldsymbol{\phi}_{ji} \tag{26}$$

From above calculation, the optimization mode is

$$\left|\begin{array}{l} find \quad \boldsymbol{x} \in E^N \\[2mm] make \quad W = \sum_{i=1}^N (a_i x_i^2 + b_i x_i) \to \min \\[2mm] s.t. \quad \sum_{i=1}^N \left|\overline{u'}_j x_i\right| \le \phi_{ji}(j = 1, \cdots, J) \\[2mm] \quad \underline{x}_i \le x_i \le \overline{x}_i (i = 1, \cdots, N) \\[2mm] \quad \omega_a \le \omega \le \omega_b \end{array}\right. \tag{13}$$

Considering that the number of design variables is frequently quite bigger than that of constraints in topology optimization of continuum structure, the programming discussed-above could be converted into dual programming according to dual programming theory in order to solve the optimal model simply.

Now, we employ the Dual Quadratic Programs to solve the optimal model (27).

$$\left|\begin{array}{l} find \quad z \\ make \; \Phi(z) \to \max \\ s.t. \quad z \ge 0 \end{array}\right. \tag{28}$$

where $\Phi(\boldsymbol{x}, z) = \min_{\underline{t}_i \le t_i \le 1} (L(\boldsymbol{x}, z))$, $L(\boldsymbol{x}, z) = \sum_{i=1}^N (a_i x_i^2 + b_i x_i) + \sum_{j=1}^J z_j \sum_{i=1}^N (\overline{u'}_j x_i - \phi_{ji})$.

From Kuhn-Tucker condition, we can get the standard quadratic programming. Then solve it and update design variables until the convergence condition of structural weight is satisfied. In this paper, a precision of convergence is prescribed to be 0.001.

Numerical example

Take T300/4211 as base structure of topology optimization, the modules of which is as fellow: E1=126GPa, E2=8GPa, G12=3.7GPa, ρ = 15600kg/m^3. The damping coefficient of the structure is 0.02 and structure size is 80×50×10mm. The exciting force is 1000N located bottom right corner, the frequency of which is 1000Hz, Divide 80 × 50 meshes. The angle of material coordinate system and geometric coordinate system are respectively 0 °, 30 °, 45 °, 60 ° and 90 °.
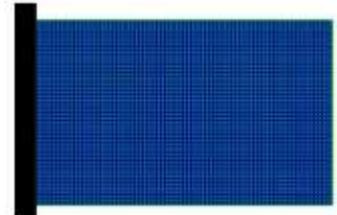
Fig.3 Base structure

| 0° | 30° | 45° | 60° | 90° |
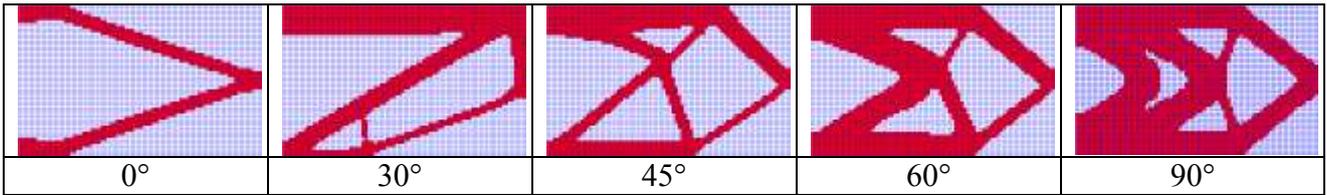
Fig.4 Optimal topology
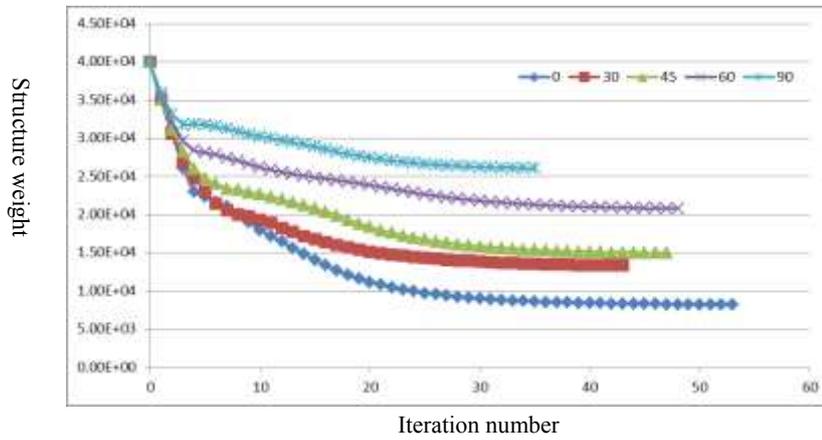
Fig.5 Iteration history of structural weight
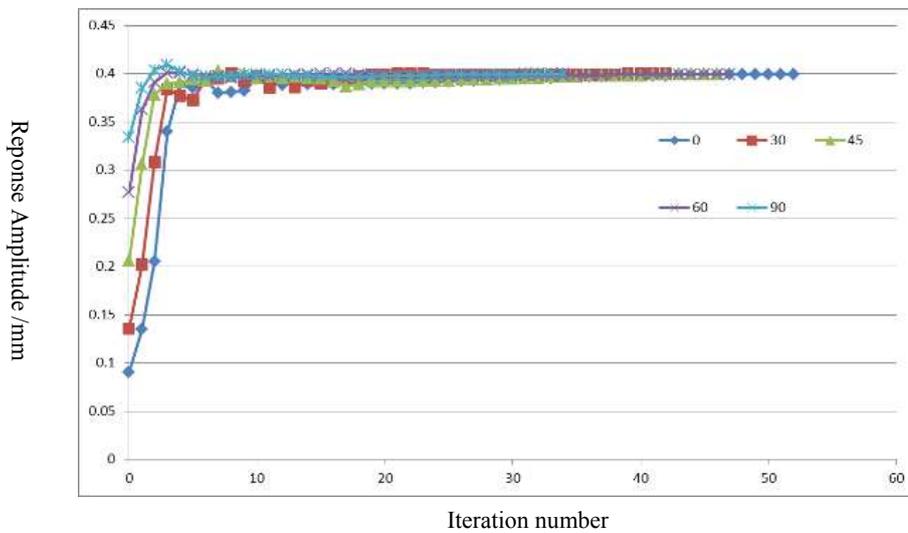
Fig.6 Iteration history of response amplitude

The optimized topology configurations for different ply angles are shown in Fig.4. Iteration history of structural weight and response amplitude with different ply angles are given in Fig.5-Fig.6. As a result, we find that the optimized topology configurations for different ply angles are different. With increasing the ply angle, the structural weight is increasing. But the optimal results are all satisfied with the response amplitude.

Conclusions

Based on ICM topology optimization method, the minimum weight subject to dynamic amplitude response with anisotropic material is established. The logarithmic type filter functions are introduced to build up the anisotropic structure topology optimization model. By using the dual quadratic programming and sequential approximation method, the mathematical model is solved. Numerical example shows that the method of this paper can effectively solve the problem of dynamic response topology optimization of anisotropic material.

Acknowledgements

References

Allahdadian S, Boroomand B, Barekatein A R(2012). Towards optimal design of bracing system of multi-story structures under harmonic base excitation through a topology optimization scheme. *Finite Elements in Analysis and Design*. 61: 60-74.

Kang Z, Zhang X, Jiang S, et al(2012). On topology optimization of damping layer in shell structures under harmonic excitations. *Structural and Multidisciplinary Optimization*. 46(1): 51-67.

Rong J H, Xie Y M，Yang X Y, et al(2000). Topology optimization of structures under dynamic response constraints. *Journal of Sound and Vibration* 234(2): 177-189.

Zhang Qiao, Zhang Weihong, Zhu Jihong. (2010) Topology Optimization of Structures under Dynamic Response Constraints, *Journal of Mechanical Engineering* ,(15):45-51

Motamarri P, Ramani A, Kaushik A(2012). Structural topology synthesis with dynamics and nonlinearities using equivalent linear systems. *Structural and Multidisciplinary Optimization*. 45(4): 545-558.

Sui Yunkang(1996). Modeling, Tansformation and Optimization: New Developments of Structural Synthesis Method. *Dalian: Dalian University of Technology Press*.92-102.

Sui Yunkang, Ye Hongling(2013). Contiuum topology optimization meyhods ICM. *Science Press.*

Legendre-collocation method for nonlinear Volterra

integral equations of the second kinds

*T.G. Zhao[1], Y.J. Li[1], and H. Q. Wu[1]

[1]School of Mathematics, Lanzhou CityUniversity, CHINA.

*Corresponding author: zhaotingg@gmail.com

Abstract

In this paper, we propose an efficient numerical method for Volterra-type nonlinear integral equations, based on Legendre-Gauss-Radau interpolation, which is easy to be implemented and possesses the spectral accuracy. We also develop a multi-step version of this approach. Numerical results demonstrate the effectiveness of these approaches.

Keywords: Legendre collocation method, Volterra integral equation, Nonlinear

Introduction

We are interested in numerically solving Volterra integral equation of the second kind in the following form:

$$u(t) = g(t) + \int_0^t K(t,\tau)F(u(\tau))d\tau, \qquad t > 0 \tag{1}$$

with $u(0) = g(0)$, where the source function $g(t)$ and the kernel function $K(t,\tau)$ are given, and $u(t)$ is the unknown function to be determined. $F()$ is a nonlinear function with certain smoothness. Let $D = \{(t,\tau) \mid 0 \le t \le \tau \le T\}$. If $K(t,\tau) \in C^\infty(D)$ and $F(u) = u$, then the solution to (1) exists, is unique,and belongs to $C^\infty(0,T)$ (Theorems 2.1.2 and 2.1.3 in [Brunner(2004)]). For the existence and uniqueness of nonlinear integral equation as (1), one can consult with [Guo and Sun(1987); Wazwaz(2011)]. We will consider the case where the solutions of (1) are sufficiently smooth -- in this case it is necessary to consider very high-order numerical methods such as spectral methods for approximating the solutions.

As we know, the spectral method employs global orthogonal polynomials as trial functions. It often provides exceedingly accurate numerical results with relatively less degree of freedoms, and thus has been widely used for scientific computation[Bernardi and Maday(1997); Boyd(2000); Guo(1998); Shen and Tang(2006); Shen et al.(2011)]. For spectral methods of (1), the first try may be Chebyshev spectral methods in [Elnagar and Kazemi(1996)]. In [Fujiwara(2006)], Chebyshev spectral methods are investigated for Fredholm integral equations of the first kind under multiple-precision arithmetic. However, no theoretical analysis is provided to justify the high accuracy obtained. In [Tang et al.(2008)], authors developed successfully spectral method and conducted the convergence analysis. Later in [Chen and Tang(2009): Wan et al.(2009); Tao et al.(2011); Xie et al.(2012)], various spectral methods proposed for integral equation with spectral accuracy. Recently, a Legendre-Gauss-Radau collocation method was proposed for initial value problems of ordinary differential equations [Wang and Guo(2012)]. Motivated by the idea in [Wang and Guo(2012)], we propose a Legendre-Gauss-Radau collocation method for a Volterra integral equation (1).

Legendre-Gauss-Radau collocation method and its implementation

Legendre-Gauss-Radau collocation method

Let $L_l(x)$ be the standard Legendre polynomial of degree l. The shifted Legendre polynomials $L_{T,l}(x)$ are defined by $L_{T,l}(t) = L_l(2t/T - 1), l = 0,1,2,\ldots$ The nodes of the standard Legendre-Gauss-Radau interpolation on the interval $[-1,1)$ are $\xi_j (0 \le j \le N)$ and the corresponding Christoffel numbers are $\rho_j (0 \le j \le N)$. The nodes of the shifted Legendre-Gauss-Radau

interpolation on the interval $[0,T)$ are $t_{T,j}^N (0 \le j \le N)$ and the corresponding Christoffel numbers are $\omega_{T,j}^N (0 \le j \le N)$. Clearly, we have links $t_{T,j}^N = \dfrac{T}{2}(\xi_j + 1)$ and $\omega_{T,j}^N = \dfrac{T}{2}\rho_j$ for $0 \le j \le N$. Let $P_N(0,T)$ be the set of polynomials of degree at most N. For any $v \in C[0,T)$, the shifted Legendre-Gauss-Radau interpolation $\Pi_N v \in P_N(0,T)$ is determined uniquely by $\Pi_N v(t_{T,j}^N) = v(t_{T,j}^N)$ ($0 \le j \le N$). For any $K(t,s) \in C[0,T)^2$, notation $\widetilde{\Pi}_N^k$ means the shifted Legendre-Gauss-Radau interpolation on $[0,t_{T,k}^N)$ with respect to s. The corresponding nodes and Christoffel numbers are $t_{t_k,j}^N$ and $\omega_{t_k,j}^N$, respectively.

The Legendre-Gauss-Radau collocation method for (1) is to seek $u^N(t) \in P_N(0,T)$ such that

$$u(t_{T,k}^N) = g(t_{T,k}^N) + \sum_{j=0}^N K(t_{T,k}^N, t_{t_k,j}^N)[\Pi_N F(u^N)](t_{t_k,j}^N)\omega_{t_k,j}^N, \qquad 1 \le k \le N \tag{2}$$

with $u(t_{T,0}^N) = g(t_{T,0}^N)$. Because of exactness of the Legendre-Gauss-Radau quadrature for polynomials of degree at most $2N$ and $\widetilde{\Pi}_N^k K(t_{T,j}^N, \tau)\Pi_N F(u^N(\tau)) \in P_{2N}(0, t_{T,k}^N)$, scheme (2) is equivalent to

$$u(t_{T,k}^N) = g(t_{T,k}^N) + \int_0^{t_{T,k}^N} \widetilde{\Pi}_N^k K(t_{T,k}^N, \tau)\Pi_N F(u^N(\tau))d\tau, \qquad 1 \le k \le N. \tag{3}$$

Implementation of the scheme (2)

We first express the approximate solution as

$$u^N(t) = \sum_{l=0}^N \widehat{u}_l L_{T,l}(t) \tag{4}$$

with $\widehat{u}_l = \dfrac{2l+1}{T}\sum_{j=0}^N u^N(t_{T,j}^N)L_{T,l}(t_{T,j}^N)\omega_{T,j}^N$ ($0 \le l \le N$). Then the nonlinear term under the integral in (3) is expressed as

$$\Pi_N F(u^N(t)) = \frac{1}{T}\sum_{j=1}^N \sum_{l=0}^N (2l+1)L_{T,l}(t_{T,j}^N)F(u^N(t_{T,j}^N))\omega_{T,j}^N L_{T,l}(t)$$

$$+ \frac{1}{T}\sum_{l=0}^N (2l+1)L_{T,l}(t_{T,0}^N)F(u^N(t_{T,0}^N))\omega_{T,0}^N L_{T,l}(t). \tag{5}$$

Inserting (5) into (2), we have

$$u(t_{T,k}^N) = g(t_{T,k}^N) + \sum_{l=0}^N \left(\sum_{i=0}^N \widetilde{\Pi}_N^k K(t_{T,k}^N, t_{t_k,i}^N)L_{T,l}(t_{t_k,i}^N)\omega_{t_k,i}^N\right)\left(\frac{2l+1}{T}L_{T,l}(t_{T,0}^N)\omega_{T,0}^N\right)F(g(t_{T,0}^N))$$

$$+ \sum_{j=1}^N \left[\sum_{l=0}^N \left(\sum_{i=0}^N \widetilde{\Pi}_N^k K(t_{T,k}^N, t_{t_k,i}^N)L_{T,l}(t_{t_k,i}^N)\omega_{t_k,i}^N\right)\left(\frac{2l+1}{T}L_{T,l}(t_{T,j}^N)\omega_{T,j}^N\right)\right]F(u^N(t_{T,j}^N)). \tag{6}$$

Further, we set

$$\mathsf{u} = [u^N(t_{T,1}^N), u^N(t_{T,2}^N),..., u^N(t_{T,N}^N)]^T, \quad \mathsf{g} = [g(t_{T,1}^N), g(t_{T,2}^N),..., g(t_{T,N}^N)]^T,$$

$$\mathsf{F}(\mathsf{u}) = [F(u^N(t_{T,1}^N)), F(u^N(t_{T,2}^N)),..., F(u^N(t_{T,N}^N))]^T,$$

$$\mathbf{B} = (b_{lj})_{(N+1)\times N} = \left(\frac{2l+1}{T}L_{T,l}(t_{T,j}^{N})\omega_{T,j}^{N}\right)_{1\le j\le N;0\le l\le N},$$

$$\mathbf{C} = (c_{kl})_{N\times(N+1)} = \left(\sum_{j=0}^{N}K(t_{T,k}^{N},t_{t_{k},j}^{N})L_{T,l}(t_{t_{k},j}^{N})\omega_{t_{k},j}^{N}\right)_{1\le k\le N;0\le l\le N},$$

$$\mathbf{h} = (h_{l})_{(N+1)\times 1} = \left(\frac{2l+1}{T}L_{T,l}(t_{T,0}^{N})\omega_{T,0}^{N}\right)_{0\le l\le N}.$$

Then we can rewrite (6) as the following compact matrix form

$$\mathbf{u} = \mathbf{g} + \mathbf{CBF}(\mathbf{u}) + \mathbf{Ch}F(g(t_{T,0}^{N})). \tag{7}$$

In actual computation, we first use (7) to evaluate $u^{N}(t_{T,k}^{N})$, $1\le k\le N$. Then we use (4) to obtain

$$u^{N}(T) = \sum_{l=0}^{N}\left(\frac{2l+1}{T}\sum_{j=0}^{N}u^{N}(t_{T,j}^{N})L_{T,l}(t_{T,j}^{N})\omega_{T,j}^{N}\right)L_{T,l}(T). \tag{8}$$

The expression above can be in matrix –vector form as

$$u^{N}(T) = \mathbf{bDv} \tag{9}$$

where

$$\mathbf{b} = (b_{l})_{1\times(N+1)} = \left(\frac{2l+1}{T}L_{T,l}(T)\right)_{0\le l\le N}, \quad \mathbf{D} = (d_{jl})_{(N+1)\times(N+1)} = \left(L_{T,l}(t_{T,j}^{N})\right)_{0\le j\le N;0\le l\le N},$$

$$\mathbf{v} = (v_{j})_{(N+1)\times 1} = \left(u^{N}(t_{T,j}^{N})\omega_{T,j}^{N}\right)_{0\le j\le N}.$$

Multi-step version of the collocation method

Let M be a positive integer number, and $N_{m}(1\le m\le M)$ be positive integer numbers. We divide the interval $[0,T]$ as $0 = t_{0} < t_{1} < ... < t_{m-1} < t_{m} < ... < t_{M} = T$. Set $\tau_{m} = t_{m} - t_{m-1}$. Replacing T and N by τ_{1} and N_{1} in (2) we obtain the local numerical solution $u_{1}^{N_{1}}(t) \in P_{N_{1}}(t_{0},t_{1})$ with the initial value $u_{1}^{N_{1}}(0) = g(0)$. Next we evaluate the local numerical solution $u_{m}^{N_{m}}(t) \in P_{N_{m}}(t_{m-1},t_{m})$ for $m = 2,3,...,M$ by

$$u_{m}^{N_{m}}(t_{\tau_{m},k}^{N_{m}}) = g(t_{\tau_{m},k}^{N_{m}}) + \int_{t_{n-1}}^{t_{\tau_{m},k}^{N_{m}}}\widetilde{\Pi}_{N}^{k}K(t_{\tau_{m},k}^{N_{m}},\tau)\Pi_{N}F(u_{m}^{N_{m}}(\tau))d\tau$$

$$+ \sum_{s=1}^{m-1}\int_{t_{s-1}}^{t_{s}}\widetilde{\Pi}_{N}^{k}K(t_{\tau_{m},k}^{N_{m}},\tau)\Pi_{N}F(u_{s}^{N_{s}}(\tau))d\tau, \quad 1\le k\le N_{m}, \tag{10}$$

with $u_{m}^{N_{m}}(t_{\tau_{m},0}^{N_{m}}) = u_{m-1}^{N_{m-1}}(t_{m-1})$. Notation $t_{\tau_{m},k}^{N_{m}}$ in (10) denote the shifted Legendre-Gauss-Radau nodes in the interval $[t_{m-1},t_{m})$ and $t_{\tau_{m},k}^{N_{m}} = \frac{\tau_{m}}{2}(\xi_{k}+1) + t_{m-1}$ (the corresponding Christoffel numbers are $\omega_{\tau_{m},k}^{N_{m}} = \frac{\tau_{m}}{2}\rho_{k}(0\le k\le N_{m})$). Now we set

$$\mathbf{u}^{\tau_{k}} = [u_{k}^{N_{k}}(t_{\tau_{k},1}^{N_{k}}),u_{k}^{N_{k}}(t_{\tau_{k},2}^{N_{k}}),...,u_{k}^{N_{k}}(t_{\tau_{k},N_{k}}^{N_{k}})]^{T}, \quad \mathbf{g}^{\tau_{k}} = [g(t_{\tau_{k},1}^{N_{k}}),g(t_{\tau_{k},2}^{N_{k}}),...,g(t_{\tau_{k},N_{k}}^{N_{k}})]^{T},$$

$$\mathbf{F}(\mathbf{u}^{\tau_{k}}) = [F(u_{k}^{N_{k}}(t_{\tau_{k},1}^{N_{k}})),F(u_{k}^{N_{k}}(t_{\tau_{k},2}^{N_{k}})),...,F(u_{k}^{N_{k}}(t_{\tau_{k},N_{k}}^{N_{k}}))]^{T},$$

$$\mathbf{B}^{\tau_k} = (b_{lj})_{(N_k+1) \times N_k} = \left(\frac{2l+1}{\tau_k} L_{\tau_k,l}(t^{N_k}_{\tau_k,j}) \omega^{N_k}_{\tau_k,j} \right)_{1 \le j \le N_k; 0 \le l \le N_k},$$

$$\mathbf{C}^{(\tau_k, \tau_n)} = (c_{kl})_{N_k \times (N_n+1)} = \left(\int_{t_{k-1}}^{t^{N_k}_{\tau_k,m}} K(t^{N_k}_{\tau_k,m}, \tau) L_{\tau_n,l}(\tau) d\tau \right)_{1 \le m \le N_k; 0 \le l \le N_n},$$

$$\mathbf{h}^{\tau_k} = (h_l)_{(N_k+1) \times 1} = \left(\frac{2l+1}{\tau_k} L_{\tau_k,l}(t^{N_k}_{\tau_k,0}) \omega^{N_k}_{\tau_k,0} \right)_{0 \le l \le N_k},$$

$$\mathbf{b}^{\tau_k} = (b_l)_{1 \times (N_k+1)} = \left(\frac{2l+1}{\tau_k} L_{\tau_k,l}(t_k) \right)_{0 \le l \le N_k},$$

$$\mathbf{D}^{\tau_k} = (d_{jl})_{(N_k+1) \times (N_k+1)} = \left(L_{\tau_k,l}(t^{N_k}_{\tau_k,j}) \right)_{0 \le j \le N_k; 0 \le l \le N_k},$$

$$\mathbf{v}^{\tau_k} = (v_j)_{(N_k+1) \times 1} = \left(u^{N_k}_k(t^{N_k}_{\tau_k,j}) \omega^{N_k}_{\tau_k,j} \right)_{0 \le j \le N_k}.$$

Then the matrix form of the multi-step version is, first we solve the system

$$\begin{cases} \mathbf{u}^{\tau_1} = \mathbf{g}^{\tau_1} + \mathbf{C}^{(\tau_1,\tau_1)}\mathbf{B}^{\tau_1}\mathbf{F}(\mathbf{u}^{\tau_1}) + \mathbf{C}^{(\tau_1,\tau_1)}\mathbf{h}^{\tau_1}F(g(0)) \\ u^{N_1}_1(t_1) = \mathbf{b}^{\tau_1}\mathbf{D}^{\tau_1}\mathbf{v}^{\tau_1} \end{cases}$$

Next, step by step, the local numerical solution $u^{N_m}_m(t) \in P_{N_m}(t_{m-1}, t_m)$ $(m > 1)$ can be obtained by

$$\begin{cases} \mathbf{u}^{\tau_m} = \mathbf{g}^{\tau_m} + \mathbf{C}^{(\tau_m,\tau_m)}\mathbf{B}^{\tau_m}\mathbf{F}(\mathbf{u}^{\tau_m}) + \mathbf{C}^{(\tau_m,\tau_m)}\mathbf{h}^{\tau_m}F(g(t^{N_m}_{\tau_m,0})) \\ \qquad + \sum_{k=1}^{m-1} \left(\mathbf{C}^{(\tau_m,\tau_k)}\mathbf{B}^{\tau_k}\mathbf{F}(\mathbf{u}^{\tau_k}) + \mathbf{C}^{(\tau_m,\tau_k)}\mathbf{h}^{\tau_k}F(g(t^{N_k}_{\tau_k,0})) \right) \\ \qquad\qquad u^{N_m}_m(t_m) = \mathbf{b}^{\tau_m}\mathbf{D}^{\tau_m}\mathbf{v}^{\tau_m} \end{cases}$$

Noting that $t^{N_k}_{\tau_k,0} = t_{k-1}$, the process above can be done without any gap.

Numerical experiments

The first issue in performing the proposed method is how to solve the nonlinear system (7). We can use a simple iterate scheme as follow:

$$\mathbf{u}^{k+1} = \mathbf{g} + \mathbf{CBF}(\mathbf{u}^k) + \mathbf{Ch}F(g(0)), \quad k = 0,1,2,.... \tag{11}$$

If the sequence $\mathbf{u}^k$ converges, we can obtain a good approximation solution to (7).

In the following examples, because the exact solutions are known, we can compute the error between numerical solution and the corresponding exact ones. The errors in L^∞ and L^2 are defined by

$$\text{Errinf} = \max_{0 \le k \le N} |u^N(t^N_{T,k}) - u(t^N_{T,k})|, \quad \text{Err2} = \sqrt{\sum_{k=0}^{N} |u^N(t^N_{T,k}) - u(t^N_{T,k})|^2 \, \omega^N_{T,k}}.$$

where $u^N(t)$ and $u(t)$ are numerical and exact solution, respectively. Another absolute error at $t = T$ is defined by $\text{Err}(T) = |u(T) - u^N(T)|$.

Example 1 The first example is concerned with a linear problem. Consider Volterra integral equation (1) with $g(t) = e^{2t} - \dfrac{e^{t(t+2)} - 1}{t+2}$, $K(t,\tau) = e^{t\tau}$ and $F(u) = u$. The exact solution is $u(t) = e^{2t}$.

With the same exact solution and kernel as above, authors test a Legendre collocation(LCM) in [Tang et al.(2008)] and also spectral Jacobi-Galerkin method (spectral Legendre-Galerkin method (SLGM) and spectral Chebyshev-Galerkin method(SCGM)) and pseudo-spectral Jacobi-Galerkin method (pseudo-spectral Legendre-Galerkin method (PSLGM) and pseudo-spectral Chebyshev-

Galerkin method(PSCGM)) in [Xie et al.(2012)] for a slightly different problem. We compare the results in table 1 for L^∞-errors. Our results are slightly better than the ones in [Tang et al.(2008)] and [Xie et al.(2012)].

Table 1.   Comparison of L^∞-errors for example 1.

N	4	6	8	10	12	14
LCM	-	3.66e-01	1.88e-02	6.57e-04	1.65e-05	3.11e-07
SLGM	5.243e-02	1.262e-03	1.753e-05	1.572e-07	9.779e-10	4.618e-12
SCGM	2.915e-02	5.696e-04	7.276e-06	5.751e-08	3.950e-10	1.737e-12
PSLGM	6.007e-03	9.386e-05	8.710e-07	6.378e-09	3.322e-11	1.323e-13
PSCGM	7.113e-03	1.003e-04	9.958e-07	6.995e-09	3.638e-11	1.492e-13
Our method	3.137e-04	1.411e-06	3.833e-09	7.665e-12	1.245e-12	1.373e-12

In table 2, we present the L^2 error and the absolute errors at $T=1$ with different N. The results show "spectral accuracy".

Table 2.   Errors with different N for example 1.

N	4	6	8	10	12
Err2	2.192e-04	9.098e-07	2.433e-09	4.534e-12	4.406e-13
Err(1)	5.508e-03	3.897e-05	1.539e-07	3.809e-10	8.383e-12

Example 2 The second example is concerned with a nonlinear problem. Consider Volterra integral equation (1) with $g(t) = \sin t + \dfrac{1-\cos t}{2} + \dfrac{\sin 2t - 2t}{8} + \dfrac{\cos t - e^{-2t} - 2\sin t}{10} + \dfrac{2 - e^{-2t} - \sin 2t - \cos 2t}{16}$,

$K(t,\tau) = -\dfrac{1 - e^{-2(t-\tau)}}{2}$ and $F(u) = u - u^2$. The exact solution is $u(t) = \sin t$. For solving the nonlinear system (7), the simple iterate method (11) is employed with tolerate $\varepsilon = 10^{-16}$.

In table 3, we present the errors with different N. The results also show "spectral accuracy".

Table 3.   Errors with different N for example 2.

N	4	6	8	10	12
Errinf	1.042e-06	2.254e-09	3.566e-12	4.385e-15	2.220e-16
Err2	7.466e-07	1.572e-09	2.521e-12	3.079e-15	1.118e-16
Err(1)	5.584e-05	9.846e-08	9.688e-11	3.948e-13	1.929e-13

Example 3 Consider Volterra equation (1) with

$$g(t) = \frac{6\pi \sin(6\pi t) - \cos(6\pi t) + 36\pi^2 - 36\pi^2 e^t + 1}{2(1 + 36\pi^2)}, \; K(t,\tau) = e^{t-3\tau}, \; F(u) = u^2.$$

The exact solution is $u(t) = e^t \sin(3\pi t)$. In table 4, we present the errors with different N. The results also show "spectral accuracy".

Table 4. Errors with different N for example 3.

N	8	12	16	20	24	28
Errinf	2.498e-02	8.997e-04	7.892e-06	2.202e-08	2.180e-11	4.219e-14
Err2	1.667e-02	6.681e-04	6.065e-06	1.712e-08	1.639e-11	1.581e-14
Err(1)	7.825e-02	3.265e-03	3.677e-05	1.240e-07	1.404e-10	2.412e-13

Conclusions

In this paper, we proposed a Legendre-Gauss-Radau collocation method for solving Volterra-type integral equations of second kind. This method is easy to be implemented for nonlinear problems. In particular, benefiting from the rapid convergence of the Legendre-Gauss-Radau interpolation, this method possesses spectral accuracy.

We also provided a multi-step version of Legendre-Gauss-Radau collocation method. We could use this process with moderate mode N to evaluate the numerical solution, step by step. This simplifies actual computation and saves work essentially. In the derivation of the existing collocation method, one could use the Lagrange interpolation which is unstable for large N. Whereas we used the Gauss-type interpolation as in [Guo and Wang(2007;2009;2010)], which makes our methods much more stable for large N. This is also confirmed by the numerical results.

Acknowledgement

References

Brunner, H. (2004) *Collocation Methods for Volterra Integral and Related Functional Equations*, Cambridge Univerisity Press, New York.

Guo, D. J. and Sun, J. X. (1987) *Nonlinear Integral Equations(in Chinese)*, Shandong Science and Technology Press, Shandong ,China.

Wazwaz, A. M. (2011) *Linear and Nonlinear Integral Equations: Methods and Applications*, Springer.

Bernardi, C. and Maday, Y. (1997) *Spectral Method*, In Handbook of Numerical Analysis, Vol. V, Techniques of Scientific Computing(Part 2), Ciarlet, P. G. and Lions, J. L. eds., pp.209-486. North-Holland, Amsterdam.

Boyd, J. P.(2000) *Chebyshev and Fourier Soectral Methods,* 2nd Edition, Dover Publication Inc., Mineola, New York.

Guo, B. Y.(1998), *The Spectral Methods and Its Applications*, World Scientific, Singapore..

Shen, J. and Tang, T. (2006) *Spectral and High-order Methods with Applications*, Science Press, Beijing.

Shen, J., Tang, T.and Wang, L. L.(2011) *Spectral Methods: Algorithms,Analysis and Applications*, Springer Series in Computational Mathematics 41, Springer-Verlag, Berlin.

Elnagar, G. N. and Kazemi, M.(1996) Chebyshev spectral solution of nonlinear Volterra-Hannerstein integral equations, *Journal of Computational and Applied Mathematics*, **76**,147-158.

Fujiwara, H.(2006)High-accurate numerical method for integral equations of the first kind under multiple-precision arithmetic, *Preprint*, RIMS, Kyoto University, 2006.

Tang,T., Xu, X. and Cheng, J.(2008) On spectral methods for Volterra integral equations and the convergence analysis, *Journal of Computational Mathematics*, **26(6)**(2008), 825-837.

Chen,Y. P. and Tang, T.(2009) Spectral methods for weakly singular Volterra integral equations with smooth solutions, *Journal of Computational and Applied Mathematics*, **233**, 938-950

Wan, Z. S., Chen, Y. P. and Huang,Y. Q. (2009) Legendre spectral Galerkin method for second-kind Volterra integral equations, *Frontiers of Mathematics in China*, **4(1)**, 181-193.

Tao, X., Xie, Z. Q. and Zhou, X. J.(2011)Spectral Petrov-Galerkin methods for the second Kind Volterra type integro-differential equations, *Numerical Mathematics: Theory, Methods and Applications*, **4(2)**, 216-236.

Xie, Z. Q., Li, X. J. and Tang T.(2012) Convergence analysis of spectral Galerkin methods for Volterra type integral equations, *Journal of Scientific Computing*, **53**, 414-434.

Wang, Z. Q. and Guo B. Y.(2012) Legendre-Gauss-Radau collocation method for solving initial value problems of first order odinary differential equations, *Journal of Scientific Computing,* **52**, 226-255.

Guo, B. Y. and Wang, Z. Q.(2007) Numerical integration based on Laguerre-Gauss interpolation, *Computer Methods in Applied Mechanics and Engineering*, **196**, 3726-3741.

Guo, B. Y. and Wang, Z. Q.(2009) Legendre-Gauss collocation methods for ordinary differential equations, *Advances in Computational Mathematics*, **30**, 249-280.

Guo, B. Y. and Wang, Z. Q.(2010) A spectral collocation method for solving initial value problems of first order ordinary differential equations, *Discrete and Continuous Dynamical Systems – SeriesB* ,**14**, 1029-1054.

Free vibration analysis of the functionally graded coated and undercoated plates

*Y. Yang[1], K.P. Kou[1], C.C. Lam[1], V.P. Iu[1]

[1]Department of Civil Engineering, Faculty of Science and Technology, University of Macau, Macau, China
*Corresponding author. E-mail addresses: yangyang.liju@hotmail.com

Abstract

In this paper, the free vibration behaviors of the functionally graded (FG) coated and undercoated plates are studied by a meshfree boundary-domain integral equation method. Based on the two-dimensional elasticity theory, the boundary-domain integral equations for each single layer of these coating-substrate plates are derived initially by using elastostatic fundamental solutions. Employ the radial integration method to transform the domain integrals into boundary integrals and achieve a meshfree scheme. By applying the multi-region boundary element method, obtain the generalized eigenvalue system of the whole plate, which involves system matrices with boundary integrals only and the complete solutions for natural frequency and vibration modes are rigidly resolved. A comparative study of FG versus homogeneous coating is conducted. The influences of material composition, material gradient, coating thickness ratio, plate aspect ratio and the boundary conditions on the natural frequencies of the FG coated and undercoated plates are evaluated and discussed.

Key words: free vibration, FG coated and undercoated plates, boundary-domain integral equations, meshfree method, multi-region boundary element method

Introduction

In many applications, especially in the space industry, energy industry and electronic industry, structures or part of structures are exposed to high temperature or high temperature gradients. Conventional metallic materials, such as carbon steels or stainless steels cannot resist such high temperature. In order to improve the resistance of metallic structures against extreme temperature conditions, without suppressing their strength and toughness, a thin layer of appropriate ceramic is generally used to cover the surface of the structures. For those structures which are subjected to constantly rolling, sliding contacts or abrasive wear, additional hardening process should be carried out within the outer surface of the materials. These two techniques are all forming the coating-substrate system, where a functional material is coated on the substrate material to increase the durability and reliability of the structures. However, due to the discontinuous of the material properties of these two or more materials, severe residual and working stresses discontinuity at the material interfaces usually cause damage to the coating, or failure due to delamination. As a remedy to the aforementioned disadvantages in coating-substrate system, a concept of functionally graded (FG) coating are proposed, where a smooth spatial gradation of the material properties are introduced from coating to substrate in order to eliminated the effect of the suddenly change of the material properties, such that stress and strain discontinuous can be mitigated in the coating-substrate system.

Due to the superiorly properties, a world-wide requirements of the application of the FG coating-substrate system triggered a series of research activities. The incorporation of functionally graded materials (FGMs) into coating design can help eliminate the mismatch of mechanical and thermo-mechanical properties between the metal plates and coating layers. Thus a number of studies existed in the literature for analyzing of the mechanical and thermo-mechanical behavior of homogeneous plates coated by an FG layer. An FG coated elastic solid under thermomechanical loading was carried out by Shodja and Ghahremaninejad [2006]. Three-dimensional elastic deformation of a

functionally graded coating/substrate system was investigated by Kashtalyan et al [2007]. Chung and Chen [2007] analyzed the bending behavior of the thin plates coated by FG layer. Several researches have addressed contact response of FG coatings. Saizonou et al. [2002] studied the subsurface stress distribution of an FG-coated elastic solid under normal and sliding contact loading by the boundary element method (BEM). Contact mechanics of the FG coated solids was analyzed by Guler and Erdogan [2004]. It should be noted that in all above studies the properties the FGM were all assumed to vary exponentially through the thickness.

Theoretical modeling of FG coatings has been focused predominantly on prediction of their fracture behaviours. Chen and Erdogan [2003] studied the interface cracks for a FG coating medium. A crack in the FG coating surface and its expansion into the substrate along the direction perpendicular to the interface between the coating and the substrate was presented by Chi and Chuang [1996]. Pindera et al. [2002, 2005] examined fracture mechanisms in thermal barrier coatings with FG bond coats under uniform cyclic thermal loading. However, the dynamic analyses of the FG coatings are very rare in the literature. Liew et al. [2006] investigated linear and non-linear vibrations of a coating-FGM-substrate cylindrical panel subjected to a temperature gradient, which were based on the first order shear deformation theory and von-Karman geometric nonlinearity. Hosseini-Hashemi et al. [2012] presented the exact closed-form solutions for both in-plane and out-of-plane free vibration of the simply supported rectangular plates coated by a FG layer, based on three-dimensional elasticity theory.

In this paper, attention is focused on investigating the free vibration behaviors of two FG coating-substrate structures. The first one involves a two-layer plate, namely an FG layer coated on a homogeneous substrate which is simply called the FG coated plate, the other involves a three-layer plate in which an FGM is employed for the inter-medium layer and different homogeneous materials are in the top and bottom layers, this is called an FG undercoated plate [Chung (2007)]. For each single layer of these plates, the boundary-domain integral equation formulations are derived initially by using the elastostatic fundamental solutions which is based on the two-dimensional elasticity theory. A meshfree scheme is achieved to apply the radial integration method to transform the domain integrals arising from the material inhomogeneous and the inertial effects to the boundary integrals. Finally, an eigenvalue system involving system matrices with boundary integrals only is obtained through assembling all the sub-layer integral equations together by employing the multi-region BEM. By the harmonious combination of this meshfree boundary-domain integral equation method and the multi-region BEM, a comparative study of FG coating versus homogeneous coating is conducted. Extensive numerical results are presented to demonstrate the influences of FG coating thickness ratio, plate aspect ratio, as well as boundary condition on the vibration characteristics of the FG coated and the FG uncoated plates.

Material properties of the coating-substrate structures

Three considered coating-substrate plates, namely, the homogeneous coated, FG coated, as well as the FG undercoated plates are schematic depicted in Fig. 1. Assume the layers of these coating-substrate plates are perfected bonded to each other. The total length and height of these coating-substrate structures are denoted by L and h_t. h_i represents the thickness of each layer. The coating and the substrate of the homogeneous coated plate as well as the top and the bottom layers of the FG undercoated plate are composed by pure ceramic and pure steel, respectively, there material parameters are described in Table 1. For the FG layer existing in FG coated and FG undercoated plates, assuming the top is ceramic rich and the bottom is steel rich, the Young's modulus and the mass density are varying continuously in the transverse direction according to an exponential function described in Eqs. (1) and (2), while the Poisson ratio is constant.

$$E(x_2) = E_b e^{\beta x_2} \quad \text{where} \quad \beta = \frac{1}{h_f} \ln(\frac{E_t}{E_b}), \tag{1}$$

$$\rho(x_2) = \rho_b e^{\gamma x_2} \quad \text{where} \quad \gamma = \frac{1}{h_f} \ln(\frac{\rho_t}{\rho_b}), \tag{2}$$

where E_t, ρ_t are the Young's modulus and mass density for the top face constituent of the FG layer, and E_b, ρ_b are for the bottom face constituent. FGM gradation parameters are represented by β and γ for Young's modulus and mass density respectively. x_2 denotes the Cartesian coordinates variable in the transvers direction and h_f is the thickness of the FG layer. The through thickness variation of the Young's modulus for the three considered coating-substrate plates is shown in Fig. 2.

Table 1. Material properties of the homogeneous ceramic and steel

Material	E(GPa)	ρ(Kg/m^3)	ν
Aluminum(Al)	70	2707	0.3
Steel(S)	210	7806	0.3
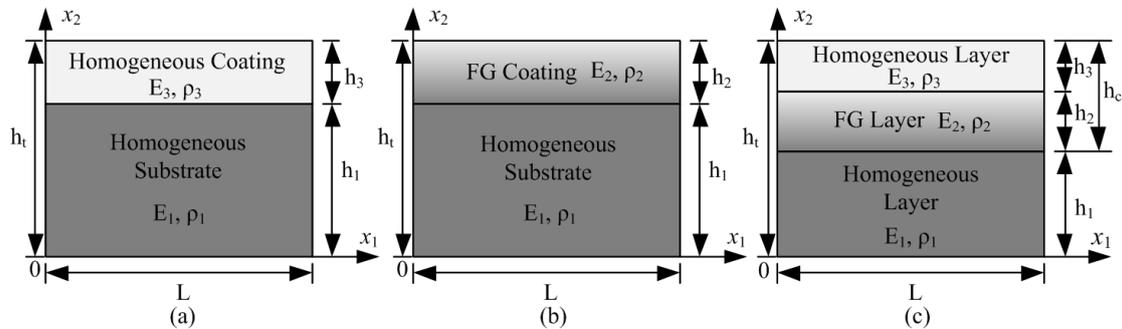
Figure 1. Coordinates and geometry of the coating-substrate plates (a) homogeneous coated plate; (b) FG coated plate; (c) FG undercoated plate
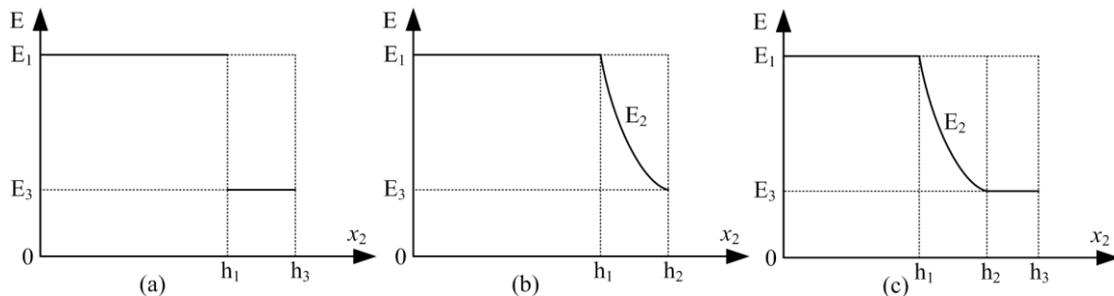
Figure 2. Variation of Young's modulus of (a) homogeneous coated plate; (b) FG coated plate; (c) FG undercoated plate

Problem formulation

The fulfillment of the free vibration analyses of the FG coated and FG undercoated plates as well as the homogeneous coated plates are by the harmonious combination of the developed meshfree boundary-domain integral equation method and the multi-region BEM.

The meshfree boundary-domain integral equation method

For each single layer of the coating-substrate plates, the governing differential equations of the steady-state elastodynamics without damping is expressed in terms of the frequency ω as

$$\sigma_{ij,j}(\boldsymbol{x}) + \omega^2 \rho u_i(\boldsymbol{x}) = 0. \tag{3}$$

Which is based on the two-dimensional elasticity theory and the stress tensor σ_{ij}, mass density ρ, displacement u_i are quantities for each layer. A comma after a quantity represents spatial derivatives and repeated indexes denote summation.

The elasticity tensor c_{ijkl} is described in the form of

$$c_{ijkl}(\boldsymbol{x}) = \mu(\boldsymbol{x})c_{ijkl}^0 \quad \text{where} \quad c_{ijkl}^0 = \frac{2v}{1-2v}\delta_{ij}\delta_{kl} + \delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}, \tag{4a, b}$$

where c_{ijkl}^0 represents the elastic tensor of the reference homogenous material, which is a "fictitious" homogeneous material with $\mu=1$. $\mu(\boldsymbol{x}) = E(\boldsymbol{x})/2(1+v)$ is the shear modulus. For the FG layer, $\mu(x)$ varies gradationally according to the coordinates, while it keeps a constant for the homogeneous layer. δ_{ij} is the Kronecker delta. By taking the elastostatic displacement fundamental solutions $U_{ij}(\boldsymbol{x}, \boldsymbol{y})$ as the weight function, the weak-form of the equilibrium Eq. (3) can be obtained as

$$\int_{\Omega}[\sigma_{jk,k} + \rho\omega^2 u_j]\cdot U_{ij}d\Omega = 0. \tag{5}$$

Application of the generalized Hooke's law $\sigma_{ij} = c_{ijkl}u_{k,l} = \mu(\boldsymbol{x})c_{ijkl}^0 u_{k,l}$ and the Gauss's divergence theorem yields the following boundary-domain integral equations

$$\tilde{u}_i(\boldsymbol{y}) = \int_{\Gamma}U_{ij}(\boldsymbol{x},\boldsymbol{y})t_j(\boldsymbol{x})d\Gamma - \int_{\Gamma}T_{ij}(\boldsymbol{x},\boldsymbol{y})\tilde{u}_j(\boldsymbol{x})d\Gamma + \int_{\Omega}V_{ij}(\boldsymbol{x},\boldsymbol{y})\tilde{u}_j(\boldsymbol{x})d\Omega$$
$$+\omega^2\int_{\Omega}\frac{\rho(\boldsymbol{x})}{\mu(\boldsymbol{x})}U_{ij}(\boldsymbol{x},\boldsymbol{y})\tilde{u}_j(\boldsymbol{x})d\Omega \tag{6}$$

In Eq. (6), the traction vector $t_i = \sigma_{ij}n_j$, n_j is the components of the outward unit normal to the boundary Γ of the considered domain Ω. $\tilde{u}_i$ is recognized as the normalized displacement vector correlating with the normalized shear modulus $\tilde{\mu}$, which are defined by [Gao (2008)]

$$\tilde{u}_i(\boldsymbol{x}) = \mu(\boldsymbol{x})u_i(\boldsymbol{x}), \qquad \tilde{\mu}(\boldsymbol{x}) = \ln[\mu(\boldsymbol{x})]. \tag{7a, b}$$

The fundamental solutions arising in equation (6) can be expressed as following, where $U_{ij}(\boldsymbol{x}, \boldsymbol{y})$ and $T_{ij}(\boldsymbol{x}, \boldsymbol{y})$ are chosen as the elastostatic displacement fundamental solutions for homogeneous, isotropic and linear elastic solids with $\mu=1$ [Gao and Davies (2002)].

$$U_{ij} = \frac{-1}{8\pi(1-v)}[(3-4v)\delta_{ij}\ln(r) - r_{,i}r_{,j}], \tag{8}$$

$$\Sigma_{ijl} = c_{rsjl}^0 U_{ir,s} = \frac{-1}{4\pi(1-v)r}[(1-2v)(\delta_{il}r_{,j} + \delta_{ij}r_{,l} - \delta_{jl}r_{,i}) + 2r_{,i}r_{,j}r_{,l}], \tag{9}$$

$$T_{ij} = \Sigma_{ijl}n_l = \frac{-1}{4\pi(1-v)r}[(1-2v)(n_ir_{,j} - n_jr_{,i}) + ((1-2v)\delta_{ij} + 2r_{,i}r_{,j})r_{,l}n_l], \tag{10}$$

$$V_{ij} = \Sigma_{ijl}\tilde{\mu}_{,l} = \frac{-1}{4\pi(1-v)r}[(1-2v)(\tilde{\mu}_{,i}r_{,j} - \tilde{\mu}_{,j}r_{,i}) + ((1-2v)\delta_{ij} + 2r_{,i}r_{,j})r_{,l}\tilde{\mu}_{,l}], \tag{11}$$

where $r=|\boldsymbol{x}\text{-}\boldsymbol{y}|$ is the distance from the field point $\boldsymbol{x}$ to the source point $\boldsymbol{y}$. Boundary-domain integral equations for boundary points can be obtained by letting $\boldsymbol{y}$ to the boundary Γ in Eq. (6). There are two domain integrals emerged in the Eq. (6), the first one is due to the material inhomogeneous and the other arises from the inertial effect.

Homogeneous layer

Respect to the homogeneous layer, in the virtue of the shear modulus is a constant through the medium, therefore, $\tilde{\mu}_{,l}$ appears a zero value in Eq. (11), which leads to the integral kernel V_{ij} inside the first domain integral of Eq. (6) vanish. Only the domain integral arises from the inertial effect

left with a constant ρ_h/μ_h (ρ_h, μ_h are mass density and shear modulus for the homogenous layer), which can be extracted out of the domain integration.

FG layer

For the FG layer, the emerged two domain integrals are all remained. In the first domain integral, $\tilde{\mu}_{,l}$ is no longer a zero value. However, in the case of an exponential law for the Young's modulus or shear modulus such as those used in this analysis, it can be seen from Eq. (7b) that $\tilde{\mu}_{,l}$ is constant and $V_{ij}(\boldsymbol{x},\boldsymbol{y})$ thus becomes very simple for integration. While the material properties ratio $\rho(\boldsymbol{x})/\mu(\boldsymbol{x})$ in the second domain integral still need to be consideration inside the domain integration due to the material properties are varying according with the coordinates.

In order to treat the domain integrals in the Eq. (6), the radial integration method (RIM) proposed by Gao [2002] is employed to transform the domain integrals into the boundary integrals over the global boundary. In the RIM, $\tilde{u}_i$ in the domain integrals of Eq. (6) are approximated by a combination of the radial basis function and the polynomials of the global coordinates as

$$\tilde{u}_i(\boldsymbol{x}) = \sum_A \alpha_i^A \phi^A(R) + a_i^k x_k + a_i^0 , \quad \sum_A \alpha_i^A = 0 , \quad \sum_A \alpha_i^A x_j^A = 0 . \qquad (12a, b, c)$$

In this analysis, the 4th order spline-type radial basis function $\phi^A(R)$ is applied. By taking all the boundary nodes (N_b) and some internal nodes (N_i) to constitute the application points A ($N_t=N_b+N_i$), and substituting the coordinates of the field points $\boldsymbol{x}$ (x_k) and the application point A (x_j^A) into Eqs. (12), if with no two coincide noes, the unknown coefficient vectors can be calculated by a set of linear algebraic equations as

$$\tilde{\boldsymbol{u}} = \boldsymbol{\phi} \cdot \boldsymbol{\alpha} , \qquad \text{and} \qquad \boldsymbol{\alpha} = \boldsymbol{\phi}^{-1} \cdot \tilde{\boldsymbol{u}} . \qquad (13a, b)$$

Subsequently determining the coefficients α_i^A, a_i^k and a_i^0, substitute Eq. (12a) into the domain integrals of Eq. (6) and apply the RIM, the two domain integrals are transformed into the boundary integrals in the form of [Yang et al. (2014)]

$$\int_\Omega V_{ij}\tilde{u}_j d\Omega = \alpha_j^A \int_\Gamma \frac{1}{r}\frac{\partial r}{\partial n} F_{ij}^{A2} d\Gamma + a_j^k \int_\Gamma \frac{r_{,k}}{r}\frac{\partial r}{\partial n} F_{ij}^{12} d\Gamma + (a_j^k y_k + a_j^0)\int_\Gamma \frac{1}{r}\frac{\partial r}{\partial n} F_{ij}^{02} d\Gamma , \qquad (14)$$

$$\omega^2 \int_\Omega \frac{\rho}{\mu} U_{ij}\tilde{u}_j d\Omega = \omega^2 \frac{\rho_b}{\mu_b} e^{(\gamma-\beta)y_2}[\alpha_j^A \int_\Gamma \frac{1}{r}\frac{\partial r}{\partial n} P_{ij}^{A2} d\Gamma + a_j^k \int_\Gamma \frac{r_{,k}}{r}\frac{\partial r}{\partial n} P_{ij}^{12} d\Gamma +$$
$$(a_j^k y_k + a_j^0)\int_\Gamma \frac{1}{r}\frac{\partial r}{\partial n} P_{ij}^{02} d\Gamma] , \qquad (15)$$

where the relation $x_i=y_i+r_{,i}r$ is used to relate x with r. By rewriting Eq. (11) with $V_{ij} = \bar{V}_{ij}/r$, the integral functions in Eqs. (14) and (15) can be expressed as [Yang et al. (2014)]

$$F_{ij}^{A2} = \int_0^r rV_{ij}\phi^{A2}dr = \bar{V}_{ij}\int_0^r \phi^{A2}dr , \quad F_{ij}^{12} = \int_0^r r^2 V_{ij}dr = \frac{1}{2}r^2\bar{V}_{ij} , \quad F_{ij}^{02} = \int_0^r rV_{ij}dr = r\bar{V}_{ij} , \quad (16a,b,c)$$

$$P_{ij}^{A2} = \int_0^r rU_{ij}\phi^A e^{(\gamma-\beta)r_2 r}dr , \quad P_{ij}^{12} = \int_0^r r^2 U_{ij}e^{(\gamma-\beta)r_2 r}dr , \quad P_{ij}^{02} = \int_0^r rU_{ij}e^{(\gamma-\beta)r_2 r}dr . \qquad (17a,b,c)$$

Since $r_{,i}$ in the above radial integrals is constant, then Eqs. (16b, c) can be evaluated analytically and other integrals are calculated by standard Gaussian quadrature formula [Yang et al. (2014)]. Therefore the displacement boundary integral equations with only boundary integrals are obtained.

$$c_{ij}\tilde{u}_j = \int_\Gamma U_{ij}t_j d\Gamma - \int_\Gamma T_{ij}\tilde{u}_j d\Gamma + [\alpha_j^A \int_\Gamma \frac{1}{r}\frac{\partial r}{\partial n} F_{ij}^{A2} d\Gamma + a_j^k \int_\Gamma \frac{r_{,k}}{r}\frac{\partial r}{\partial n} F_{ij}^{12} d\Gamma$$

$$+ (a_j^k y_k + a_j^0) \int_\Gamma \frac{1}{r}\frac{\partial r}{\partial n} F_{ij}^{02} d\Gamma] + \omega^2 \frac{\rho_b}{\mu_b} e^{(\gamma-\beta)y_2} [\alpha_j^A \int_\Gamma \frac{1}{r}\frac{\partial r}{\partial n} P_{ij}^{A2} d\Gamma + a_j^k \int_\Gamma \frac{r_{,k}}{r}\frac{\partial r}{\partial n} P_{ij}^{12} d\Gamma . \quad (18)$$

$$+ (a_j^k y_k + a_j^0) \int_\Gamma \frac{1}{r}\frac{\partial r}{\partial n} P_{ij}^{02} d\Gamma]$$

After the spatial discretization of the each layer boundary into quadratic boundary elements with N_b boundary nodes, collocating the resulting boundary integral equations at the N_t boundary and internal nodes, two sets of discretized boundary integral equations are obtained, which can be expressed in the matrix from as

$$([H_b \quad 0]-[V_b])_{2N_b\times 2N_t} \begin{Bmatrix} \tilde{u}_b \\ \tilde{u}_i \end{Bmatrix}_{2N_t\times 1} - \omega^2 [P_b]_{2N_b\times 2N_t} \{\tilde{u}\}_{2N_t\times 1} = [G_b]_{2N_b\times 2N_b} \{t_b\}_{2N_b\times 1}, \quad (19a)$$

$$([H_i \quad I]-[V_i])_{2N_i\times 2N_t} \begin{Bmatrix} \tilde{u}_b \\ \tilde{u}_i \end{Bmatrix}_{2N_t\times 1} - \omega^2 [P_i]_{2N_i\times 2N_t} \{\tilde{u}\}_{2N_t\times 1} = [G_i]_{2N_i\times 2N_b} \{t_b\}_{2N_b\times 1}, \quad (19b)$$

where I is the identity matrix with the size of $2N_i$ x $2N_i$. $\{\tilde{u}_b\}$, $\{\tilde{u}_i\}$ and $\{\tilde{u}\}$ are the displacement vectors of the boundary nodes, internal nodes and applications points respectively. By considering the boundary conditions, the sub-columns of the coefficient matrices respect to the known displacements nodes should be interchanged with that respected to the tractions, so do the displacements and the tractions vectors. Meanwhile, it is noticed that the sub-columns of the matrices $[V_b]$, $[P_b]$ and $[V_i]$, $[P_i]$ corresponding to the known boundary displacement nodes should be taken as zero. Then Eqs. (19) lead to the following system of linear algebraic equations

$$[A_b]\begin{Bmatrix} \bar{x}_b \\ \tilde{u}_i \end{Bmatrix} - \omega^2 [P_b]\{\tilde{u}\} = [B_b]\{\bar{y}_b\} \quad \text{for boundary nodes}, \quad (20a)$$

$$[A_i]\begin{Bmatrix} \bar{x}_b \\ \tilde{u}_i \end{Bmatrix} - \omega^2 [P_i]\{\tilde{u}\} = [B_i]\{\bar{y}_b\} \quad \text{for internal nodes.} \quad (20b)$$

It is convenient to find that the traction vector $\{\bar{y}_b\}$ in Eq. (20b) which is for the internal nodes is the same with that for the boundary nodes, such that the boundary nodes traction vector can be expressed in the terms of the coefficient matrices of Eq. (20a) by multiply the $[B_b]^{-1}$, and a new relationship can be set up by the equations

$$[\bar{A}_b]\begin{Bmatrix} \bar{x}_b \\ \tilde{u}_i \end{Bmatrix} - \omega^2 [\bar{P}_b]\{\tilde{u}\} = \{\bar{y}_b\} \quad \text{for boundary nodes}, \quad (21a)$$

$$[\bar{A}_i]\begin{Bmatrix} \bar{x}_b \\ \tilde{u}_i \end{Bmatrix} - \omega^2 [\bar{P}_i]\cdot\{\tilde{u}\} = 0 \quad \text{for internal nodes}, \quad (21b)$$

where

$$[\bar{A}_b] = [B_b]^{-1}[A_b] \qquad\qquad [\bar{P}_b] = [B_b]^{-1}[P_b], \quad (22a)$$

$$[\bar{A}_i] = [A_i] - [B_i][B_b]^{-1}[A_b] \qquad [\bar{P}_i] = [P_i] - [B_i][B_b]^{-1}[P_b]. \quad (22b)$$

In each single region, all the nodes could be divided into three sets as shown in Fig 3. The first set includes the boundary nodes solely associated with a single region. This set nodes are denoted by 's'. The remaining boundary nodes reside on region-to-region interfaces belong to the second set which are denoted by 'c'. The third set is formed by the internal nodes which are denoted by 'i'. Then Eqs. (21) can be rewritten in the form of these three sets nodes as

$$\begin{bmatrix} \bar{A}_{bss} & \bar{A}_{bsc} & \bar{A}_{bsi} \\ \bar{A}_{bcs} & \bar{A}_{bcc} & \bar{A}_{bci} \\ \bar{A}_{is} & \bar{A}_{ic} & \bar{A}_{ii} \end{bmatrix} \begin{Bmatrix} \bar{x}_{bs} \\ \tilde{u}_{bc} \\ \tilde{u}_{i} \end{Bmatrix} - \omega^2 \begin{bmatrix} \bar{P}_{bss} & \bar{P}_{bsc} & \bar{P}_{bsi} \\ \bar{P}_{bcs} & \bar{P}_{bcc} & \bar{P}_{bci} \\ \bar{P}_{is} & \bar{P}_{ic} & \bar{P}_{ii} \end{bmatrix} \begin{Bmatrix} \bar{x}_{bs} \\ \tilde{u}_{bc} \\ \tilde{u}_{i} \end{Bmatrix} = \begin{Bmatrix} \bar{y}_{bs} \\ t_{bc} \\ 0 \end{Bmatrix}. \tag{23}$$

In Eq. (23), $\bar{x}_{bs}$ contains the unknown normalized displacements and the unknown traction vectors for the first set boundary nodes, while $\bar{y}_{bs}$ contains all the known vectors. In the free vibration analysis, only the homogeneous system of the linear algebraic equations is needed, which can be obtained by taking the vectors $\bar{y}_{bs}$ containing the known normalized boundary displacements as well as the known boundary tractions to be zero. However, the normalized displacements $\tilde{u}_{bc}$ and traction t_{bc} for the set two nodes are all unknown.
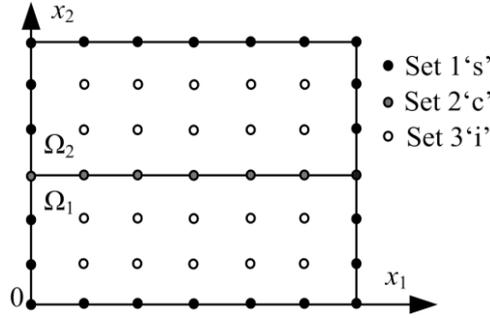
Figure 3. Node sets of the coting-substrate plates

Assemble the system of equations by the multi-region BEM

After obtaining the system linear algebraic equations for the each single layer separately, the multi-region BEM is then employed to assemble the stiffness matrix and mass matrix for the whole coating-substrate plates. Taking a two-layer FG coated plate as an example. The divided boundaries are described in Fig. 4. The boundary of the each layer is discretized into two sub-boundaries Γ_1, Γ_2 and Γ_3, Γ_4, where Γ_2 and Γ_3 are the common interface. Let u_i and t_i denote the nodal displacement and the traction vectors on boundary Γ_i respectively. The boundary integral equations can be written together in the matrix form

$$\begin{bmatrix} \bar{A}_{b11} & \bar{A}_{b12} & \bar{A}_{b1i1} \\ \bar{A}_{b21} & \bar{A}_{b22} & \bar{A}_{b2i1} \\ \bar{A}_{i11} & \bar{A}_{i12} & \bar{A}_{i1i} \end{bmatrix} \begin{Bmatrix} \bar{x}_{b1} \\ \tilde{u}_{b2} \\ \tilde{u}_{i1} \end{Bmatrix} - \omega^2 \begin{bmatrix} \bar{P}_{b11} & \bar{P}_{b12} & \bar{P}_{b1i1} \\ \bar{P}_{b21} & \bar{P}_{b22} & \bar{P}_{b2i1} \\ \bar{P}_{i11} & \bar{P}_{i12} & \bar{P}_{i1i} \end{bmatrix} \begin{Bmatrix} \bar{x}_{b1} \\ \tilde{u}_{b2} \\ \tilde{u}_{i1} \end{Bmatrix} = \begin{Bmatrix} 0 \\ t_{b2} \\ 0 \end{Bmatrix} \quad \text{for } \Omega_1, \ m=1,2, \tag{24a}$$

$$\begin{bmatrix} \bar{A}_{b33} & \bar{A}_{b34} & \bar{A}_{b3i2} \\ \bar{A}_{b43} & \bar{A}_{b44} & \bar{A}_{b4i2} \\ \bar{A}_{i33} & \bar{A}_{i34} & \bar{A}_{i2i} \end{bmatrix} \begin{Bmatrix} \tilde{u}_{b3} \\ \bar{x}_{b4} \\ \tilde{u}_{i2} \end{Bmatrix} - \omega^2 \begin{bmatrix} \bar{P}_{b33} & \bar{P}_{b34} & \bar{P}_{b3i2} \\ \bar{P}_{b43} & \bar{P}_{b44} & \bar{P}_{b4i2} \\ \bar{P}_{i33} & \bar{P}_{i34} & \bar{P}_{i2i} \end{bmatrix} \begin{Bmatrix} \tilde{u}_{b3} \\ \bar{x}_{b4} \\ \tilde{u}_{i2} \end{Bmatrix} = \begin{Bmatrix} t_{b3} \\ 0 \\ 0 \end{Bmatrix} \quad \text{for } \Omega_2, \ m=3,4 \tag{24b}$$

Then takes into account the interface equilibrium and compatibility conditions for the tractions and displacements shared by Ω_1 and Ω_2

$$t_{b2} = -t_{b3}, \qquad \tilde{u}_{b2} = \tilde{u}_{b3} \qquad \text{between } \Omega_1 \text{ and } \Omega_2 \tag{25}$$

Finally yields a 2N x 2N (N contains all the boundary nodes and the internal nodes for the whole coating-substrate plate) generalized eigenvalue system.

$$[K]\{X\} = \omega^2 [M]\{X\}, \tag{26}$$

where

$$[K] = \begin{bmatrix} \bar{A}_{b11} & \bar{A}_{b12} & 0 & \bar{A}_{b1i1} & 0 \\ \bar{A}_{b21} & \bar{A}_{b22} + \bar{A}_{b33} & \bar{A}_{b34} & \bar{A}_{b2i1} & \bar{A}_{b3i2} \\ 0 & \bar{A}_{b43} & \bar{A}_{b44} & 0 & \bar{A}_{b4i2} \\ \bar{A}_{i11} & \bar{A}_{i12} & 0 & \bar{A}_{ili} & 0 \\ 0 & \bar{A}_{i33} & \bar{A}_{i34} & 0 & \bar{A}_{i2i} \end{bmatrix} \quad [M] = \begin{bmatrix} \bar{P}_{b11} & \bar{P}_{b12} & 0 & \bar{P}_{b1i1} & 0 \\ \bar{P}_{b21} & \bar{P}_{b22} + \bar{P}_{b33} & \bar{P}_{b34} & \bar{P}_{b2i1} & \bar{P}_{b3i2} \\ 0 & \bar{P}_{b43} & \bar{P}_{b44} & 0 & \bar{P}_{b4i2} \\ \bar{P}_{i11} & \bar{P}_{i12} & 0 & \bar{P}_{ili} & 0 \\ 0 & \bar{P}_{i33} & \bar{P}_{i34} & 0 & \bar{P}_{i2i} \end{bmatrix} \quad (27ab)$$

$$\{X\} = \{\bar{x}_{b1} \quad \tilde{u}_{b2} \quad \bar{x}_{b4} \quad \tilde{u}_{i1} \quad \tilde{u}_{i2}\}^T . \tag{27c}$$
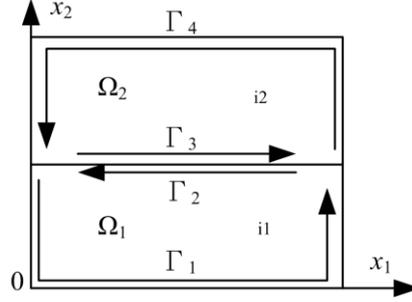
Figure 4. Boundary discretization of a two-layered FG coated plates

By resolving this general eigenvalue equation, the eigenvalue ω and the eigenvector $\{X\}$ for the coating-substrate plates can be obtained numerically.

Numerical analysis and discussion

Two numerical analysis examples make up this section. The first one is conducted by the comparative study of the 2D homogeneous coated and the FG coated plates. The other investigates the free vibration behaviors of the FG undercoated plates. The boundary conditions for these three plates are notated by the combination of four edges boundary situations moving counter clockwise starting from the edges x_2=0. The simply supported (S), fixed (C) and free (F) boundary conditions are imposed as below and shown in Fig. 5.

S: t_{x1}=0, v=0, on x_1=0;   C: u=v=0, on x_1=0;   F: t_{x1}= t_{x2}=0, on x_1=0.

A developed FORTRAN program [Yang et al. (2014)] is using to fulfill this numerical evaluation and plane-strain condition is considered throughout this study. The natural frequencies are all normalized by

$$\varpi = \omega h_t \sqrt{\rho_{Al}/E_{Al}} , \tag{28}$$

where h_t is the total thickness of the analyzed plate.
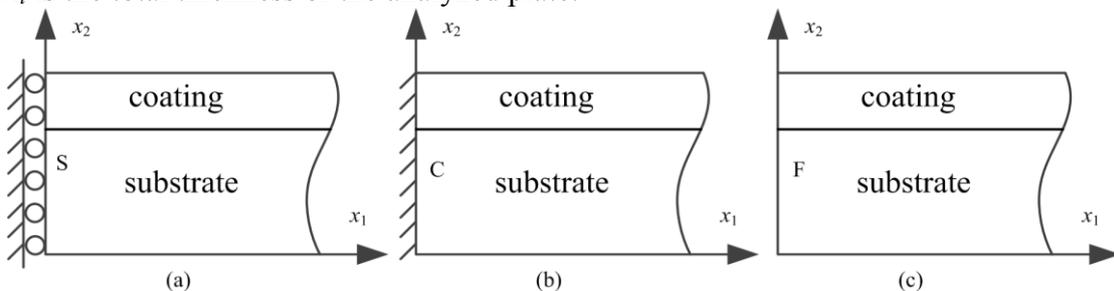
Figure 5. Different supports for coating-substrate plates (a) simply supported; (b) fixed; (c) free

Verify the accuracy of the results

In order to verify the accuracy of the present method, the results evaluated by the developed meshfree boundary-domain integral equation method are used to compare with that calculated by the traditional finite element method (FEM). The first ten normalized natural frequencies of the SSSS supported homogeneous coated plate (HCP), FG coated plate (FCP) and the FG undercoated plate (FUCP) with $L/h_t=1$ are shown in Table 2. In this study, the coating thickness ratio considered for HCP is $h_3/h_t=0.5$, for FCP is $h_2/h_t=0.5$, and for FUCP is $h_c/h_t=0.5$, $h_2/h_t=0.6$. From the Table 2, it can be seen that the results of the present methods have a great agree with that of the FEM, even for the high frequencies.

Table 2. Comparison the normalized frequencies of the homogeneous coated plate, FG coated plate and the FG undercoated plate

HCP	FEM	Error(%)	FCP	FEM	Error(%)	FUCP	FEM	Error(%)
1.8456	1.8408	0.26	1.8057	1.7999	0.32	1.8070	1.8021	0.28
1.9856	1.9772	0.42	1.9378	1.9270	0.56	1.9042	1.8946	0.50
2.9478	2.9427	0.17	3.0070	2.9999	0.23	3.1112	3.0580	1.74
3.8201	3.8151	0.13	3.9569	3.9179	1.00	3.9257	3.8885	0.96
3.9252	3.9161	0.23	4.0271	3.9850	1.06	4.1889	4.1533	0.86
4.2923	4.2861	0.14	4.4794	4.4641	0.34	4.5159	4.5055	0.23
4.4885	4.4773	0.25	4.6930	4.6709	0.47	4.7561	4.7356	0.43
5.2895	5.2786	0.21	5.4499	5.4207	0.54	5.4261	5.4075	0.34
5.5976	5.5928	0.09	5.6657	5.6671	0.02	5.5483	5.5491	0.01
5.8012	5.7909	0.18	5.9679	5.8794	1.51	5.8907	5.8831	0.13

Comparative study for the homogeneous coated plates and the FG coated plates

For the sake of understanding the free vibration behaviors of the FG coated plates in a more comprehensive view, the free vibration of the homogeneous coated plates is also analyzed to do the comparative study. The square and the rectangular homogeneous coated and the FG coated plates with five coating thickness ratios ($h_i/h_t=0.1, 0.2, 0.3, 0.4$ and 0.5) as well as six boundary conditions are investigated in details. The square coating-substrate plates with six different boundary conditions are described in Fig. 6.
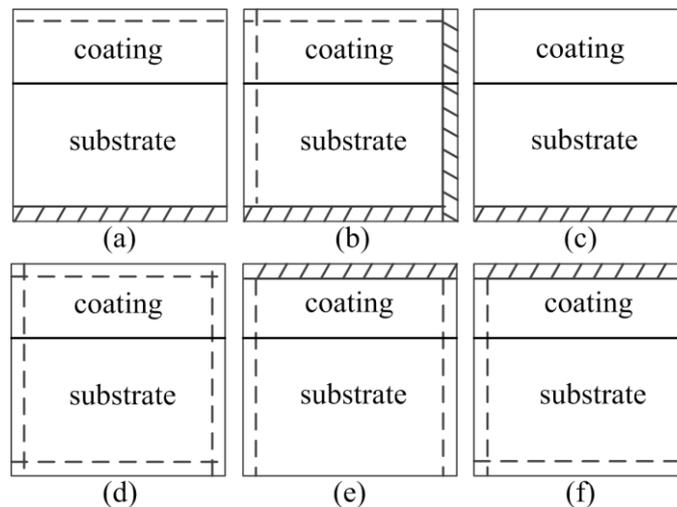
Figure 6. Boundary conditions for the FG coated plates (a) CFSF; (B)CCSS; (C)CFFF; (D)SSSS; (E)FSCS; (F)SFCS

The normalized fundamental frequency versus the coating thickness ratio of the homogeneous coated and the FG coated plates with $L/h_t=1$ and 2 are drawing in Fig. 7 and 8, respectively. From figures 7(a) and 8(a), It can be seen that, with increasing the coating thickness ratio, the normalized fundamental frequencies of the CCSS and CFFF coated plates are increased, but it is decreased for the FSCS and SFCS coated plates, while, for the other two coated plates it varies in a parabolic tendency, which is for the square coated plates. Nevertheless, for the rectangular coated plates, the variation trend of the normalized frequency according with the coating thickness ratio is the similar with that of the square one, except for the SSSS coated plates, which decrease with increasing of h_i/h_t. The plate aspect ratio effects the normalized frequencies of the coated plates in a way like that, for the CFSF and CFFF coated plates, increase the normalized frequencies with increasing the plates aspect ratios, however, it effects the CCSS and SSSS coated plates in an opposite tendency. What is more, the normalized fundamental frequencies for the FSCS coated plates makes no difference with that of the SFCS coated plates, then it can be concluded that the plate aspect ratio effects less for the SFCS and FSCS coated plates.
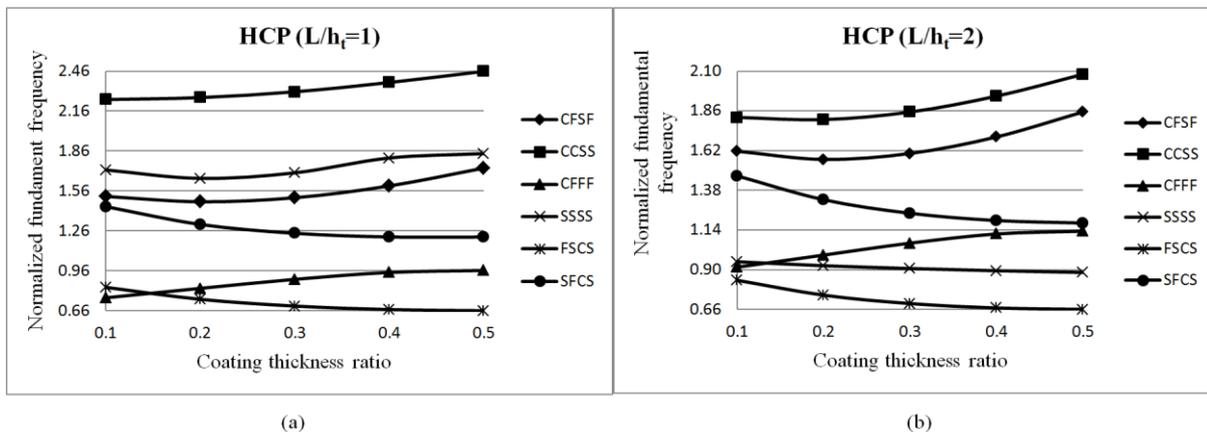
Figure 7. The normalize fundamental frequency versus coating thickness ratio of the homogeneous coated plates with different boundary conditions (a) $L/h_t=1$; (b) $L/h_t=2$
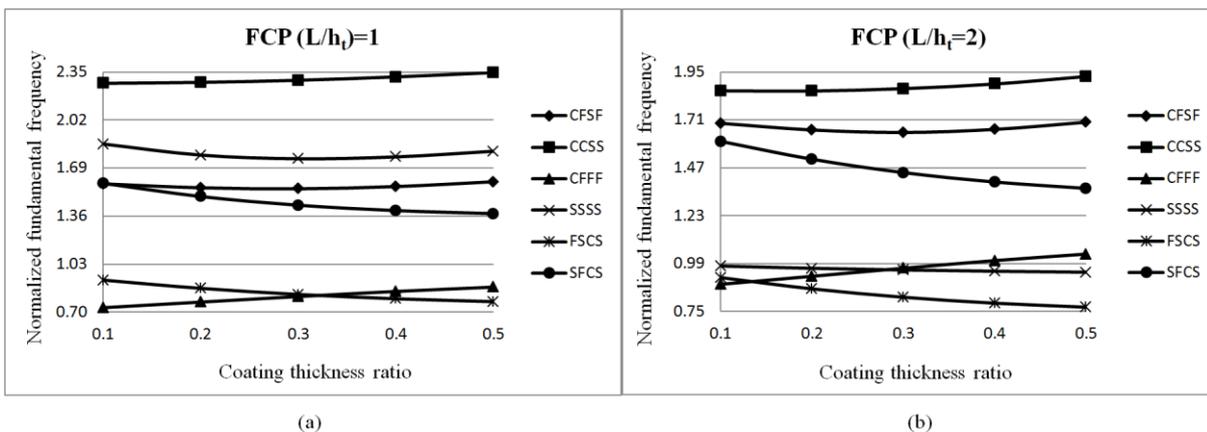
Figure 8. The normalized fundament frequency versus coating thickness ratio of the FG coated plates with different boundary conditions (a) $L/h_t=1$; (b) $L/h_t=2$

From Fig 7 and 8, it can be seen that the variation trend of the homogeneous coated plates and the FG coated plates with different aspect ratios, coating thickness ratios and boundary conditions almost in an identical way, which shed lights on that the free vibration behaviors of the coating-substrate plates will be determined based on the associated effects of the different kinds of the variables, and contrast to the others, the material properties play a weaker role. But compare the

free vibration behaviors of these two coating-substrate plates in a more detail, it will be found that, with the variation of the important parameters, the changing of the normalized natural frequencies for the FG coated plates are more temperately, that can be seen from Fig. 9. Then it can be concluded that, the coating thickness ratio, plate aspect ratio and the boundary conditions have a less effectiveness on the FG coated plates than the homogeneous one.
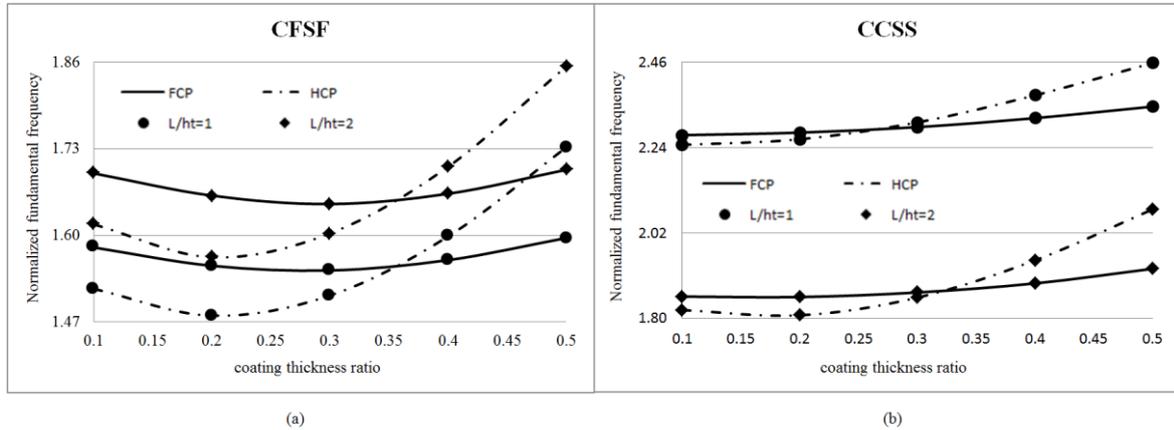
Figure 9.   Normalized fundamental frequencies versus coating thickness ratio (a) CFSF coated plates; (b) CCSS coated plates

Free vibration behaviors of the FG undercoated plates

In this section, the three-layered FG undercoated plates are investigated. For the FG undercoated plates, the thickness of the whole coating is denoted by h_c, in which the top homogeneous coating has the thickness h_3, and the bottom FG layer has the thickness h_2. In this study, fixes the coating thickness ratio, $h_c/h_t=0.5$, meanwhile, six variational $h_2/h_c=0$, 0.2, 0.4, 0.6, 0.8 and 1 are considered in the parametric study. Respect to the $h_2/h_c=0$, that is the FG layer thickness is changing to zero which refers to the homogenous coated plate and when the $h_2/h_c=1$, it refers to the FG coated plates. Two plate aspect ratios and six different boundary conditions are still used to simulate the free vibration behaviors of this FG undercoated plates. The normalized fundamental frequency versus h_2/h_c of the considered FG undercoated plates are plotted in Fig. 10.
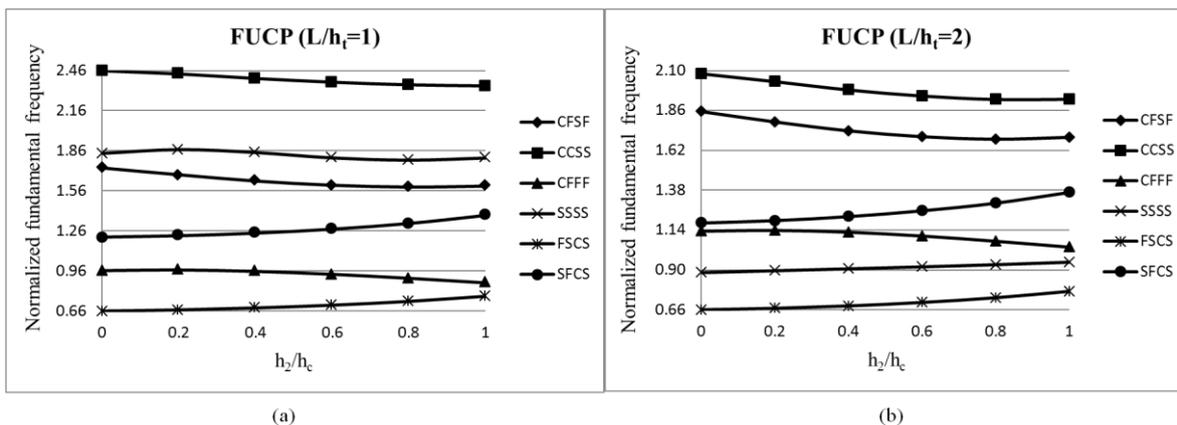
Figure 10.   The normalized fundamental frequency versus h_2/h_c of the FG undercoated plates

It is important to be noted that, the parameter h_2/h_c changing from *0* to *1*, represents the thickness of the FG layer changing from *0* to h_c. In the meantime, the rising of this parameter makes the covered proportion of the steel constituent enlarged, which directly leads to the grown up of the young's modulus and mass density for the entire FG undercoated plates. It illustrates that the larger the

528

parameter h_2/h_c, the stiffer of the FG undercoated plates. Then it can be obtained from the Fig 10 that, with the increasing of the h_2/h_t, decrease the normalized fundamental frequencies of the CFSF, CCSS and CFFF FG undercoated plates, while increase that of the FSCS and SFCS FG undercoated plates, these characters are fitting for both of the square and the rectangular FG undercoated plates. However, the SSSS FG undercoated plates is a special case, that is for the $L/h_t=1$ the normalized fundamental frequency is in an upward trend and for the $L/h_t=2$, it plays an opposite trend.

Conclusions

In this paper, the free vibration of the FG coated and the FG undercoated plates are analyzed by the developed meshfree boundary-domain integral equation method. The homogeneous coated plates are also considered to do the comparative study. These numerical analyses demonstrate that the present method is accuracy and efficiency. Based on the parametric studies, it obtained that, the free vibration behaviors of the FG coated plates and the FG undercoated plates are influenced by the associated effects of the different kinds of the important parameters, and these parameters affect the homogeneous coated plates a lot.

Acknowledgment

Reference

Shodja, H. M., Ghahremaninejad, A. (2006) An FGM coated elastic solid under thermomechanical loading: a two dimensional linear elastic approach, *Surface and Coatings Technology* **200**(12–13), 4050-4064.

Kashtalyan, M., Menshykova, M.. (2007) Three-dimensional elastic deformation of a functionally graded coating/substrate system, *International Journal of Solids and Structures* **44**(16), 5272-5288.

Chung, Y. L., Chen, W. T.. (2007) Bending behavior of FGM-coated and FGM-undercoated plates with two simply supported opposite edges and two free edges. *Composite Structures* **81**(2), 157-167.

Saizonou, C., Kouitat-Njiwa, R., von Stebut, J.. (2002) Surface engineering with functionally graded coatings: a numerical study based on the boundary element method. *Surface and Coatings Technology* 153(2–3), 290-297.

Guler, M. A., Erdogan, F.. (2004) Contact mechanics of graded coatings, *International Journal of Solids and Structures* **41**(14), 3865-3889.

Chi, S. h., Chung, Y. L.. (2003) Cracking in coating–substrate composites with multi-layered and FGM coatings. *Engineering Fracture Mechanics* **70**(10), 1227-1243.

Chen, Y. F., Erdogan, F.. (1996) The interface crack problem for a nonhomogeneous coating bonded to a homogeneous substrate. *Journal of the Mechanics and Physics of Solids* **44**(5), 771-787.

Pindera, M. J., Aboudi, J., Arnold, S. M.. (2002) Analysis of spallation mechanism in thermal barrier coatings with graded bond coats using the higher-order theory for FGMs. *Engineering Fracture Mechanics* **69**(14–16), 1587-1606.

Pindera, M. J., Aboudi, J., Arnold, S. M.. (2005) Analysis of the spallation mechanism suppression in plasma-sprayed TBCs through the use of heterogeneous bond coat architectures. *International Journal of Plasticity* **21**(6), 1061-1096.

Liew, K. M., Yang, J., Wu, Y. F.. (2006) Nonlinear vibration of a coating-FGM-substrate cylindrical panel subjected to a temperature gradient. *Computer Methods in Applied Mechanics and Engineering* **195**(9–12), 1007-1026.

Hosseini-Hashemi, S., Salehipour, H., Atashipour, S. R.. (2012) Exact three-dimensional free vibration analysis of thick homogeneous plates coated by a functionally graded layer. *Acta Mech* **223**(10), 2153-2166.

Gao, X. W., Zhang, C., Sladek, J., Sladek, V.. (2008) Fracture analysis of functionally graded materials by a BEM. *Composites Science and Technology* **68**(5), 1209-1215.

Gao, X. W., Davies, T. G.. (2002) Boundary Element Programming in Mechanics. MI: Cambridge University Press.

Yang. Y. et al.. (2014) Free vibration analysis of two-dimensional functionally graded structures by a meshfree boundary-domain integral equation method. *Composite Structures* **110**, 342-353.

Gao, X. W.. (2002) The radial integration method for evaluation of domain integrals with boundary-only discretization. *Engineering Analysis with Boundary Elements* **26**(10), 905-916.

Study on mechanical model of Nafion membrane

*I. Riku, K. Marui and K. Mimura

Graduate School of Engineering, Osaka Prefecture University,
1-1, Gakuen-cho, Naka-ku, Sakai City 599-8231, Japan

*Corresponding author: riku@me.osakafu-u.ac.jp

Abstract

To clarify the effect of density of entanglement points of molecular chains on mechanical behavior of Nafion membrane, we at first employ molecular dynamic (MD) method to constitute the computational models for Nafion membranes with different density of entanglement points of molecular chains. And then, MD simulation for Nafion membrane under simple tension is performed. The results show that relatively high deformation resistance together with a distinct yield point appears in the Nafion membrane, which has a high density of entanglement points of molecular chains.

Keywords: Nafion membrane, Mechanical behavior, Entanglement point, Molecular chain, MD

Introduction

Because of the high power density, high efficiency, fast start-up, and zero emission at the point of use, proton exchange membrane fuel cells (PEMFCs) are the most promising candidates for replacing internal combustion engines in automobiles, and are also being developed for portable and distributed stationary power generation applications. However, the life of PEMFCs is currently limited by the mechanical endurance of polymer electrolyte membranes (PEMs) [1].

The failure of PEM is believed to be the result of a combined chemical and mechanical effect acting together [2]. While chemical degradation of the membrane has been investigated and reported extensively in literature, there is little work published on mechanical degradation of the membrane. Recently, it is found that cyclic hydration of the membrane during the operation cycles (start/shut down) of the fuel cell may cause mechanical degradation of the membrane [3]. To investigating such mechanical degradation of the membrane subjected to fuel cell cycles, some microstructure analyses have been done for the membrane made from the sulfonated tetrafluoroethylene copolymer with the trade name Nafion® [4]. Nafion® consists of a hydrophobic polytetrafluoroethylene (PTFE)-like backbone and pendent chains with sulfonated (SO3−) end groups. Under humidified conditions, the hydrophilic end groups segregate into nano-sized clusters, which imbibe water and cause the swelling of the ionomer [5]. To account for the effect of such change of the microstructure of the membrane on the mechanical response, Benziger et al. [6] proposed that membrane swelling and relaxation processes work as an interfacial contact switch between the membrane and the catalyst layer. Moreover, a viscoelastic model has been developed by Lai et al. and the mechanical response predictions upon implementing the model using the data for Nafion® NR111 have been validated with stress measurements from a relaxation test performed at small initial strain (3%) in the linear elastic region [3]. However, in these studies, the computational models are just phenomenological ones and the change in the entanglement situation for the physical linkages of the molecular chain has not been accounted for explicitly.

Therefore, in this paper, to clarify the effect of density of entanglement points of molecular chains on mechanical behavior of Nafion membrane, we at first employ molecular dynamic (MD) method to constitute the computational models for Nafion membranes with different density of

entanglement points of molecular chains. And then, MD simulation for Nafion membrane under simple tension is performed and the relationship between the macroscopic yield behaviour and the movement of molecular chains is discussed.

MD Simulation Model

Figure 1 shows the structure formula of Nafion®. In this paper, polymer chains of Nafion membrane are represented by coarse-grained model, in which each bead corresponds to a group of atoms such as CF, CF_2, CF_3. The total potential function of the molecular chain of Nafion membrane are given by

$$U_{total} = U_{bond}(r) + U_{angle}(\theta) + U_{torsion}(\phi) + U_{nonbond}(\bar{r}) + U_{coulomb}(\bar{r})$$

where $U_{bond}(r)$, $U_{angle}(\theta)$, $U_{torsion}(\phi)$, $U_{nonbond}(\bar{r})$ and $U_{coulomb}(\bar{r})$ represents bond stretching energy, bending energy of successive bonds, torsion energy, van der Waals potential and Coulomb potential, respectively. These potential functions are defined as below:

$$U_{bond}(r) = \frac{1}{2}k_b(r - r_0)^2$$

$$U_{angle}(\theta) = \frac{1}{2}k_\theta(\theta - \theta_0)^2$$

$$U_{torsion}(\phi) = \sum_{n=0}^{4}\frac{1}{2}V_n\left[1 - (-1)^n\cos(n\phi)\right]$$

$$U_{nonbond}(\bar{r}) = D_0\left[\left(\frac{\sigma_0}{\bar{r}}\right)^{12} - \left(\frac{\sigma_0}{\bar{r}}\right)^{6}\right]$$

$$U_{coulomb}(\bar{r}) = \frac{q_i q_j}{4\pi\varepsilon_0\bar{r}^2}$$

where k_b, k_θ, V_n, D_0 are constants, r_0 is equilibrium bond length, σ_0 is Lennard-Jones diameter, θ_0 is equilibrium angle, q_i, q_j are the electric charge held by ith and jth bead, and ε_0 is the vacuum conductivity.

The number of the group of monomers shown in Figure 1 is prescribed as $m = 4$, $x = 10$. The number of the beads of Nafion membrane is 21,000 and the number of molecular chains is 100. To clarify the effect of the density of entanglement points of molecular chains, two different models of membrane are constituted. One is a point-poor model, in which the number of entanglement point is poor, and the other is a point-rich model, in which 50 entanglement points have been added to the point-poor model equally. Figure 2 shows the configuration of molecular chains of the point-rich model. The equation of motion is solved using the velocity Verlet algorithm with time step 2fs. The simulation is performed using a periodic boundary condition for the x and y axial directed walls of the simulation model. All the simulations are done using the coarse-grained molecular dynamics program OCTA/COGNAC [7]. Relaxation of the simulation model is carried out for 50,000 time steps under constant-temperature of 300K and density conditions (NVT-Nose Hoover ensemble).

Simulation Result

Figure 3 shows the macroscopic stress-strain relation of Nafion membrane. The point-poor model shows an elastic-like response whereas the point-rich model shows an elastoplastic-like response.

To clarify the effect of the density of entanglement points on such macroscopic response, the microscopic movement of certain molecular chains has been investigated. Figure 4 shows the difference between the position of several beads of one molecular chain of point-rich model and its corresponding position when the molecular chain moves as a rigid solid. It can be understood that there are two different patterns of the movement of the molecular chain. One pattern is shown in Figure 4(a) that the difference of the position of all the investigated beads increases gradually. The other pattern is shown in Figure 4(b) that the difference of the position of several beads once increases dramatically at the deformation stage, marked as "A", and decreases quickly at the subsequent deformation stage whereas the difference of the position of the other beads is negligible. As a result, the molecular chains that behave in the pattern shown in Figure 4(a) attribute to the elongation of Nafion membrane whereas the molecular chains behave in the pattern shown in Figure 4(b) have no attribution to the deformation of Nafion membrane but have considerable relation with the macroscopic yield behavior of Nafion membrane. Table 1 shows the number of the corresponding molecular chains of each pattern for the point-poor and the point-rich model. The fraction of the molecular chains that behaves in the pattern shown in Figure 4(b) increases from 27% to 47% when the number of entanglement points of the molecular chain increases. Based on the results shown above, we imply that the increase of the number of entanglement points of molecular chain leads to the tendency of the network of molecular chains to deform more locally at the microscopic region and consequently much more deformation of molecular chains behaving in pattern shown in Figure 4(a) and much more possibility of movement of molecular chains behaving in pattern shown in Figure 4(b), i.e. macroscopic yield behavior of Nafion membrane.
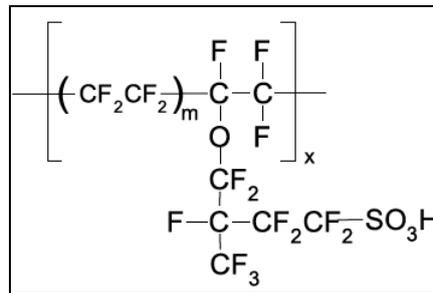
Figure 1. Structure formula of Nafion®
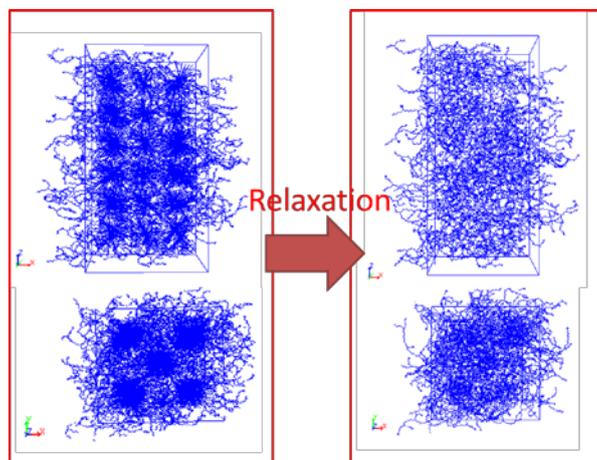
Figure 2. Configuration of molecular chains of the point-rich model of Nafion membrane
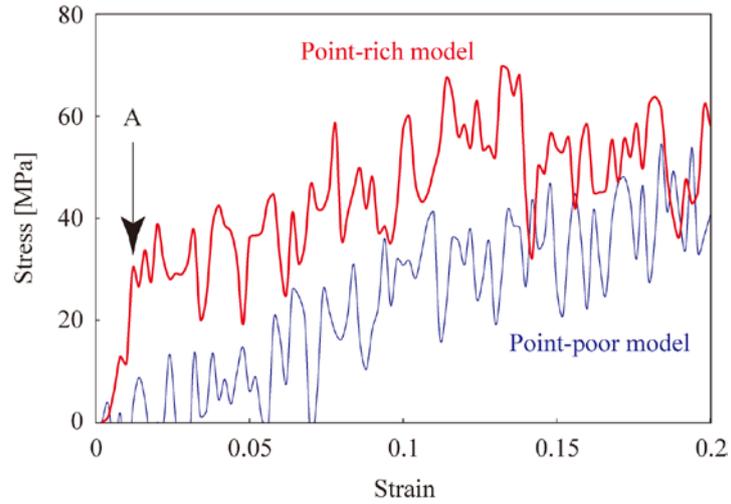
Figure 3. Macroscopic stress-strain relation of Nafion membrane
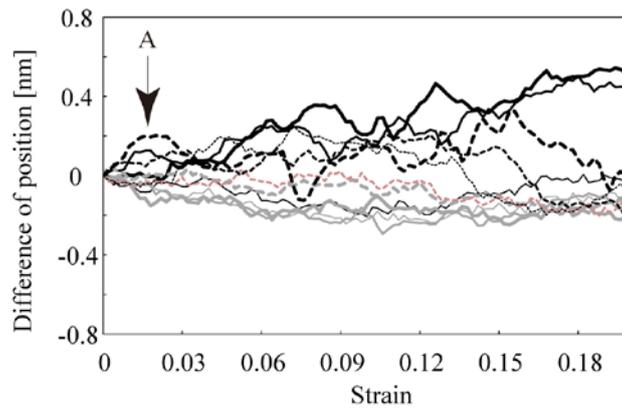
(a)
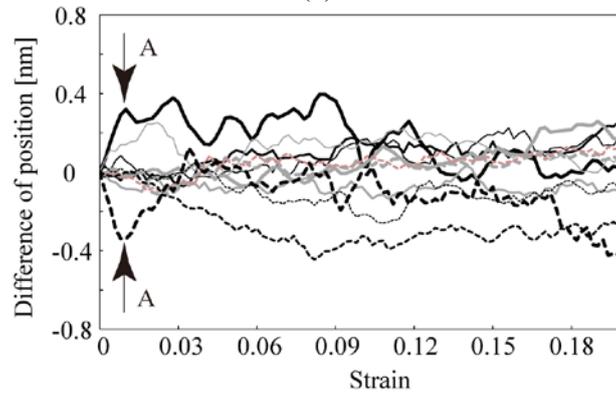
(b)

Figure 4. Movement of the beads inside the molecular chain of point-rich model of Nafion membrane

Table 1. Number of molecular chains

	Pattern shown in Figure 4(a)	Pattern shown in Figure 4(b)
Point-poor model	73	27
Point-rich model	53	47

Conclusions

In this paper, we constituted the MD computational models for Nafion membranes with different density of entanglement points of molecular chains and employed such models to clarify the relationship between the macroscopic yield behaviour and the movement of molecular chains. We found that the increase of the density of entanglement points of molecular chain leads to the tendency of the network of molecular chains to deform more locally at the microscopic region and consequently high deformation resistance together with a distinct macroscopic yield point of Nafion membrane.

References

1. Huang, X. Y., Solasi, R., Zou, Y., Feshler, M., Reifsnider, K., Condit, D., Burlatsky, S., Madden, T., Mechanical endurance of polymer electrolyte membrane and PEM fuel cell durability, J. Polym. Sci., B44, 2346-2357, (2006).
2. Liu, W., Ruth, K., Rusch, G., The membrane durability in PEM fuel cells, J. New Mater. Electrochem. Syst., 4, 227-231, (2001).
3. Lai, Y. H., Gittleman, C. S., Mittelsteadt, C. K., Dillard, D. A., Viscoelastic stress model and mechanical characterization of perfluorosulfonic acid (PFSA) polymer electrolyte membranes, Proceedings of the Third International Conference on Fuel Cell Science, Engineering and Technology, 153-159, Ypsilanti, Michigan, USA, May 23-25 (2005).
4. Gebel, G., Aldebert, P., Pineri, M., Swelling study of perfluorosulphonated ionomer membranes, Polymer, 34, 333-339, (1993).
5. Liu, D., Kyriakides, S., Case, S. W., Lesco, J. J., Li, Y. X., Mcgrath, J. E., Tensile behavior of Nafion and sulfonated poly(arylene ether sulfone) copolymer membranes and its morphological correlations, J. Polym. Sci., B44, 1453-1465, (2006).
6. Benziger, J., Chia, E., Moxley, J. F., Kevrekidis, I. G., The dynamic response of PEM fuel cells to changes in load, Chem. Eng. Sci., 60, 1743-1759, (2005).
7. Information on http://octa.jp/

Unstable mixed convection in an inclined porous channel with uniform wall heat flux

*A. Barletta[1], and M. Celli[1]

[1]Department of Industrial Engineering, Alma Mater Studiorum Università di Bologna,
Viale Risorgimento 2, 40136 Bologna, Italy

*Corresponding author: antonio.barletta@unibo.it

Abstract

The aim of this paper is to analyse the onset of convective instability in a plane porous channel inclined to the horizontal. A net upslope or downslope flow is considered, so that mixed convection takes place as caused by the uniform and symmetric heat fluxes prescribed on the impermeable bounding walls. The thermoconvective instability of the basic flow is studied versus small-amplitude wavelike perturbations. The hybrid analytical-numerical technique adopted in this paper, in order to track and illustrate the parametric changes of neutral stability curves, is Galerkin's method of weighted residuals. Numerical values at significant points on the neutral stability curves are obtained by employing an accurate Runge-Kutta solver combined with the shooting method.

Keywords: Porous Medium, Convective Instability, Mixed Convection, Normal Modes

Introduction

Convective instability induced by thermal gradients is a subject widely explored in the literature. The typical setup giving rise to unstable behaviour is one where the vertical temperature gradient is directed downward. Such a configuration may pertain to a motionless basic state, as in the classical Rayleigh-Bénard problem and its many variants [Drazin and Reid (2004)], as well as to basic forced or mixed convection flow states.

A wide research work has been done over the last sixty years to investigate thermal instability in fluid saturated porous media. Surveys of the present knowledge in this field have been written by [Nield and Bejan (2013)], as well as by [Rees (2000)], by [Tyvand (2002)] and more recently by [Barletta (2011)]. Thermoconvective instability of a basic motionless state, much like as a porous medium version of the Rayleigh-Bénard problem, has been first studied by [Horton and Rogers (1945)] and [Lapwood (1948)]. These studies were relative to a horizontal layer with impermeable and isothermal walls kept at different temperatures, and they defined what is now well-known as the Darcy-Bénard problem.

A direct extension of the Darcy-Bénard problem arises when the plane porous layer is inclined to the horizontal. The pioneering papers on this subject are [Bories and Combarnous (1973)] and [Weber (1975)]. Important developments were obtained more recently by [Rees and Bassom (2000)]. The main effect of the layer inclination is that the basic state cannot be motionless, but it is given by a stationary and parallel buoyant flow with a zero mass flow rate.

The aim of this paper is to go beyond the analysis of the Darcy-Bénard problem in an inclined porous channel by devising a setup where both walls are impermeable and symmetrically heated or cooled. The analysis to be carried out is an extension of what has been done by [Barletta (2012); (2013)], with reference to the special cases of a horizontal or vertical layer.

Mathematical model

Let us consider an inclined porous channel with infinite width and thickness H. The channel boundaries are the planes $z = 0, H$, tilted an angle $\phi \in [0, 90°]$ to the horizontal (see Fig. 1). We will assume that these boundaries are impermeable and subject to symmetric wall heat fluxes, q_0. As the value of q_0 can be either positive or negative, this may result in either a net fluid heating or cooling, respectively.

Let us define the dimensionless quantities through the scaling

$$
\begin{aligned}
\text{coordinates:} \quad & (x, y, z)\, \frac{1}{H} \to (x, y, z), \\
\text{time:} \quad & t\, \frac{\varkappa}{\sigma H^2} \to t, \\
\text{velocity:} \quad & \boldsymbol{u} = (u, v, w)\, \frac{H}{\varkappa} \to (u, v, w) = \boldsymbol{u}, \\
\text{temperature:} \quad & \frac{T - T_0}{\Delta T} \to T,
\end{aligned}
\tag{1}
$$

where T_0 is a reference constant temperature, and ΔT is a reference constant temperature difference defined as

$$
\Delta T = \frac{q_0 H}{\lambda},
\tag{2}
$$

while λ is the effective thermal conductivity, $\varkappa$ is the effective thermal diffusivity, and σ is the dimensionless ratio between the average heat capacity per unit volume of the porous medium and that of the fluid. Thus, according to the Oberbeck-Boussinesq approximation, and to Darcy's law for the momentum transfer in a porous medium, we may write the local balance equations of mass, momentum and energy in the dimensionless form

$$
\boldsymbol{\nabla} \cdot \boldsymbol{u} = 0,
\tag{3a}
$$

$$
\boldsymbol{\nabla} \times \boldsymbol{u} = R\, \boldsymbol{\nabla} \times [T (\sin \phi\, \hat{\boldsymbol{e}}_x + \cos \phi\, \hat{\boldsymbol{e}}_z)],
\tag{3b}
$$

$$
\frac{\partial T}{\partial t} + \boldsymbol{u} \cdot \boldsymbol{\nabla} T = \nabla^2 T.
\tag{3c}
$$

Equation (3b) is obtained by evaluating the curl of both sides of the local momentum balance equation in order to encompass the dependence on the pressure field. The symbols $(\hat{\boldsymbol{e}}_x, \hat{\boldsymbol{e}}_y, \hat{\boldsymbol{e}}_z)$ denote the unit vectors along the $(x, y, z)-$axes, while the Darcy-Rayleigh number, R, is defined as

$$
R = \frac{g\, \beta\, \Delta T\, K H}{\nu\, \varkappa},
\tag{4}
$$

where g is the modulus of the gravitational acceleration $\boldsymbol{g}$, ν is the kinematic viscosity, and β is the thermal expansion coefficient of the fluid.

The boundary conditions are given by

$$
\begin{aligned}
z = 0: \quad & w = 0, \quad && \frac{\partial T}{\partial z} = -1, \\
z = 1: \quad & w = 0, \quad && \frac{\partial T}{\partial z} = 1.
\end{aligned}
\tag{5}
$$
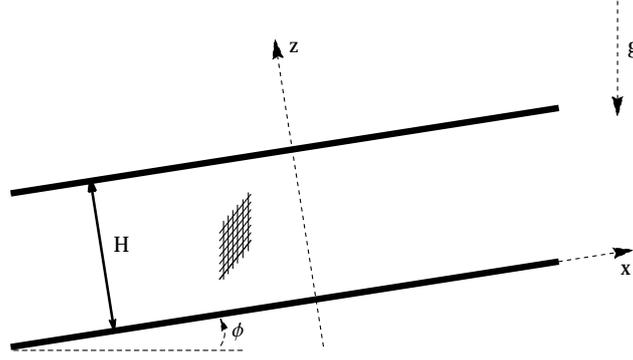
Figure 1. A sketch of the inclined porous layer

Basic solution

Equations (3) and (5) admit a time-independent basic solution expressed as

$$\boldsymbol{u}_b = F(z)\,\hat{\boldsymbol{e}}_x,$$
$$\boldsymbol{\nabla}T_b = \frac{2}{P}\,\hat{\boldsymbol{e}}_x + \frac{1}{P}\,G(z)\,\hat{\boldsymbol{e}}_z, \tag{6}$$

where "b" stands for "basic solution", and

$$F(z) = \frac{\Omega}{2}\left[P\coth\left(\frac{\Omega}{2}\right) + 2\tanh\left(\frac{\Omega}{2}\right)\cot\phi\right]\cosh(\Omega z) - \frac{\Omega}{2}\,(P + 2\cot\phi)\sinh(\Omega z), \tag{7a}$$

$$G(z) = 2\cot\phi + \left[P\coth\left(\frac{\Omega}{2}\right) + 2\tanh\left(\frac{\Omega}{2}\right)\cot\phi\right]\sinh(\Omega z) - (P + 2\cot\phi)\cosh(\Omega z), \tag{7b}$$

while the parameter Ω is given by

$$\Omega = \sqrt{\frac{2R\sin\phi}{P}}, \tag{8}$$

and P is the Péclet number defining the average velocity in the porous channel,

$$P = \int_0^1 F(z)\,dz. \tag{9}$$

Equations (6) and (7) define a horizontal through flow in the $x-$direction, with a dimensionless flow rate given by P. Equation (6) reveals that this time-independent basic solution is possible only with a nonvanishing P. This feature reflects the obvious fact that, due to the heat flux either supplied ($R > 0$) or subtracted ($R < 0$) at both boundary walls, a stationary state is possible if and only if the net heat supplied/subtracted is convected along the channel.

From Eqs. (6)–(8), a symmetry of the basic solution is revealed, namely

$$\begin{cases} z \to 1 - z \\ R \to -R \\ P \to -P \end{cases} \implies \begin{cases} \boldsymbol{u}_b \to -\boldsymbol{u}_b \\ \boldsymbol{\nabla}T_b \to -\boldsymbol{\nabla}T_b \end{cases}. \tag{10}$$

As a consequence of this symmetry, it is not restrictive to consider fluid heating conditions ($R > 0$) with either upslope mean flow ($P > 0$) or downslope mean flow ($P < 0$).
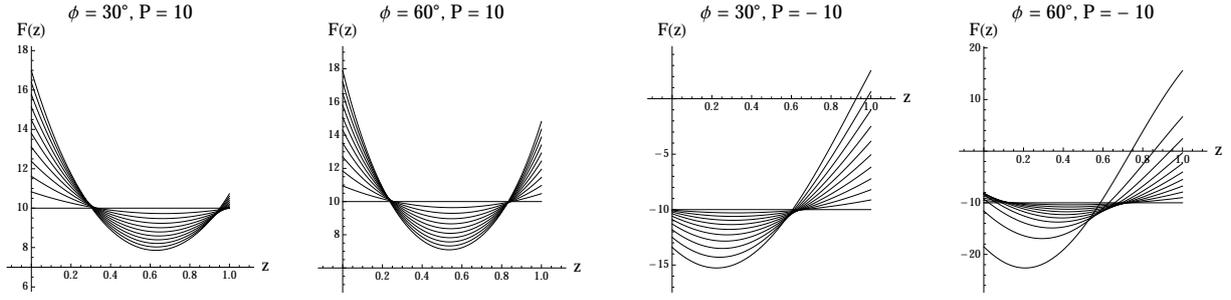
Figure 2. Plots of $F(z)$ for $P = \pm 10$, and $R = 0$ to 50 in steps of 5, with $\phi = 30°$ and $60°$

Eqs. (6)–(8) tend to the basic solution found by [Barletta (2012)] in the case of a horizontal channel ($\phi = 0°$) with the same boundary conditions.

Fig. 2 illustrates, for $\phi = 30°$ and $60°$, how the uniform velocity profiles with $R = 0$ and $P = \pm 10$ are continuously deformed as R increases up to 50, in steps of 5. A phenomenon of flow reversal is evidenced in Fig. 2 when $P = -10$ and R becomes sufficiently high.

Linear stability analysis

Let us consider small-amplitude perturbations of the basic state,

$$\boldsymbol{u} = \boldsymbol{u}_b + \varepsilon \boldsymbol{U}, \qquad T = T_b + \varepsilon \Theta, \tag{11}$$

where ε is a perturbation parameter, $\varepsilon \ll 1$, and $\boldsymbol{U} = (U, V, W)$.

Substitution of Eq. (11) in Eqs. (3) and (5), by taking into account Eqs. (6)–(8), neglecting terms $O(\varepsilon^2)$, and introducing an auxiliary scalar field Ψ such that

$$\boldsymbol{U} = R \left[\Theta \left(\sin \phi \ \hat{\boldsymbol{e}}_x + \cos \phi \ \hat{\boldsymbol{e}}_z \right) \right] - \boldsymbol{\nabla} \Psi, \tag{12}$$

yields

$$\nabla^2 \Psi = R \left(\sin \phi \ \frac{\partial \Theta}{\partial x} + \cos \phi \ \frac{\partial \Theta}{\partial z} \right), \tag{13a}$$

$$\nabla^2 \Theta = \frac{\partial \Theta}{\partial t} + F(z) \frac{\partial \Theta}{\partial x} + \frac{2}{P} \left(R \Theta \sin \phi - \frac{\partial \Psi}{\partial x} \right) + \frac{1}{P} G(z) \left(R \Theta \cos \phi - \frac{\partial \Psi}{\partial z} \right), \tag{13b}$$

$$z = 0, 1: \qquad \frac{\partial \Psi}{\partial z} = R \Theta \cos \phi, \qquad \frac{\partial \Theta}{\partial z} = 0. \tag{13c}$$

Normal mode analysis

We rely on the conclusion, drawn for the case $\phi = 0°$ (horizontal channel) and for the case $\phi = 90°$ (vertical channel) in [Barletta (2012); (2013)], to reckon that non-oscillatory longitudinal modes are the most unstable at onset. Thus, we will focus our stability analysis to these special normal modes,

$$\left\{ \begin{array}{c} \Psi \\ \Theta \end{array} \right\} = \left\{ \begin{array}{c} \psi(z) \\ \theta(z) \end{array} \right\} \exp\left[i(ky - \omega t)\right], \tag{14}$$

where k is the real-valued wave number, and ω is the complex-valued frequency. The real part of ω vanishes for non-oscillatory modes. The imaginary part of ω is the exponential growth parameter. If $\mathrm{Im}(\omega) < 0$, the normal mode describes a stability condition. On the other hand, if $\mathrm{Im}(\omega) > 0$, the normal mode yields instability. In the following, we will be interested in the marginal stability condition, so that we will assume $\mathrm{Im}(\omega) = 0$.

By defining the new functions,

$$f(z) = \psi'(z) - R\,\theta(z)\cos\phi, \qquad h(z) = k\,\theta(z), \tag{15}$$

and substituting Eq. (14) into Eqs. (13), we obtain the linear stability eigenvalue problem, namely

$$f'' - k^2 f - R k h \cos\phi = 0, \tag{16a}$$

$$h'' - \left(k^2 + \frac{2R}{P}\sin\phi\right) h + \frac{k}{P}\, G(z)\, f = 0, \tag{16b}$$

$$z = 0, 1: \qquad f = 0, \qquad h' = 0. \tag{16c}$$

In Eqs. (15) and (16), the primes denote derivatives with respect to z.

Limiting case $|P| \gg 1$

When the Péclet number is very large we obtain a dramatic simplification of Eqs. (16). In fact, on account of Eq. (7b), we may approximate

$$\frac{G(z)}{P} \approx -(1 - 2z) + \mathrm{O}(P^{-1}). \tag{17}$$

Thus, when $P \gg 1$, Eqs. (16) assume the asymptotic form

$$f'' - k^2 f - R k h \cos\phi = 0, \tag{18a}$$

$$h'' - k^2 h - k(1 - 2z) f = 0, \tag{18b}$$

$$z = 0, 1: \qquad f = 0, \qquad h' = 0. \tag{18c}$$

Eqs. (18) allow us to infer that the neutral stability condition, expressed as $R\cos\phi$ versus k, is independent of the channel inclination angle ϕ. In other words, the inclination angle influences the neutral stability condition merely by rescaling $R(k)$, as obtained for a horizontal channel, with a factor $1/\cos\phi$. The immediate consequence of this behaviour is that a gradually increasing inclination has a stabilising effect. In fact, this conclusion applies in the asymptotic regime $|P| \gg 1$, while things are more complicated for finite values of $|P|$ as we will discuss in the next sections.

Method of weighted residuals

Equations (16) can be solved by expressing f and h as,

$$f(z) = \sum_{n=1}^{N} f_n \varphi_n(z), \qquad h(z) = \sum_{n=1}^{N} h_n \chi_n(z), \tag{19}$$

where N is the truncation order, $\{f_n, h_n\}$ are complex coefficients, and the test functions $\{\varphi_n(z), \chi_n(z)\}$ have to be chosen so that the boundary conditions Eq. (16c) are satisfied,

$$z = 0, 1: \qquad \varphi_n = 0, \qquad \chi_n' = 0, \tag{20}$$

for every positive integer n. Normalised test functions are chosen, defined by

$$\varphi_n(z) = \sqrt{2}\sin[(n-1)\pi z], \qquad \chi_n(z) = \begin{cases} 1, & n = 1 \\ \sqrt{2}\cos[(n-1)\pi z], & n > 1 \end{cases} . \tag{21}$$

Substitution of Eq. (19) into Eqs. (16a) and (16b) yields the residuals

$$\sum_{n=1}^{N} \left[f_n \varphi_n''(z) - k^2 f_n \varphi_n(z) - R k h_n \chi_n(z) \cos\phi \right] = E(z), \tag{22a}$$

$$\sum_{n=1}^{N} \left[h_n \chi_n''(z) - \left(k^2 + \frac{2R}{P}\sin\phi \right) h_n \chi_n(z) + \frac{k}{P} G(z) f_n \varphi_n(z) \right] = \tilde{E}(z). \tag{22b}$$

A set of $2N$ algebraic equations is obtained by imposing that the weighted residuals are zero. Here, we adopt Galerkin's method, *i.e.* we assume that the weight functions coincide with the test functions. Thus, we can write

$$\int_0^1 E(z)\varphi_n(z)\,dz = 0, \qquad \int_0^1 \tilde{E}(z)\chi_n(z)\,dz = 0, \quad n = 1, 2, \ldots, N. \tag{23}$$
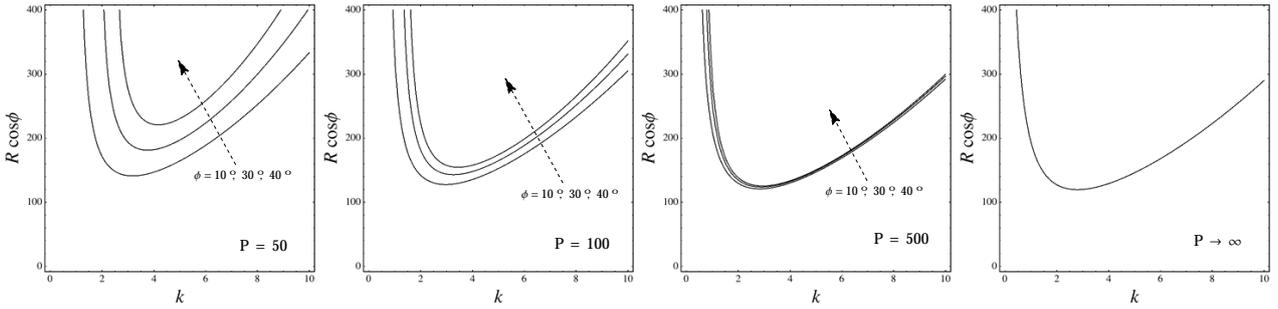
Figure 3. Neutral stability curves for $P = 50,\ 100,\ 500,\ \infty$, and $\phi = 10°$, $30°$ and $40°$
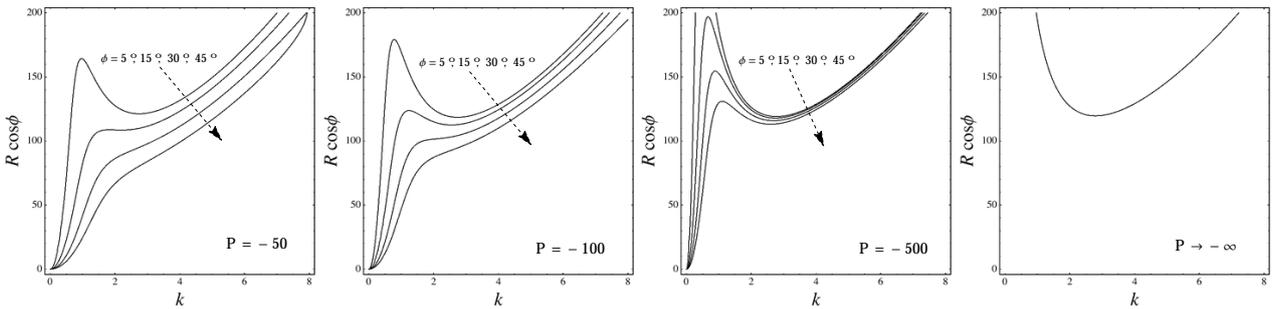
Figure 4. Neutral stability curves for $P = -50,\ -100,\ -500,\ -\infty$, and $\phi = 5°$, $15°$ $30°$ and $45°$

Eq. (23) can be rewritten in a matrix form as

$$\mathsf{M} \cdot \mathsf{c} = 0, \tag{24}$$

where M is a $2N \times 2N$ matrix and $\mathsf{c} = (f_1, f_2, \ldots, f_N, h_1, h_2, \ldots, h_N)$ is the vector of the expansion coefficients. Solving the eigenvalue problem (16) means determining non-vanishing vectors c satisfying Eq. (24), and this is accomplished when M has a vanishing determinant. Setting $\det(\mathsf{M}) = 0$

provides us with the neutral stability condition. This condition can be displayed graphically through plots of $R\cos\phi$ versus k, with prescribed values of (ϕ, P).

In Figs. 3 and 4, neutral stability curves in the $(k, R\cos\phi)$ plane are displayed, for either positive or negative Péclet numbers. Here, the computation has been carried out with a truncation order $N = 7$, which is sufficient for graphical purposes. Assessment of the solution method precision and of its convergence can be achieved by comparison with a highly accurate numerical solution based on a 4th order Runge-Kutta solver and on shooting method.

Numerical solution: combined Runge-Kutta method and shooting method

Another strategy for the solution of the eigenvalue problem (16) is achieved by combining use of an initial value ODE solver, *viz.* the 4th order Runge-Kutta method, with a technique employed to solve two-point problems, *viz.* the shooting method. This procedure is a standard approach in stability analyses. A survey can be found, for instance, in Chapter 9 of [Straughan (2010)].

The basic idea is to employ the 4th order Runge-Kutta method to solve Eqs. (16a) and (16b) with initial conditions based on an expanded form of the boundary conditions at $z = 0$ defined by Eq. (16c),

$$f(0) = 0, \quad f'(0) = \eta, \quad h(0) = 1, \quad h'(0) = 0. \tag{25}$$

Only the first and the last of these four initial conditions stem from Eq. (16c); the second one is just the definition of an unknown parameter η, while the third is a scale-fixing constraint for the eigenfunctions (f, h). In fact, the eigenfunctions are scale invariant due to the homogeneous nature of Eqs. (16), so that the setting $h(0) = 1$ can be invoked without any loss of generality. The parameter η, together with the eigenvalue $R\cos\phi$, is determined for every 3–tuple of input data (k, ϕ, P). Shooting method is employed for this task, with target constraints given by the boundary conditions at $z = 1$ and expressed by Eq. (16c), namely

$$f(1) = 0, \quad h'(1) = 0. \tag{26}$$

This procedure can be developed entirely within the *Mathematica 9* software environment (© Wolfram Research). Built-in functions `NDSolve` and `FindRoot` are the basic tools for implementing the 4th order Runge-Kutta solver, and the shooting method, respectively. This solution strategy is far more accurate than the method of weighted residuals described in the preceding section, but it requires longer computational times. Then, the method of weighted residuals is quicker and more effective for drawing plots of the neutral stability curves, while the present numerical method allows one to check its convergence and to determine accurate numerical data with given special parametric data (k, ϕ, P).

Table 1 reports the results of a convergence test for the eigenvalue $R\cos\phi$, with $\phi = 30°$ and $k = 6$. The acronym MWR denotes method of weighted residuals, while RK4 stands for combined 4th order Runge-Kutta method and shooting method. While the latter method turns out to be accurate to more than six significant figures, the former is far less precise when $N \leq 8$. On the other hand, as $N > 8$, the computational time increases significantly. Overall, the agreement between the two methods is fairly satisfactory.

Discussion of the results and concluding remarks

The main features of the linear stability analysis can be inferred from inspection of Figs. 3 and 4. The effect of a gradually increasing inclination angle ϕ is dissimilar when P is either positive or negative. When $P > 0$, the effect of inclination is stabilising, as the neutral stability values of $R\cos\phi$ increase

Table 1. Convergence test for the eigenvalue $R \cos \phi$, with $\phi = 30°$ and $k = 6$

Method	$P = -100$	$P = -50$	$P = 50$	$P = 100$
MWR ($N = 5$)	158.5	153.3	231.3	193.1
MWR ($N = 6$)	156.7	151.4	226.1	190.0
MWR ($N = 7$)	155.8	150.5	223.7	188.6
MWR ($N = 8$)	155.4	150.1	222.5	187.9
RK4	154.5902	149.3223	220.3801	186.5640

with the inclination angle ϕ. This feature reflects what we already noted in the discussion of the asymptotic case $|P| \gg 1$. When $P < 0$, even a small inclination to the horizontal may trigger the onset of instability, with a vanishingly small critical value of R. The instability is rather activated with small wave numbers, or large spatial wavelengths, as it is clearly evidenced in Fig. 4.

This behaviour implies a substantial reconsideration of the results obtained in the special case of a horizontal channel, as drawn by [Barletta (2012)]. In fact, when designing an experimental setup to test the flow stability in the horizontal case, a minimal misalignment in the channel inclination may have dramatic effects. Even a very small accidental inclination, with resulting downslope flow ($P < 0$), would imply instability. This is the case whatever small the wall heat flux and, hence, the value of R may be.

Finally, we note that the scaling law $R \sim 1/\cos \phi$, proved for the asymptotic case $|P| \gg 1$, does not hold when P is finite. Departure from this scaling law becomes more and more evident when $|P|$ is finite and decreases.

References

Barletta, A. (2011), Thermal instabilities in a fluid saturated porous medium. In Öchsner, A., Murch, G. E., editors, *Heat Transfer in Multi–Phase Materials*, pages 381–414. Springer, New York.

Barletta, A. (2012), Thermal instability in a horizontal porous channel with horizontal through flow and symmetric wall heat fluxes. *Transport in Porous Media*, **92**, 419–437.

Barletta, A. (2013), Instability of mixed convection in a vertical porous channel with uniform wall heat flux. *Physics of Fluids*, **25**, 084108.

Bories, S. A., Combarnous, M. A. (1973), Natural convection in a sloping porous layer. *Journal of Fluid Mechanics*, **57**, 63–79.

Drazin, P. G., Reid, W. H. (2004), *Hydrodynamic Stability*, 2nd edition. Cambridge University Press, New York, NY.

Horton, C. W., Rogers, F. T. (1945), Convection currents in a porous medium. *Journal of Applied Physics*, **16**, 367–370.

Lapwood, E. R. (1948), Convection of a fluid in a porous medium. *Proceedings of the Cambridge Philosophical Society*, **44**, 508–521.

Nield, D. A., Bejan, A. (2013), *Convection in Porous Media*, 4th edition. Springer, New York.

Rees, D. A. S. (2000), The stability of Darcy–Bénard convection. In Vafai, K., Hadim, H. A., editors, *Handbook of Porous Media*, chapter 12, pages 521–558. CRC Press, New York.

Rees, D. A. S., Bassom, A. P. (2000), Onset of Darcy–Bénard convection in an inclined layer heated from below. *Acta Mechanica*, **144**, 103–118.

Straughan, B. (2010), *Stability and Wave Motion in Porous Media*. Springer.

Tyvand, P. A. (2002), Onset of Rayleigh–Bénard convection in porous bodies. In Ingham, D. B., Pop, I., editors, *Transport Phenomena in Porous Media II*, chapter 4, pages 82–112. Pergamon, New York.

Weber, J. E. (1975), Thermal convection in a tilted porous layer. *International Journal of Heat and Mass Transfer*, **18**, 474–475.

Effect of GPU Communication-Hiding for SpMV Using OpenACC

*Olav Aanes Fagerlund[1], Takeshi Kitayama[2,3], Gaku Hashimoto[2] and Hiroshi Okuda[2]

[1] Department of Systems Innovation, School of Engineering, The University of Tokyo,
7-3-1 Hongo Bunkyo-ku, Tokyo 113-8656, Japan.
[2] Graduate School of Frontier Sciences, The University of Tokyo, 5-1-5 Kashiwanoha, Kashiwa,
Chiba 277-8563, Japan.
[3] Japan Science and Technology Agency, CREST, 7, Gobancho, Chiyoda-ku,
Tokyo 102-0076, Japan.

*Corresponding author: olav@multi.k.u-tokyo.ac.jp

Abstract

In the finite element method simulation we often deal with large sparse matrices. Sparse matrix-vector multiplication (SpMV) is of high importance for iterative solvers. During the solver stage, most of the time is in fact spent in the SpMV routine. The SpMV routine is highly memory-bound; the processor spends much time waiting for the needed data.

In this study, we discuss overlapping possibilities of SpMV in cases where the sparse matrix data does not fit into the memory of the discrete GPU, by using OpenACC. With GPUs one can take advantage of their relatively high memory bandwidth capabilities. However, data needs to be transferred over the relatively slow PCI express (PCIe) bus. This transfer time can to a certain degree be hidden. We concurrently perform computation on one set of data while another set of data is being transferred. Parameters such as the size of each subdivision being transferred - the number of matrix subdivisions, and the whole matrix size, are adjustable. We generate matrices modeling one, three and six degrees of freedom. It is observed how these parameters affect performance. We analyze the improved performance as a result of communication-hiding with OpenACC, and a profiler is used to provide us with additional insight. This is of direct relevance for a block Krylov solver, for instance a block Cg solver. Here, one can benefit from streaming of data with SpMV and overlap while doing so. Each streamed subdivision is used several times with different vectors.

When using a discrete GPU with an ordinary (non-block) Krylov solver, one has to run SpMV once over the whole matrix (or subdivision) for each solver iteration, so there will be no benefit if the matrix does not fit the GPU memory. This is due to the fact that streaming the matrix over the PCIe bus for each of the solver iterations incurs a too big overhead.

For instance, in the case of three degrees of freedom and modeling 2,097,152 nodes, we observe a just above 40% performance increase by applying communication-hiding in our benchmarking routine. This gives us close to 33 GFLOP/s on the AMD Tahiti GPU architecture, in double precision. When modeling the same amount of nodes with a 'synthetic' six degrees of freedom, up to ~65.7% is observed in increased performance when hiding parts of the data transfer time. This underlines the importance of applying such techniques in simulations, when it is suitable with the algorithmic structure of the problem in relation to the underlying computer architecture.

Keywords: SpMV, OpenACC, GPGPU, communication-hiding, overlapping, FEM, block Krylov

Introduction

Sparse Matrix-Vector multiplication (SpMV) is of high importance in many engineering applications. In iterative solvers it plays an important part, as this is the routine where most of the solver-time is spent. In the Finite Element Method (FEM) very large problems are frequently handled. Direct solvers will consume unmanageable amounts of memory when the problems are large, so a iterative solver is the choice here. The time to solution partly depends on a fast SpMV routine. It is an operation that can be parallelized. However, the main challenge is the memory

bottleneck found in the current day systems. For each computation in need of being performed, a significant amount of data must be read from memory. So, we cannot achieve a performance any close to the peak performance of the processor itself. The memory subsystem, and how well it is utilized, gives the limitation as to what speed can be achieved.

To partly get around the problem several sparse matrix storage formats have been developed over the years. These typically only store the non-zero elements and their locations in the matrix, thus saving storage space and bandwidth usage. The storage format of a sparse matrix can greatly influence the performance achieved, as it will dictate the memory access pattern under execution, which again affects performance according to how well that particular memory access pattern suits the underlying architecture executing the code. Today, GPUs deliver unrivaled memory bandwidth paired with massively parallel processors, so running these operations on GPUs can give a substantial advantage. Due to its importance in engineering simulation there has been extensive research in optimizing the SpMV routine [Bell and Garland (2008)]. In order to parallelize our code and also get a more performance-portable code-base we have parallelized our routines by taking advantage of a modern OpenACC compiler and its associated runtime.

One of the main obstacles that remain when using discrete GPUs is the need to transfer data over the relatively slow PCI express (PCIe) data-bus. The latency and bandwidth of this bus is at least an order of magnitude slower than the internal communication on the GPU card between the GPU chip and GPU memory ('global memory' in OpenCL terms, or 'VRAM' in more traditional terms).

For a 16-lane PCI express connection (one 16x slot) the speeds are as follows:
- PCIe 1.x : 4 GB/s
- PCIe 2.x : 8 GB/s
- PCIe 3.0 : 15.75 GB/s
- PCIe 4.0 : 31.51 GB/s (Not released yet)

Even 'on-GPU-card', where the bandwidth between the global memory and the GPU chip can be about 300 GB/s, we are severely memory bandwidth limited. We can only attain a small percentage of the peak GPU performance as we wait for the data needed for the computation. This illustrates well the problem the PCIe bus adds.

To minimize this effect we implement communication-hiding by performing needed memory transfers over the PCIe bus while a previous set of data is computed upon. The aim is to drastically reduce the overall execution time, effectively raising the performance of the SpMV operation as a whole. This can significantly benefit block Krylov solvers, described in [Saad (2003)], where one do multiplications with the same matrix and different vectors several times versus only one time with one vector per iteration as in standard Cg. The many SpMV operations can be performed while a 'new' subdivision is transferred over the PCIe bus, as they combined are time-consuming enough so that the solver as a whole benefits from the overlapping – in cases where the data simply cannot fit the available GPU memory. For instance, 'Stochastic FEM using the Seed Method', described in [Sato and Okuda (2008)], is one such block solver that could benefit from this - in its GPU-based incarnation.

Implementation of communication-hiding

A cube, consisting of smaller cubes, is modeled. The problem is a classical cantilever problem with a force applied. The cube has a certain amount of nodes, set by the number of nodes per side. The number of nodes and one, three or six degrees of freedom for all nodes are parameters set prior to matrix generation. The matrices generated will mimic stiffness matrices commonly encountered in FEM problems, and are highly sparse. Hexahedra elements are used, i.e. the nodes are in the corners of each cube, and each element has six faces. Note that for this problem, six degrees of freedom is a "synthetic" scenario. However, its results are of interest nevertheless.

We keep our data in the CSR-format, a format that is more deeply explained in [Vuduc (2003)], and from where Fig. 1 is inspired. The amount of GPU memory can serve as a limitation, as the stiffness matrix can be many times larger than the total GPU memory. Therefore, we divide the stiffness matrix up into several equal-sized subdivisions, upon transfer to the GPU over the PCI express bus. This serves two purposes:
1. Each subdivision will fit in GPU memory.
2. Communication-hiding can be performed on subdivisions subsequent to the first subdivision transfer. To enable this efficiently two blocks must fit into memory at any given time, in addition to other vectors of the CSR format needed.
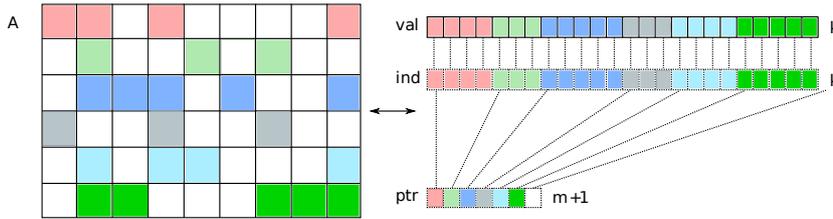
Figure 1. The CSR format layout.

The more nodes we model, the bigger the stiffness matrix becomes. The degrees of freedom decide the amount of blocked non-zero elements per 'node-to-node-connection'. One degree of freedom gives 1 x 1 elements, three degrees of freedom gives 3 x 3 elements, and six degrees of freedom gives 6 x 6 blocked non-zero elements per 'node-to-node-connection'. As we persistently use double precision, each element of the matrix consumes 8 Bytes. The basic structure of the communication-hiding is as follows. Here, for example, we explain the details of how four subdivisions are read on the fly from system main memory:

1. The stiffness matrix is divided into subdivisions. It must be, at a minimum, be divided into enough subdivisions so that two subdivisions can reside in GPU memory concurrently. In our implementation all other vectors of the CSR format are kept in GPU memory without being broken up, as the sizes are of a significantly lesser magnitude than the stiffness matrix.
2. First subdivision is transferred to GPU memory, synchronously.
3. Computation over the first subdivision (2.) is commencing once the transmission of the first subdivision is completed. However, immediately before this takes place the transfer of the second subdivision is initiated asynchronously.
4. Once computation over the first subdivision is completed, the memory space consumed by the first subdivision is de-allocated.
5. It is ensured that the transfer of the second subdivision has completed. Transfer of the third subdivision of data is initiated asynchronously.
6. Computation over the second subdivision of data is commenced and completed.
7. Once the previous step (6.) is completed, the associated second subdivision of data is de-allocated. It is ensured that the transfer of the third subdivision has completed.
8. Transfer of the fourth subdivision of data is initiated asynchronously.
9. Computation over the third subdivision of data is commenced and completed.
10. Once the previous step (9.) is completed, the associated third subdivision of data is de-allocated. It is ensured that the transfer of the forth subdivision has completed.
11. Computation over the forth subdivision of data is commenced and completed. Forth subdivision of data is de-allocated once done. All processing is completed; the complete result vector residing in GPU memory is copied into system main memory.
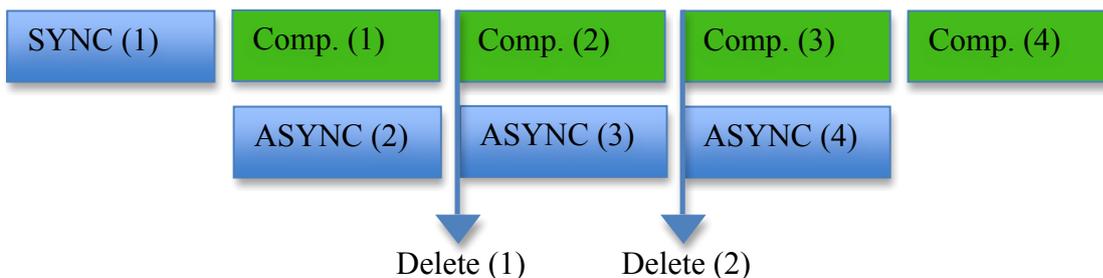
Figure 2. Illustration of the sequence with 4 subdivisions used in communication-hiding.

Fig. 2 shows the sequence explained in detail above, point 1-11. Blue boxes illustrates data transfers, while the green boxes illustrates the computation over the data of the same ID number.

Throughout these stages appropriate 'book keeping' is done to ensure correctness, required when the stiffness matrix is broken up and computed upon piece by piece. The sequences perform in a similar pattern when the stiffness matrix is divided into more subdivisions.

It is important to note that our routine will do computations over the data 50 times, for each subdivision transferred over the PCI Express bus. For each of these 50 iterations, all of the values in the vector being multiplied with the matrix are modified, so each iteration yield different result. Thus, the transfer time over the PCI Express bus is partly amortized. With the overlapping applied it is further partly hidden. The behavior mimics that of a block Krylov solver performing SpMV. Of course, if we reduce the amount of iterations, and then consequently the amount of GPU computational work for each block transfer, the communication time over the PCI Express bus will dominate more and influence the performance. This underlines the general importance of reducing the communication, and to perform as much communication-hiding as possible, in heterogeneous computing. With GPU computing we cannot get around the fact that we have to be able to amortize the communication costs over the 'slow' PCI Express bus, if the GPU cannot directly access the system main memory with higher bandwidths and lower latencies than the PCI Express bus can supply.

Methodology of benchmarking

Parameters such as the size of each subdivision of data being transferred - the number of matrix subdivisions, and the whole matrix size, are adjustable. We generate matrices modeling one, three and six degrees of freedom.

In our performance measurements we will look at the performance both with and without communication-hiding enabled. Further, we will vary our parameters:

- Degrees of freedom: one, three or six degrees
- How many subdivisions of the stiffness matrix (directly affecting the size of each subdivision of data being transferred)
- The whole matrix size: number of non-zeroes, as a result of 128^3 (~2 mill.), 256^3 (~16.8 mill.), 160^3 (~4.1 mill.) or 96^3 (~0.9 mill.) nodes and the degrees of freedom (as mentioned in the first point). We have to stay within the bounds of the system memory for the largest cases.

The latter property will affected the minimum amount of subdivisions needed, in order for two subdivisions to concurrently fit into the GPU memory available. Table 1 shows the properties of the stiffness matrices generated, and the space consumed by the non-zero elements alone.

Table 1. Stiffness matrix properties

Degrees of freedom	Number of nodes	Number of non-zeroes	Size of non-zeroes in the stiffness matrix, in GB
1	128^3	55,742,968 x 1	0.42
1	256^3	449,455,096 x 1	3.39
3	128^3	55,742,968 x 9	3.74
3	160^3	109,215,352 x 9	7.32
6	96^3	23,393,656 x 36	6.27
6	128^3	55,742,968 x 36	14.95

Since SpMV is a highly memory bound type of application we can estimate the highest performance theoretically possible to achieve on a certain piece of hardware if we know its memory bandwidth. Of course, we also have to take the algorithm's bytes-per-flop property into

consideration, for this estimation. By counting the number of double precision operations in our routine, and how many reads we have of doubles, we can find the absolute upper limit for the performance that can be achieved with this routine, for a particular hardware device. The calculation is done as follows:

$$\text{FLOP / FLIO x Bandwidth = Performance theoretically achievable} \tag{1}$$

Here, 'FLOP' is number of floating point operations and 'FLIO' is read operations in bytes. For the AMD card we get

$$\text{2 FLOP / 16 byte x 288 GB/s = 36 GFLOP/s} \tag{2}$$

$$\text{18 FLOP / 96 byte x 288 GB/s = 54 GFLOP/s} \tag{3}$$

$$\text{72 FLOP / 336 byte x 288 GB/s = 61.71 GFLOP/s} \tag{4}$$

where (2) is for the one DOF SpMV kernel, (3) is for the three DOF SpMV kernel, and (4) is for the 6 DOF SpMV kernel. These serves as the maximum limits possible to achieve, intra-GPU.

For all benchmarking the GPU in use is 'AMD RADEON HD7970 GHz Edition', of the 'Tahiti' architecture. This is connected on a PCI Express 2.0 bus. All parallel work is configured to use all physical compute units available intra-GPU, and the largest possible number of threads per compute unit (or work-group). In this case, we have 32 'gangs', each of 256 'workers'. The processor is an Intel Nehalem i7-920. For OpenACC the PGI Accelerator C/C++ Workstation compiler release 14.3 was used, with latest AMD GPU drivers (as of April 25th 2014; Catalyst v.14.4). We used AMD CodeXL v.1.3 for GPU profiling.

Results

As a comparison we have measured the serial version, for each of the three degrees of freedom. Number of nodes is set to 128^3 for all cases, and we use 'g++' with the '–O2' optimization flag:
- 1 DOF: 1.08 GFLOP/s
- 3 DOF: 2.67 GFLOP/s
- 6 DOF: 3.20 GFLOP/s

When running on the GPU, all cases inform us that we clearly benefit from the communication-hiding. *Note that the x-axis in all figures below is the amount of subdivisions applied.* In Fig. 3, the size of data is fairly small. We notice a bit varying performance. The small amount of data cannot keep the GPU busy in a sustained manner, so performance is limited.
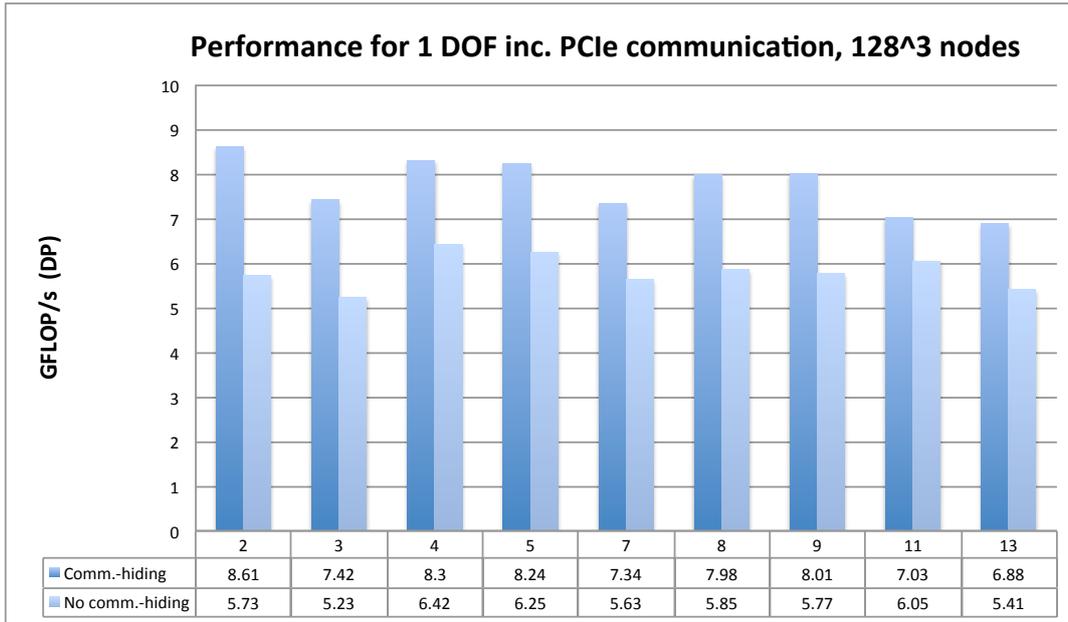
Figure 3. Results for 1 DOF, 128^3 nodes.

In Fig. 4 we have drastically increased the number of nodes, and as a consequence the number of non-zero elements. The GPU is kept busy to a higher degree. The sparse data does not reside blocked (limiting coalesced reads to GPU memory), and the usual high byte-to-flop ratio limits the performance. The in-GPU bandwidth cannot be exploited to a high degree.
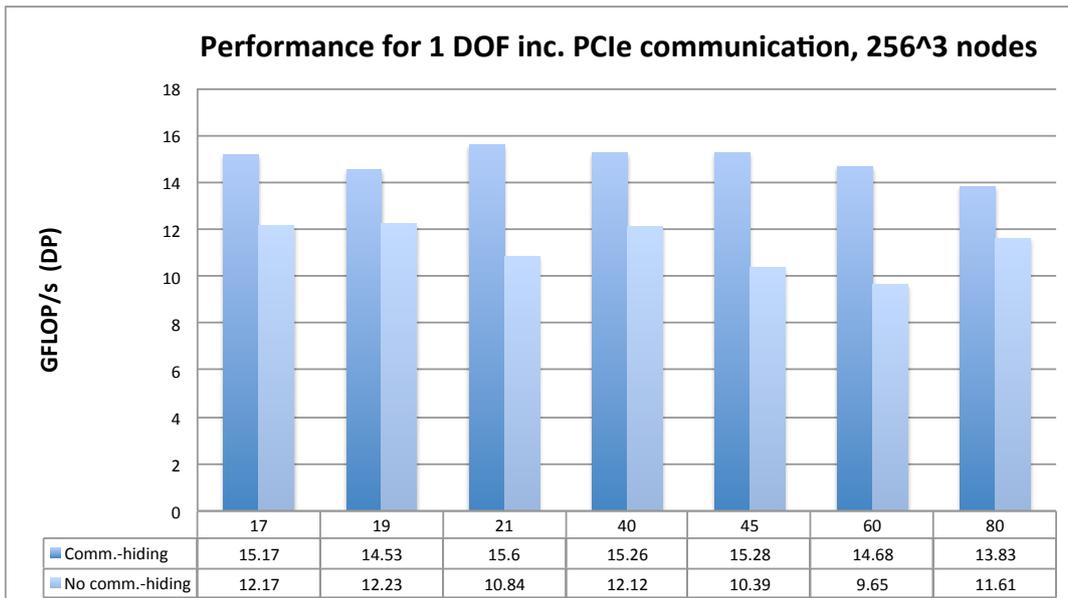
Figure 4. Results for 1 DOF, 256^3 nodes.

In Fig. 5 we move on to a higher degree of freedom. The data in the sparse matrix is here in 3 x 3 blocks. This gives the possibility of coalesced reads on-GPU-card, and thus a better utilization of the bandwidth the card offers. We see much improved performance. Up to about 13 subdivisions the performance is quite stable, then up to 41 subdivisions it gradually drops with several GFLOP/s.
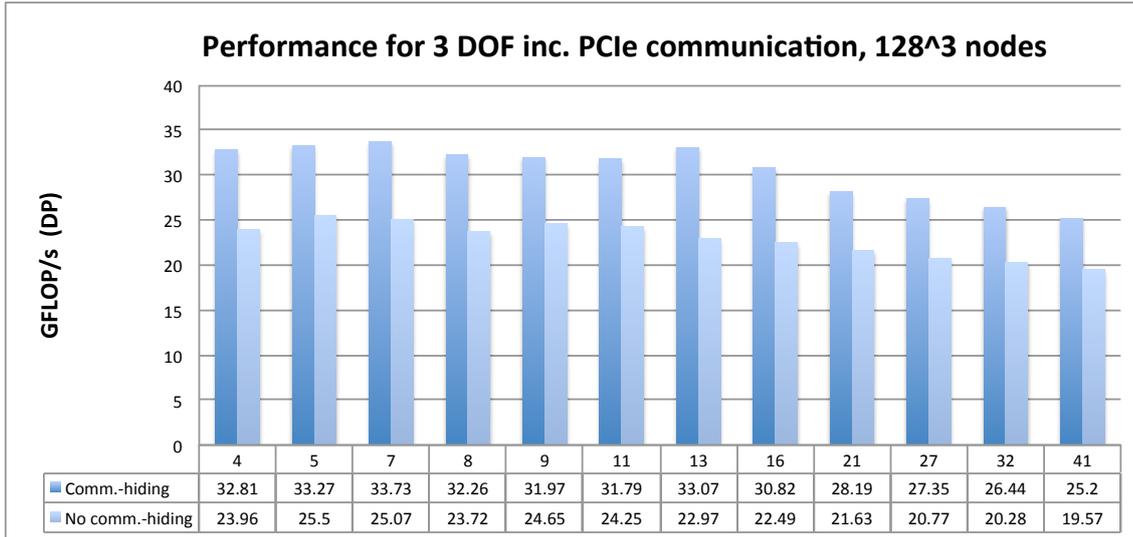
Figure 5. Results for 3 DOF, 128^3 nodes.

In Fig. 6., with more nodes, each subdivision number will give larger subdivisions. The performance is stable from 12 up to 41 subdivisions. Performance with communication-hiding is significantly higher than that of Fig. 5. Total data-size is about twice, and the data amount will possibly keep the GPU better utilized than in Fig. 5.
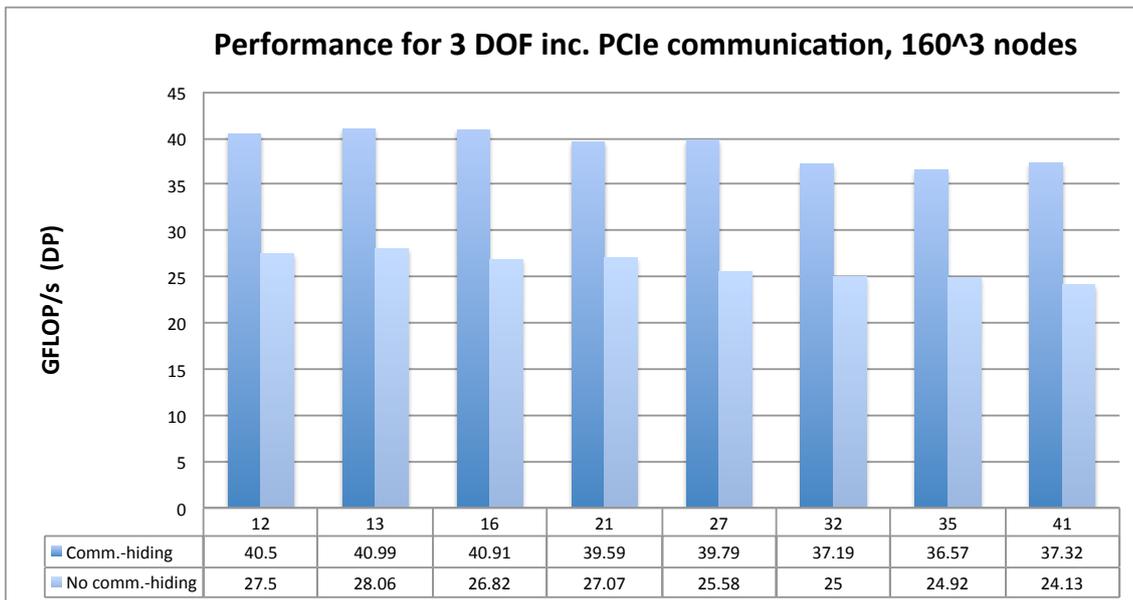
Figure 6. Results for 3 DOF, 160^3 nodes.

In Fig. 7 we have six degrees of freedom. That means even better conditions for coalesced reads of data on-GPU, and this is reflected in a performance jump compared to Fig. 6 and 5. We also observe a performance decrease as number of subdivisions increases.

	7	8	11	16	21	31	36	45
■ Comm.-hiding	43.99	44.22	43.02	42.34	41.39	40.72	37.34	35.22
■ No comm.-hiding	31.47	32.77	31.25	28.88	28.61	25.97	25.6	25.6
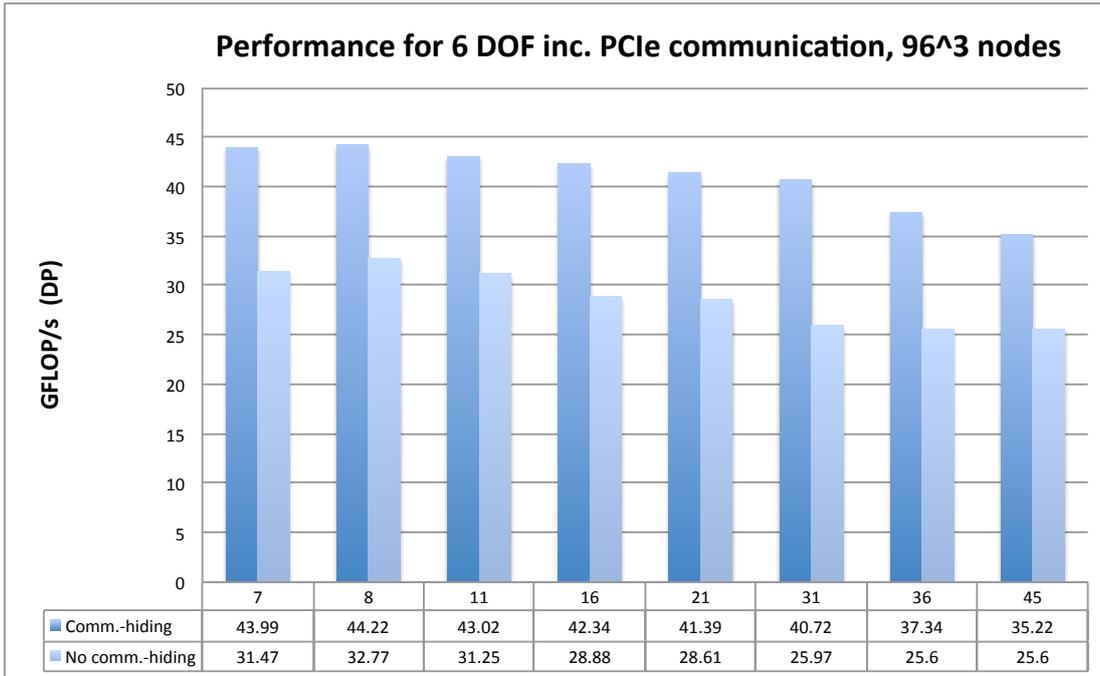
Figure 7. Results for 6 DOF, 96^3 nodes.

In Fig. 8 the combined non-zero size is at the largest, close to 15 GB. Combined with the large amount of coalesced reads, we here observe the highest performance. This applies to both with and without communication-hiding, when "competing" with the other alternatives of parameters and their associated performance. Here, we also observe the most significant gain of performance, when enabeling communication-hiding; up to ~65.7% increased performance for 61 subdivisions.

	21	27	31	36	41	47	51	57	61	71
■ Comm.-hiding	52.88	53.07	52.04	49.87	50.75	50.25	49.51	49.86	49.44	48.99
■ No comm.-hiding	35.37	34.35	34.62	33.63	32.79	32.49	32.26	31.77	29.87	31.49
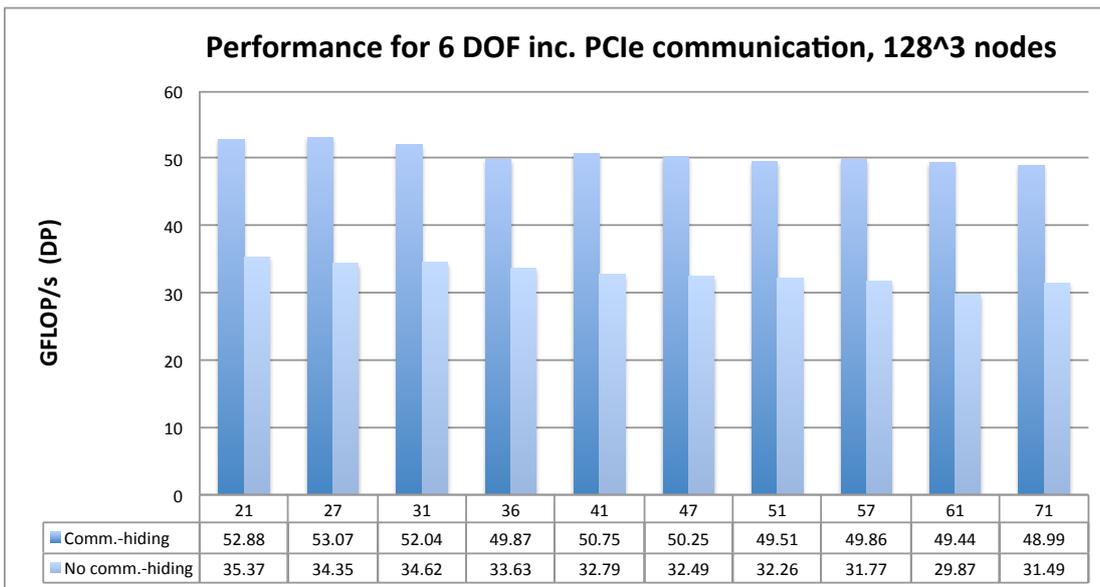
Figure 8. Results for 6 DOF, 128^3 nodes.

When using the profiler, we can get better insight into what is actually happening on the GPU. Particularly, how the asynchronous and synchronous data transfers behave, and differ. It gives us a "window" to see how the communicatin-hiding actually works. In the profiler, the timeline stretches horizontally. In both Fig. 9 and 10 the blue blocks illustrates the transfers of the subdivisions, while green blocks illustrates GPU computation, i.e. the SpMV being executed, in our case.

Figure 9. Profiling when communication-hiding is disabled, all memory transfers are synchronous.

One can easily see how the data-transfer command is blocking, or synchronous, in Fig. 9. No other GPU related work can happen while the synchronous call is being executed. In Fig. 10 we clearly see the effect of using asynchronous, non-blocking, calls. While data is in-flight to the GPU memory, the SpMV routine is executed on a set of data already in-memory. This contributes to the improved performance, for the whole parameter-space we set up.

Figure 10. Profiling when communication-hiding is enabled. All memory transfers except the initial one is asynchronous.

Conclusions

When we consider the peak performance possible intra-GPU for these data-intensive algorithms and with the hardware used, we come impressively close when applying our communication-hiding. The closest is achieved with 6 degrees of freedom and 128^3 number of nodes; 53.07 GFLOP/s is achieved - whereas the peak possible intra-GPU is 61.71 GFLOP/s. On the opposite side, we are the farthest away from the peak possible when having one degree of freedom and 128^3 nodes.

We have seen how SpMV can achieve improved performance by communication-hiding – or, overlapping of computation and data-transfers. This is of direct relevance for, for instance, block Krylov solvers, and makes streaming over the PCI express bus more efficient. This is useful especially when the amount of non-zeroes is so large that all data cannot fit the GPU memory. The communication-hiding scheme becomes more efficient when the amount of subdivisions is high enough so that the first synchronous data-transfer performed in order to buffer does not dominate the total time used for computation and communication. Also, there is a "sweet-spot" where the computation done over each subdivision takes as similar amount of time as possible as the current subdivision being transferred takes to transfer. This controls the rate of efficiency.

It was observed a performance improvement over the whole parameter-space tested. The performance improvement increased as the total amount of non-zero data increased and the possibility of intra-GPU coalesced reads increased. This means that the more efficiently the GPU memory sub-system is utilized, the better effect is observed from the communication-hiding. At the very best, we found an increase in performance of ~65.7%. This underlines the importance of applying such techniques in simulations, when it is suitable with the algorithmic structure of the problem in relation to the underlying computer architecture – as examples here in relation to our SpMV; block Krylov and attached discrete GPUs with a separate memory hierarchy.

Future work

A thorough study of the performance portability is of importance. From the start, the tools selected and the use of them were done so to maintain a high possibility of performance portability. One could, of course, want to run the routines on a future GPU architecture, from the same or a different vendor. Doing such should have as little burden on the developer as possible.

Automatically finding the optimal parameters, for a given problem, is of interest. That is to say, the number of subdivisions that gives the best communication-hiding (matrix size and degrees of freedom already determined by the problem).

As part of a MPI multi-node solver, communication-hiding can be applied at several levels, or layers. Multi-layer communication-hiding is of interest for our research.

References

Sato, Y. and Okuda, H. (2008) Grid-Aided Stochastic Finite Element Using Ninf-G, *8th World Congress on Computational Mechanics (WCCM8)*.

Bell, N. and Garland, M. (2008) Efficient sparse matrix-vector multiplication in CUDA. Vol. 2. No. 5. *NVIDIA Technical Report NVR-2008-004*, NVIDIA Corporation.

Saad, Y. (2003) Iterative methods for sparse linear systems. Siam.

Vuduc, R. W. (2003) Automatic performance tuning of sparse matrix kernels, *Dissertation*, University of California.

Numerical Simulation of Nonlinear Ultrasonic Wave Generation

by an Interface Crack of Bi-material

*T. Maruyama[1], T. Saitoh[2], and S. Hirose[1]

[1]Department of Mechanical & Environmental Informatics, Tokyo Institute of Technology
Ookayama 2-12-1, Meguro-ku, Tokyo 152-8552, Japan
[2]Division of Environmental Engineering Science, Gunma University Tenjin 1-5-1, Kiryu, Gunma 376-8515, Japan

*Corresponding author: maruyama.t.ag@m.titech.ac.jp

Abstract

Nonlinear ultrasonic nondestructive testing using contact acoustic nonlinearity has been developed over the last decade. However, although nonlinear waves such as higher- and sub-harmonics are considered to be generated by the interaction of the crack faces such as clapping motion or friction, the mechanism of generation has not been understood clearly from theoretical view point yet. 1-D and 2-D numerical simulations have been conducted, and 3-D axisymmetric problems have been numerically solved so far. However, no full 3-D analysis has been done. Therefore, in this research, the boundary integral equation for an interface crack with nonlinear boundary conditions in 3-D medium is formulated, and solved numerically using a time-domain boundary element method. The Fourier spectra of received waves are evaluated in the form of far-field scattered waves because the received points are usually located far from the defects in NDT.

Keywords: Time-domain BEM, 3-D nonlinear ultrasonic simulation, Nondestructive testing (NDT), Higher-harmonics, Sub-harmonics, Contact acoustic nonlinearity

Introduction

The nonlinear nondestructive testing (NNDT) using contact acoustic nonlinearity (CAN) is considered as one of the effective methods for the evaluation of closed cracks in metal or on interface of bi-material. Thus, some NNDT methods using CAN have been developed in order to detect cracks and measure the length of closed part of a crack [Ohara at el. (2011)]. The generation of nonlinear ultrasonic waves by the CAN was advocated over thirty years ago [Buck at el. (1978)]. At this stage, higher-harmonics were considered to be generated by the interaction of the crack faces such as clapping motions or friction due to large amplitudes of incident waves [Solodov at el. (2011)]. However, the generation mechanism of sub-harmonics has not been understood clearly yet. Therefore, it is needed to investigate the mechanism more from the theoretical or numerical point of view.

In previous researches, 2-D numerical simulations were carried out [Hirose (1994); Saitoh at el. (2011)]. 3-D axisymmetric problem of a penny-shaped crack subjected to normal incidence of a longitudinal wave was solved numerically [Hirose at el. (1993)]. However, no full 3-D analysis has been done. Therefore, in this research, the 3-D boundary integral equation (BIE) is formulated for an interface crack of bi-material with nonlinear boundary conditions and numerically solved using the time-domain boundary element method (BEM) in order to investigate the relation between the analysis conditions, such as frequencies of an incident wave and size of a crack, and the generation of higher- and sub-harmonics.

In the proposed numerical method, the implicit Runge-Kutta (IRK) based convolution quadrature method (CQM) [Maruyama at el. (2013)] is applied to discretization of convolution integrals in BIE. Application of CQM to the discretization improves numerical accuracy and stability behavior of the time-marching process of time-domain BEM. In addition, far-field scattered waves are evaluated as received waves and used for the Fourier spectrum analysis because the received points are usually located far from the defects compared with the defect size and wave length in NDT.
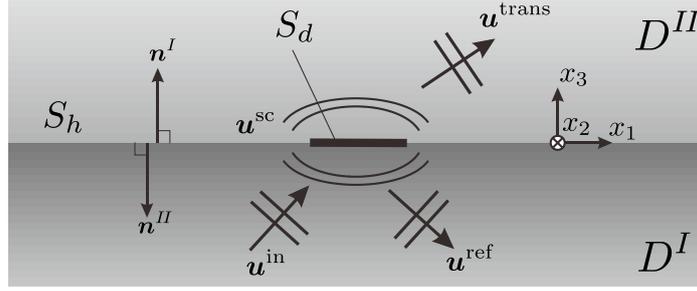
Figure 1. Debonding area of bi-material interface subjected to an incident plane wave.

Formulation of boundary integral equations

A 3-D boundary element analysis model for nonlinear ultrasonic simulation is considered for two-layer problems including the debonding area as shown in Fig. 1. This model consists of two semi-infinite domains, D^I and D^{II}, and the interfaces between them, S_h and S_d, denote bonding and debonding areas, respectively. In addition, $\boldsymbol{n}^I$ and $\boldsymbol{n}^{II}$ are defined as unit vectors pointing into outer normal directions from respective domains where the upper subscripts, I and II, indicate the respective domains. In this analysis, a plane wave is used as the incident wave to investigate fundamental motions of the nonlinear crack.

For the layered problems subjected to an incident plane wave, the free field formulation is usually used in BEM. Assuming that the interface S_h is flat, the free field $\boldsymbol{u}^{\text{free}}$ which consists of incident wave $\boldsymbol{u}^{\text{in}}$, reflected wave $\boldsymbol{u}^{\text{ref}}$, and transmitted wave $\boldsymbol{u}^{\text{trans}}$ can be calculated analytically. The scattered wave $\boldsymbol{u}^{\text{sc}}$ is defined as the disturbance of $\boldsymbol{u}^{\text{free}}$ by the debonding area S_d, and following equations are obtained:

$$\boldsymbol{u}^{\text{free};I} = \boldsymbol{u}^{\text{in};I} + \boldsymbol{u}^{\text{ref};I}, \qquad \boldsymbol{u}^{\text{free};II} = \boldsymbol{u}^{\text{trans};II}, \qquad \boldsymbol{u} = \boldsymbol{u}^{\text{free}} + \boldsymbol{u}^{\text{sc}}, \tag{1}$$

where $\boldsymbol{u}$ is the total wave. Since $\boldsymbol{u}^{\text{sc}}$ satisfies the radiation condition, the BIE is formulated for $\boldsymbol{u}^{\text{sc}}$ as follows:

$$
\begin{aligned}
\frac{1}{2}\boldsymbol{u}^{\text{sc};I(II)}(\boldsymbol{x},t) = &\int_0^t \int_{S_h+S_d} \boldsymbol{U}^{I(II)}(\boldsymbol{x},\boldsymbol{y},t-\tau)\boldsymbol{t}^{\text{sc};I(II)}(\boldsymbol{y},\tau)dS_y\, d\tau \\
&- \int_0^t \text{p.v.} \int_{S_h+S_d} \boldsymbol{T}^{I(II)}(\boldsymbol{x},\boldsymbol{y},t-\tau)\boldsymbol{u}^{\text{sc};I(II)}(\boldsymbol{y},\tau)dS_y\, d\tau,
\end{aligned}
\tag{2}
$$

where $\boldsymbol{t}$ is the traction force, and $\boldsymbol{U}$ and $\boldsymbol{T}$ are the fundamental solutions for displacement and traction, respectively in 3-D elastodynamics. The symbol p.v. indicates the Cauchy's principle integral. Substituting Eq. (1c) into Eq. (2), the BIE is expressed by $\boldsymbol{u}$ and $\boldsymbol{u}^{\text{free}}$ and can be numerically solved using discretization methods for time and space and appropriate interface conditions on S_h and S_d. In addition, S_h is truncated by finite area in numerical analysis.

Discretization of BIE using IRK based CQM

In solving the BIE (2) numerically, the convolution integrals are evaluated by means of the IRK based CQM [Lubich et al. 1993] and the surface integrals over S_h and S_d are discretized by the piecewise constant boundary elements. If the m-stage Radau IIA method, which is one of the IRK methods, is used in the IRK based CQM, and the interface including the debonding area is divided into M boundary elements, the discretized BIE at the n-step and the i-sub-step in time are shown as follows:

$$\frac{1}{2}\boldsymbol{u}_\gamma^{I(II)}\big((n+c_i)\Delta t\big) = \frac{1}{2}\boldsymbol{u}_\gamma^{\text{free};I(II)}\big((n+c_i)\Delta t\big)$$

$$+\sum_{k=0}^{n}\sum_{\alpha=1}^{M}\sum_{j=1}^{m}\Big[\boldsymbol{A}_{\gamma\alpha}^{ij;n-k}\big\{\boldsymbol{t}_\alpha^{I(II)}\big((k+c_j)\Delta t\big) - \boldsymbol{t}_\alpha^{\text{free};\,I(II)}\big((k+c_j)\Delta t\big)\big\} \tag{3}$$

$$-\boldsymbol{B}_{\gamma\alpha}^{ij;n-k}\big\{\boldsymbol{u}_\alpha^{I(II)}\big((k+c_j)\Delta t\big) - \boldsymbol{u}_\alpha^{\text{free};\,I(II)}\big((k+c_j)\Delta t\big)\big\}\Big],$$

where subscripts, α and γ, are the indexes of boundary elements and c_i is the parameter in IRK method corresponding to the sub-step. In addition, $\boldsymbol{A}_{\gamma\alpha}^{ij;\kappa}$ and $\boldsymbol{B}_{\gamma\alpha}^{ij;\kappa}$ are influence functions expressed as follows:

$$\boldsymbol{A}_{\gamma\alpha}^{ij;\kappa} = \frac{\mathcal{R}^{-\kappa}}{L}\sum_{l=0}^{L-1}\left[\sum_{\beta=1}^{m}\{\boldsymbol{E}_\beta(z_l)\}_{ij}\int_{S_\alpha}\widehat{\boldsymbol{U}}^{I(II)}(\boldsymbol{x}_\gamma,\boldsymbol{y},\lambda_\beta^l)dS_y\right]e^{-\frac{2\pi i\kappa l}{L}}, \tag{4}$$

$$\boldsymbol{B}_{\gamma\alpha}^{ij;\kappa} = \frac{\mathcal{R}^{-\kappa}}{L}\sum_{l=0}^{L-1}\left[\sum_{\beta=1}^{m}\{\boldsymbol{E}_\beta(z_l)\}_{ij}\ \text{p. v.}\int_{S_\alpha}\widehat{\boldsymbol{T}}^{I(II)}(\boldsymbol{x}_\gamma,\boldsymbol{y},\lambda_\beta^l)dS_y\right]e^{-\frac{2\pi i\kappa l}{L}}, \tag{5}$$

where (^) indicates the function in the Laplace-domain, i is the imaginary unit, and the last arguments λ_β^l of $\widehat{\boldsymbol{U}}$ and $\widehat{\boldsymbol{T}}$ correspond to Laplace parameters. In Eqs. (4) and (5), $\lambda_\beta^l, \mathcal{R}, L, z_l$, and $\boldsymbol{E}_\beta$ are the parameters of IRK based CQM [Maruyama et al. (2013)]. The matrix-vector products on the right side of Eq. (3) are effectively calculated by means of the fast multipole method (FMM), which is one of the acceleration methods for BEM.

Nonlinear interface conditions

The interface condition on the bonding area S_h is the continuity of displacement and traction as

$$\boldsymbol{u}^I = \boldsymbol{u}^{II}, \qquad \boldsymbol{t}^I = -\boldsymbol{t}^{II}. \tag{6}$$

For the debonding area S_d, three types of interface conditions, "separation", "stick", and "slip", are considered [Hirose (1994); Saitoh at el. (2011)]. "separation" means that two surfaces of upper and lower materials are separated with no traction, while "stick" and "slip" are contact conditions under compressive normal stress state. For the "stick" condition, the surfaces of two materials move with no relative velocity. On the other hand, the "slip" condition allows a relative tangential movement with dynamic friction force. Therefore, these three conditions are described as follows:

$$\boldsymbol{t}^I = \boldsymbol{t}^{II} = \boldsymbol{0} \qquad\qquad\qquad\qquad\qquad\qquad : \text{separation,} \tag{7}$$

$$[u_3] = 0, \qquad \boldsymbol{t}^I = -\boldsymbol{t}^{II}, \qquad [\dot{\boldsymbol{u}}_t] = \boldsymbol{0} \qquad\qquad : \text{stick,} \tag{8}$$

$$[u_3] = 0, \qquad t_3^I = -t_3^{II}, \qquad \boldsymbol{t}_t^I = -\boldsymbol{t}_t^{II} = \frac{[\dot{\boldsymbol{u}}_t]}{|[\dot{\boldsymbol{u}}_t]|}\mu_d(-t_3^I) \qquad : \text{slip,} \tag{9}$$

where $[\boldsymbol{u}]$ is the crack opening displacement and expressed by $[\boldsymbol{u}] = \boldsymbol{u}^{II} - \boldsymbol{u}^I$, (˙) indicates the time differentiation, and the subscript t means tangential components in x_1 and x_2 directions. In addition, μ_d is the dynamic friction coefficient.

Numerical procedure

The numerical algorithm is shown in Fig. 2. At the beginning of a time step in the IRK based CQM, the discretized BIE (3) is solved assuming that the interface conditions on each element are the same as those in the previous time step. If the additional constraint conditions enclosed by the double
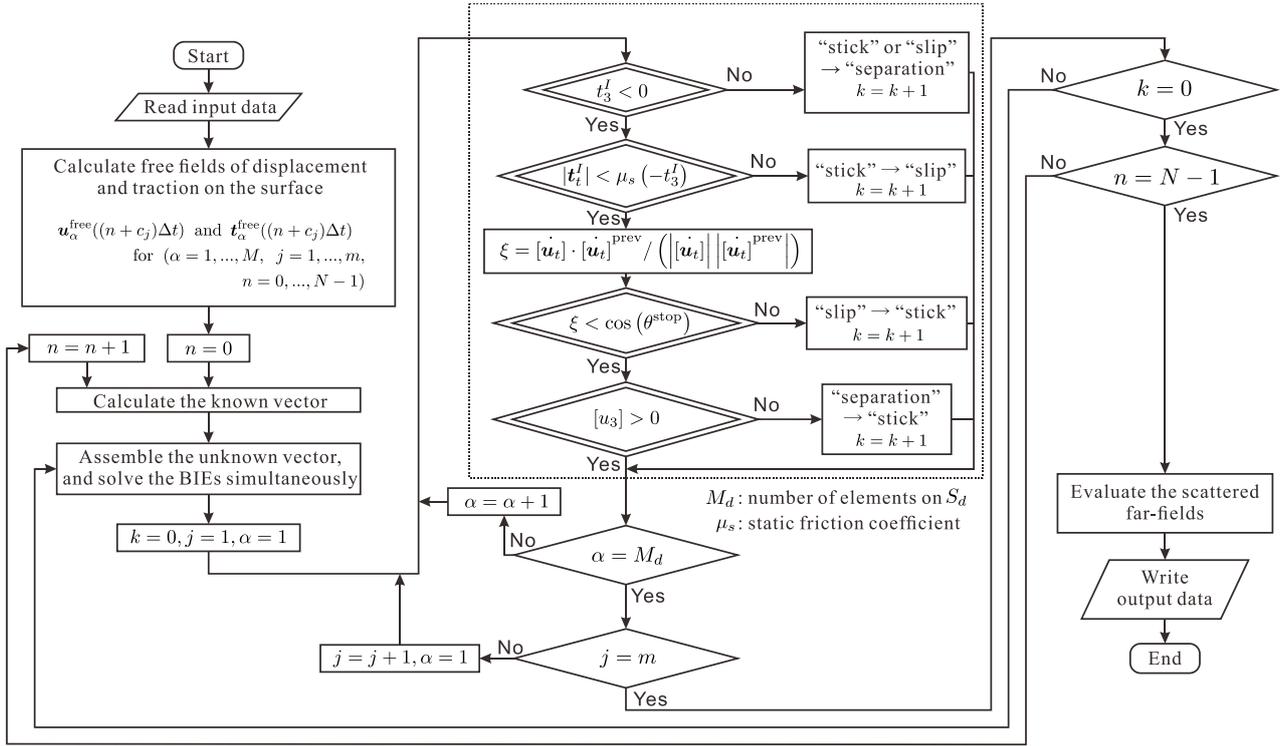
Figure 2. Numerical algorithm.

rhombuses in Fig. 2 are not satisfied, the interface condition on the element, which is one of "separation", "stick", and "slip", is changed into one of the other conditions and then the system of equations is assembled and solved again. After conducting the iterative calculations, if both the interface conditions and the additional constraint conditions on all elements at all sub-steps are satisfied, the time step proceeds to the next one.

Some remarks concerning the numerical calculations are given below. At the initial time step, the interface condition of "stick" is given on all elements on the debonding area assuming that the interface is closed before the wave incidence. There are two possible phase shifts from "separation" to one of two contact conditions, i.e., "slip" and "stick". In the present study, the priority is given to the change from "separation" to "stick", if $[u_3] > 0$ for the "separation" is violated on the element. In numerical calculations, it is difficult to achieve the condition $[\dot{u}_t] = \mathbf{0}$ exactly in the transition from "stick" to "slip". Therefore, we set $[\dot{u}_t] = \mathbf{0}$ unless the following condition is satisfied:

$$\xi < \cos(\theta^{\text{stop}}), \qquad \xi = [\dot{u}_t] \cdot [\dot{u}_t]^{\text{prev}} / (|[\dot{u}_t]||[\dot{u}_t]^{\text{prev}}|), \tag{10}$$

where $[u]^{\text{prev}}$ is the crack opening displacement at the previous time step. Eq. (10) means that the transition from "slip" to "stick" occurs when there is a big change in the slip direction. In this study, θ^{stop} is given by 90 degrees.

Far-field scattered wave

In this research, the scattered wave by an interface crack at far-field [Hirose at el. (1989)] is calculated to investigate the generation of nonlinear ultrasonic waves. For example, when x is the receiver point and y is the point on a crack, the far-field scattered L wave by an interface crack of bi-material $u_L^{\text{sc,far};I(II)}$ is given by
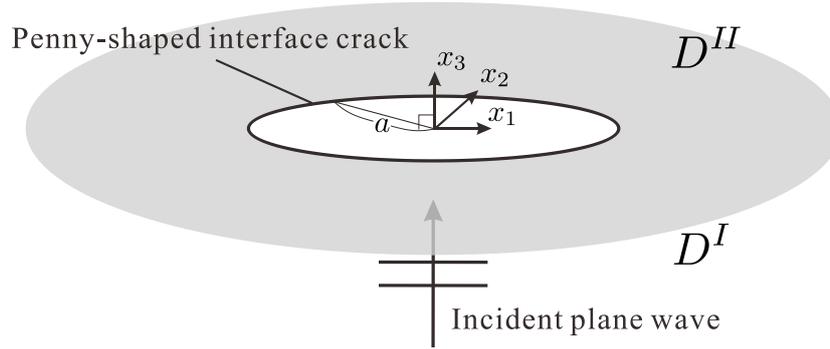
Figure 3. Scattering of an incident plane wave by a penny-shaped interface crack of
bi-material.

$$u_L^{\text{sc,far};I(II)}(\pmb{x},t) \simeq \frac{1}{4\pi x}\Omega_L\left(\hat{\pmb{x}},t-\frac{x}{c_L^{I(II)}}\right), \tag{11}$$

where $x = |\pmb{x}|$ and $\hat{\pmb{x}} = \pmb{x}/x$. Ω_L is the far-field amplitude of L wave, which is expressed for $y_3 = 0$ as follows:

$$\Omega_L\left(\hat{\pmb{x}},t-\frac{x}{c_L^{I(II)}}\right) = \sum_{\alpha=L,TV}\frac{C_{pjkq}^{II(I)}}{\mu^{II(I)}}A_k^{\pm\alpha}\frac{\zeta_q^{\pm}}{c_L^{I(II)}}T^{\alpha,L}(\pmb{\zeta}^{\pm})\frac{|\hat{x}_3|}{\nu}$$

$$\times\int_{S_d}n_j^I(\pmb{y})[\dot{u}_p]\left(\pmb{y},t-\left[\frac{x}{c_L^{I(II)}}-\frac{\pmb{\zeta}^{\pm}\cdot\pmb{y}}{c_\alpha^{II(I)}}\right]\right)dS_y\,, \tag{12}$$

where C_{ijkl} is the elastic constant tensor and μ is the shear modulus. $T^{\alpha,L}(\pmb{\zeta}^{\pm})$ is the transmission coefficient with incident wave propagation vector $\pmb{\zeta}^{\pm}$ into $x_3 = 0$ plane when the wave mode is changed from α to L and the wave is propagating from $D^{II(I)}$ to $D^{I(II)}$. In addition, $\pmb{\zeta}^{\pm}$, ν, and $A_k^{\pm\alpha}$ are given by

$$\pmb{\zeta}^{\pm} = \left(\frac{c_\alpha^{II(I)}}{c_L^{I(II)}}\,\hat{x}_1, \frac{c_\alpha^{II(I)}}{c_L^{I(II)}}\,\hat{x}_2, \pm\nu\right),\qquad \nu = \sqrt{1-\left(\frac{c_\alpha^{II(I)}}{c_L^{I(II)}}\,\hat{x}_1\right)^2-\left(\frac{c_\alpha^{II(I)}}{c_L^{I(II)}}\,\hat{x}_2\right)^2}, \tag{13}$$

$$A_k^{\pm L} = \left(\frac{c_T^{II(I)}}{c_L^{II(I)}}\right)^2\zeta_k^{\pm},\qquad \pmb{A}^{\pm TV} = \hat{\pmb{d}}^{TH}\times\pmb{\zeta}^{\pm}. \tag{14}$$

In Eq. (14), $\hat{\pmb{d}}^{TH}$ is the displacement vector of the TH wave propagating to $\pmb{\zeta}^{\pm}$ direction. In Eqs. (12)-(14), $\pm$ is decided by the positional relation between $\pmb{x}$ and $\pmb{y}$ in derivation of the Green's function, and $+$ and $-$ correspond to $x_3 > y_3$ and $x_3 < y_3$, respectively [Achenbach at el. (1982)]. In addition, the TV and TH wave components of the far-field scattered waves are described by analogous formulas. In this study, Ω_L is used for the Fourier spectrum analysis, because the far-field scattered waves do not include the truncated error of S_h if only $[\dot{\pmb{u}}]$ is calculated accurately.

Numerical examples

Scattering of an incident plane wave by a penny-shaped nonlinear interface crack with radius a, as shown in Fig. 3, is analyzed by the proposed method. The material constants are shown in Table 1

Table 1. Material constants

	c_L [m/s]	c_T [m/s]	ρ [kg/m^3]
D^I (steel)	5800	3100	7690
D^{II} (aluminum)	6400	3040	2700
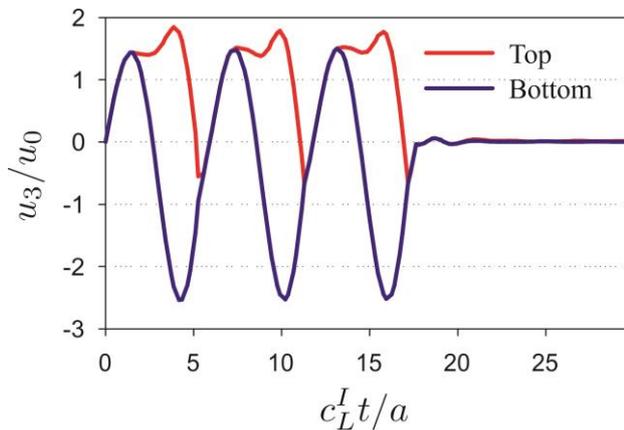
Figure 4. Vertical displacements at the center points on top and bottom surfaces of an interface crack of bi-material as a function of time.
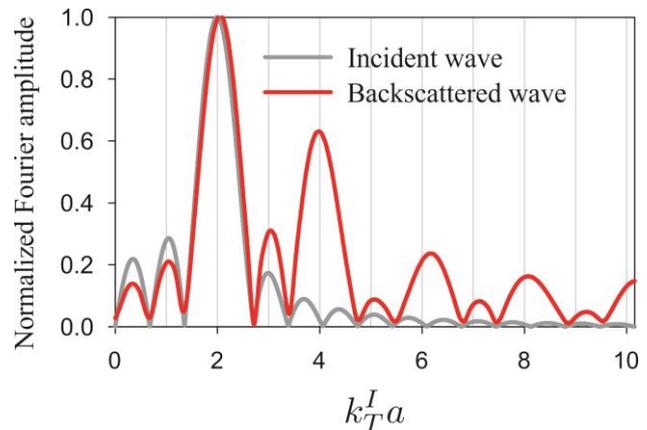
Figure 5. Normalized Fourier amplitudes of backscattered L wave at far-field and incident wave as a function of wave number.

and the static and dynamic friction coefficients, μ_s and μ_d, are given by $\mu_s = 0.61$ and $\mu_d = 0.47$, respectively. The incident plane wave is given by a three cycle sinusoidal wave with amplitude u_0.

Fig. 4 shows the vertical displacements at the center points on top and bottom surfaces of the interface crack subjected to the normal incident L wave with the normalized wave number $k_T^I a = 2\pi a f / c_T^I = 2.0$ where f is the center frequency of the incident wave. In Fig. 4, the clapping motion occurs at the crack face. The crack opening displacement rapidly decreases and then vanishes when the crack is completely closed. Fourier spectra of the backscattered far-field amplitude $\Omega_L/(au_0)$ and the incident wave are shown in Fig. 5. These spectra are normalized by their maximum values. It is observed that large higher-harmonics components are included in the backscattered wave.

Conclusions

In this paper, the boundary integral formulation, interface conditions, and numerical algorithm for the simulation of an interface crack of bi-material subjected to an incident plane wave are presented. Moreover, the calculation method of far-field amplitude for the two-layer problem, and numerical results of normal incidence of L wave are shown. From the numerical results, the generation of higher-harmonics by CAN was confirmed using the proposed method. Additional numerical examples, such as oblique incidence, will be shown in near future.

References

Achenbach, J. D., Gautesen, A. K. and McMaken, H. (1982) Ray methods for waves in elastic solids, *Pitman*.
Buck, O., Morris, W. L. and Richardson, J.M. (1978) Acoustic harmonic generation at unbonded interfaces and fatigue cracks, *Applied Physics Letters* **33**, 371–373.

Hirose, S. and Achenbach, J. D. (1989) Time-domain boundary element analysis of elastic wave interaction with a crack, *International Journal for Numerical Methods in Engineering*, **28**, 628–644.

Hirose, S. and Achenbach, J. D. (1993) Higher harmonics in the far-field due to dynamic crack-face contacting, *Journal of the Acoustical Society of America*, **93**, 142–147.

Hirose, S. (1994) 2-D scattering by a crack with contact boundary conditions, *Wave Motion*, **19**, 37–49.

Lubich, C. and Ostermann A. (1993) Runge-Kutta methods for parabolic equations and convolution quadrature, *Mathematics of Computation*, **60**, 105-131.

Maruyama, T., Saitoh, T. and Hirose, S. (2013) Implicit Runge-Kutta based convolution quadrature time-domain fast multipole boundary element method for 3-D scalar wave problems, *Journal of JSCE Applied Mechanics*, **69**, I_175–I_185 (in Japanese).

Ohara, Y., Horinouchi, S., Hashimoto, M., Shintaku, Y. and Yamanaka, K. (2011) Nonlinear ultrasonic imaging method for closed cracks using subtraction of responses at different external loads, *Ultrasonics*, **51**, 661-666.

Saitoh, T., Furuta, Y., Hirose, S. and Nakahata, K. (2011) Simulation of higher harmonics on nonlinear ultrasonic testing using convolution quadrature time-domain boundary element method in 2-D elastodynamics, *Journal of JSCE Applied Mechanics*, **67**, I_161–I_169 (in Japanese).

Solodov, I., Doring, D. and Busse, G. (2011) New opportunities for NDT using non-linear interaction of elastic waves with defects, *Journal of Engineering Mechanics*, **57**, 169–182.

Non-linear stability analysis of a Darcy flow with viscous dissipation

*M. Celli[1], L. De B. Alves[2], A. Barletta[1]

[1]Department of Industrial Engineering, Alma Mater Studiorum Università di Bologna,
Viale Risorgimento 2, 40136 Bologna, Italy

[2]Departamento de Engenharia Mecânica, Universidade Federal Fluminense,
Niterói, Rio de Janeiro, 24210-240, Brasil

*Corresponding author: michele.celli3@unibo.it

Abstract

The nonlinear convective instability of a flow in a fluid saturated impermeable and rectangular porous channel of arbitrary aspect ratio is here investigated by taking into account the effect of viscous dissipation. Darcy's law and Oberbeck-Boussinesq approximation are assumed. The vertical boundaries are assumed to be adiabatic and the horizontal boundaries are taken to be isothermal with the cold face placed on top. The system is characterised by two sources of thermal instability: the buoyancy activated by the non trivial temperature distribution due to the internal heat generation by the viscous dissipation and the buoyancy triggered by the non linear temperature distribution due to the temperature gap between the horizontal boundaries. The novel feature introduced in the present paper is the fully nonlinear approach to the stability analysis. The results obtained by the linear stability analysis are here used as a reference. The purpose of this paper is to analyse the system with the aim of finding possible subcritical instabilities. The technique employed in order to investigate the nonlinear problem is the generalized integral transform technique. The computational task relative to the integral transformation procedure and the solution of the ordinary differential equations obtained are carried out by *Mathematica 9* (© Wolfram Research).

Keywords: Nonlinear Stability, Generalised Integral Transform Technique, Porous Media, Viscous Dissipation, Thermal Convection

Introduction

The study of the onset of the thermal instabilities is an important topic with a deep engineering impact. In particular, the stability analyses of fluid saturated porous media have indeed several applications in a widespread range of scientific fields: from the oil extraction engineering to geological and geophysical studies and biological tissues convection heat and mass transfer. The source of thermal convection in fluid saturated porous media consists, typically, in the buoyancy force built up by heating from below boundary condition or an internal heat generation [Nield and Bejan (2013)]. In this paper, the buoyancy force is identified as the source of thermal instability and an internal generation effect, *viz.* the viscous dissipation, is employed in order to yield the buoyancy force. Analyses of this topic have been published by [Nield (2007)], [Storesletten and Barletta (2009)] and [Barletta, Celli and Rees (2009)]. The novel feature introduced in the present paper is the fully nonlinear approach to the stability analysis. The results obtained by the linear stability analysis are here used as a reference [Nield and Barletta (2010)]. In order to solve the nonlinear problem a specific mathematical

technique is here employed: the generalized integral transform technique (GITT) [Cotta (1998)] and [Pontes, Alves and Cotta (2002)]. The GITT is an alternative hybrid numerical analytical technique based on the eigenfunction expansion in the spatial variables of the problem fields, *i.e.* velocity and temperature. The ordinary differential equations obtained by the integral transformation procedure constitute an initial value problem that is solved numerically. The computational task relative to the integral transformation procedure and the solution of the initial value problem are carried on by software that allow for mixed symbolic and numerical computations such as *Mathematica 9* (©) Wolfram Research).

Mathematical model

A rectangular porous channel saturated by fluid with arbitrary aspect ratio is here investigated. A throughflow of give rate is assumed. The channel is considered impermeable while, for what concerns the thermal boundary conditions, the vertical channel walls are assumed to be adiabatic and the horizontal channel walls are assumed to be isothermal. A temperature gap, $\Delta T = T_h - T_c$, is imposed between the horizontal boundaries. The cold face, T_c, is placed on top and the hot face, T_h, is placed on the lower boundary. The Oberbeck-Boussinesq approximation is assumed, Darcy's law is employed in order to define the momentum balance equation and the viscous dissipation contribution inside the energy balance equation is taken into account as internal heat source. The curl operator is applied to Darcy's law so that the governing equations lose the pressure gradient contribution. The dimensionless governing equations that describe the system together with the relative boundary conditions are

$$
\begin{aligned}
\boldsymbol{\nabla} \cdot \boldsymbol{u} &= 0, \\
\boldsymbol{\nabla} \times \boldsymbol{u} &= \boldsymbol{\nabla} \times (T\boldsymbol{e}_y), \\
\frac{\partial T}{\partial t} + \boldsymbol{u} \cdot \boldsymbol{\nabla} T &= \nabla^2 T + Ge\,\boldsymbol{u} \cdot \boldsymbol{u}, \\
x = 0, s: \quad u &= 0, \quad \frac{\partial T}{\partial x} = 0, \\
y = 0, 1: \quad v &= 0, \quad T = R, 0.
\end{aligned}
\tag{1}
$$

where $\boldsymbol{u} = (u, v, w)$ is the velocity vector, T is the temperature, $\boldsymbol{e}_y$ is the unit vector of the y-axis, t is the time, Ge the Gebhart number, R the Rayleigh number and s is the aspect ratio. A sketch of the geometry and a description of the boundary conditions is reported in Fig. 1. The scaling employed in order to obtain the dimensionless formulation is the following

$$
\begin{aligned}
\bar{t} &= \frac{\sigma H^2}{\alpha} t, \quad \bar{x} = Hx, \quad \bar{u} = \frac{\alpha}{H} u, \quad \overline{T} = T_c + \Delta T \frac{T}{R}, \\
Ge &= \frac{g \beta H}{c}, \quad R = \frac{g \beta \Delta T H K}{\nu \alpha}, \quad s = \frac{L}{H},
\end{aligned}
\tag{2}
$$

where the dimensional quantities are over-lined, σ is the dimensionless ratio between the average heat capacity per unit volume of the porous medium and the average heat capacity per unit volume of the fluid, H is the height of the channel, α is the effective thermal diffusivity, ν is the kinematic viscosity, K is the permeability, c is the specific heat, g is the modulus of gravity acceleration, β is the thermal expansion coefficient and L is the width of the channel.
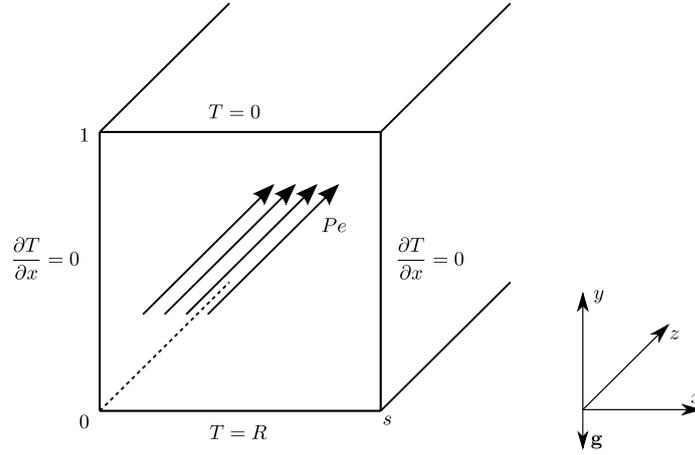
Figure 1. A sketch of the porous channel and its boundary conditions

Basic stationary solution

The first step in this stability analysis consists in redefining the velocity and temperature fields as composed by two contributions: a stationary basic state and a perturbed field, namely

$$\boldsymbol{u} = \boldsymbol{u}_b + \boldsymbol{U}, \quad T = T_b + \Theta, \tag{3}$$

where the subscript b refers to the basic state. In the following we will need also the initial values of velocity and temperature perturbed fields. These initial values are defined as a combination of an initial value of the perturbed field plus the basic stationary flow contribution.

$$t = 0: \qquad \boldsymbol{u} = \boldsymbol{u}_b + \boldsymbol{U}_0, \quad T = T_b + \Theta_0. \tag{4}$$

The stability analysis is here performed with respect to a particular stationary solution of the governing equations (1). A constant throughflow in the z-direction is assumed and the temperature field of this fully developed flow is assumed to be dependent only on the y-coordinate, namely

$$\boldsymbol{u}_b = \{0, 0, Pe\}, \quad T_b = \frac{(1-y)(2R+\Lambda y)}{2}. \tag{5}$$

The Péclet number is defined through the average velocity over the channel, $Pe = w_b L/\alpha$, and the parameter $\Lambda = Ge\,Pe^2$. Inside the setup just described we may identify two possible mechanisms capable to generate thermal instabilities: the coupling between the buoyancy force and the heat generated by viscous dissipation and the coupling between the buoyancy force and the vertical temperature gradient produced by ΔT. Each mechanism is regulated by means of a nondimensional parameter: Λ regulates the strength of the buoyancy force due to the viscous dissipation contribution whereas R regulates the strength of the buoyancy force due to Darcy-Bènard-like mechanism coming from the temperature gap ΔT.

Perturbed equations

On applying the Eq. (3) to Eq. (1) and subtracting the basic stationary state contribution one obtains

$$\nabla \cdot \boldsymbol{U} = 0,$$
$$\nabla \times \boldsymbol{U} = \nabla \times (\Theta \boldsymbol{e}_y),$$
$$\frac{\partial \Theta}{\partial t} + V \frac{\partial T_b}{\partial y} + Pe \frac{\partial \Theta}{\partial z} + \boldsymbol{U} \cdot \nabla \Theta = \nabla^2 \Theta + 2 Ge\, Pe\, W + Ge\, \boldsymbol{U} \cdot \boldsymbol{U}. \tag{6}$$

The investigation may now be reduced in complexity by focusing our attention only on the longitudinal rolls and disregarding the other possible inclinations of the disturbances. Since the longitudinal rolls lie on the (x,y)-plane, the contributions of those term in the equations that refer to the z-direction are thus neglected. On introducing the streamfunction $U = \partial \Psi / \partial y$ and $V = -\partial \Psi / \partial x$, the governing equations (6), the relative boundary conditions and the initial values reduce to

$$\nabla^2 \Psi = -\frac{\partial \Theta}{\partial x}, \tag{7a}$$

$$\frac{\partial \Theta}{\partial t} - \frac{\partial T_b}{\partial y}\frac{\partial \Psi}{\partial x} + \frac{\partial \Psi}{\partial y}\frac{\partial \Theta}{\partial x} - \frac{\partial \Psi}{\partial x}\frac{\partial \Theta}{\partial y} = \frac{\partial^2 \Theta}{\partial x^2} + \frac{\partial^2 \Theta}{\partial y^2} + Ge \left(\frac{\partial \Psi}{\partial y} \right)^2 + Ge \left(\frac{\partial \Psi}{\partial x} \right)^2, \tag{7b}$$

$$
\begin{aligned}
x = 0, s: &\qquad \Psi = 0, &\qquad \frac{\partial \Theta}{\partial x} = 0, \\
y = 0, 1: &\qquad \Psi = 0, &\qquad \Theta = 0, \\
t = 0: &\qquad \Psi = \Psi_0(x,y), &\qquad \Theta = \Theta_0(x,y).
\end{aligned} \tag{7c}
$$

Since Eq. (7a) does not show a time dependency, the initial value for the streamfunction filed is not necessary. On the other hand, the shape of Θ_0 is chosen to be equal to the temperature field configuration of a single longitudinal roll occupying the whole channel, $\Theta_0(x,y) = \cos(\pi x/s)\sin(\pi y)$. One may note that the value of the perturbation amplitude for Θ_0 is of $O(1)$. The choice of this order of magnitude comes from the necessity to distinguish this analysis from the linear stability one. The linear stability analysis requires, indeed, to employ perturbations small enough so that the nonlinear terms in the perturbations inside the governing equations may be neglected.

The Generalised Integral Transform Technique

In order to perform the nonlinear stability analysis the Generalised Integral Transform Technique (GITT) is employed. The GITT starts with the eigenfunction expansion of the problem potential on the spatial variables. For what concerns the streamfunction field, the so-called auxiliary eigenvalue problems in the x and y-directions are

$$\frac{d^2 \bar{\psi}_i(x)}{dx^2} + \lambda_i^2 \bar{\psi}_i(x) = 0, \qquad \bar{\psi}_i(0) = \bar{\psi}_i(s) = 0, \tag{8a}$$

$$\frac{d^2 \tilde{\psi}_j(y)}{dy^2} + \omega_j^2 \tilde{\psi}_j(y) = 0, \qquad \tilde{\psi}_j(0) = \tilde{\psi}_j(1) = 0. \tag{8b}$$

The relative eigenfunctions and eigenvalues are

$$\bar{\psi}_i(x) = \sqrt{\frac{2}{s}} \sin(\lambda_i x), \quad \lambda_i = \frac{i\pi}{s}, \quad i = 1, 2, \dots \tag{9a}$$

$$\tilde{\psi}_j(y) = \sqrt{2}\sin(\omega_j y), \quad \omega_j = j\pi, \quad j = 1,2,.... \tag{9b}$$

For what concerns the temperature field, the so-called auxiliary eigenvalue problems in the x and y-directions are

$$\frac{d^2\bar{\theta}_m(x)}{dx^2} + \gamma_m^2\bar{\theta}_m(x) = 0, \qquad \frac{d\bar{\theta}_m(0)}{dx} = \frac{d\bar{\theta}_m(s)}{dx} = 0, \tag{10a}$$

$$\frac{d^2\tilde{\theta}_n(y)}{dy^2} + \sigma_n^2\tilde{\theta}_n(y) = 0, \qquad \bar{\theta}_n(0) = \bar{\theta}_n(1) = 0. \tag{10b}$$

The relative eigenfunctions and eigenvalues are

$$\bar{\theta}_0(x) = \frac{1}{s}, \quad \bar{\theta}_m(x) = \sqrt{\frac{2}{s}}\cos(\gamma_m x), \quad \gamma_m = \frac{m\pi}{s}, \quad m = 1,2,... \tag{11a}$$

$$\tilde{\theta}_n(y) = \sqrt{2}\sin(\sigma_n y), \quad \sigma_n = n\pi, \quad n = 1,2,....,. \tag{11b}$$

The GITT is based on the expansion of Eq. (7) by means of the eigenfuctions and eigenvalues Eqs. (9) and Eqs. (11). The next step in the solution procedure consists in integral transforming Eq. (7). The streamfunction transform relations pair, and the relative inverse relations, are defined as follows

$$\bar{\Psi}_i(y,t) = \int_0^s \bar{\psi}_i(x)\Psi(x,y,t)dx, \qquad \Psi(x,y,t) = \sum_{i=1}^{\infty}\bar{\psi}_i(x)\bar{\Psi}_i(y,t),$$

$$\tilde{\bar{\Psi}}_{i,j}(t) = \int_0^1 \tilde{\psi}_j(y)\bar{\Psi}_i(y,t)dy, \qquad \bar{\Psi}_i(y,t) = \sum_{j=1}^{\infty}\tilde{\psi}_j(y)\tilde{\bar{\Psi}}_{i,j}(t). \tag{12}$$

The temperature transform relations pair, and the relative inverse relations, are

$$\bar{\Theta}_m(y,t) = \int_0^s \bar{\theta}_m(x)\Theta(x,y,t)dx, \qquad \Theta(x,y,t) = \sum_{m=0}^{\infty}\bar{\theta}_m(x)\bar{\Theta}_m(y,t),$$

$$\tilde{\bar{\Theta}}_{m,n}(t) = \int_0^1 \tilde{\theta}_n(y)\bar{\Theta}_m(y,t)dy, \qquad \bar{\Theta}_m(y,t) = \sum_{n=1}^{\infty}\tilde{\theta}_n(y)\tilde{\bar{\Theta}}_{m,n}(t). \tag{13}$$

Integral transform procedure

In order to perform the integral transformation of Eq. (7), we start working on the streamfunction equation. We first multiply Eq. (7a) by the eigenfunction of the auxiliary problem for the streamfunction in the x-direction $\bar{\psi}_i(x)$ of Eqs. (9a) and then we integrate over x to obtain

$$\int_0^s \bar{\psi}_i(x)\left[\frac{\partial^2\Psi(x,y,t)}{\partial x^2} + \frac{\partial^2\Psi(x,y,t)}{\partial y^2}\right]dx = -\int_0^s \bar{\psi}_i(x)\frac{\partial\Theta(x,y,t)}{\partial x}dx. \tag{14}$$

Equation (14) can be integrated by applying the integration by parts, by applying the boundary conditions in Eqs. (7c) and by applying the inverse definition in Eqs. (12). The integration yield to the following expression

$$\frac{\partial^2\bar{\Psi}_i(y,t)}{\partial y^2} - \lambda_i^2\bar{\Psi}_i(y,t) = -\sum_{m=0}^{\infty}\bar{A}_{i,m}\bar{\Theta}_m(y,t), \tag{15}$$

With the integral transform coefficient $A_{i,m}$ defined as

$$\bar{A}_{i,m} = \int_0^s \bar{\psi}_i(x)\frac{\mathrm{d}\bar{\theta}_m(x)}{\mathrm{d}x}\mathrm{d}x \tag{16}$$

We thus multiply Eq. (15) by the eigenfunction of the auxiliary problem for the streamfunction in the y-direction $\tilde{\psi}_j(y)$ of Eqs. (9b) and then we integrate over y to obtain

$$\int_0^1 \tilde{\psi}_j(y)\frac{\partial^2 \bar{\Psi}_i(y,t)}{\partial y^2}\mathrm{d}y - \lambda_i^2 \int_0^1 \tilde{\psi}_j(y)\bar{\Psi}_i(y,t)\mathrm{d}y = -\sum_{m=0}^{\infty}\bar{A}_{i,m}\int_0^1 \tilde{\psi}_j(y)\bar{\Theta}_m(y,t)\mathrm{d}y. \tag{17}$$

On following the same procedure employed to obtain Eq. (15) one can write

$$(\lambda_i^2 + \omega_j^2)\bar{\tilde{\Psi}}_{i,j}(t) = \sum_{m=0}^{\infty}\sum_{n=0}^{\infty}A_{i,j,m,n}\bar{\tilde{\Theta}}_{m,n}(t), \tag{18}$$

where the integral transform coefficient $A_{i,j,m,n}$

$$A_{i,j,m,n} = \bar{A}_{i,m}\tilde{A}_{j,n}, \qquad \tilde{A}_{j,n} = \int_0^1 \tilde{\psi}_j(y)\tilde{\theta}_n(y)\mathrm{d}y \tag{19}$$

We may now start transforming the equation for the perturbed temperature. We thus multiply Eq. (7b) by the eigenfunction of the auxiliary problem for the temperature in the x-direction $\bar{\theta}_m(x)$ of Eqs. (10a) and then we integrate over x to obtain

$$\int_0^s \frac{\partial \Theta(x,y,t)}{\partial t}\bar{\theta}_m(x)\mathrm{d}x - \frac{\partial T_b}{\partial y}\int_0^s \frac{\partial \Psi(x,y,t)}{\partial x}\bar{\theta}_m(x)\mathrm{d}x + \int_0^s \frac{\partial \Psi(x,y,t)}{\partial y}\frac{\partial \Theta(x,y,t)}{\partial x}\bar{\theta}_m(x)\mathrm{d}x$$

$$-\int_0^s \frac{\partial \Psi(x,y,t)}{\partial x}\frac{\partial \Theta(x,y,t)}{\partial y}\bar{\theta}_m(x)\mathrm{d}x = \int_0^s \frac{\partial^2 \Theta(x,y,t)}{\partial x^2}\bar{\theta}_m(x)\mathrm{d}x + \int_0^s \frac{\partial^2 \Theta(x,y,t)}{\partial y^2}\bar{\theta}_m(x)\mathrm{d}x \tag{20}$$

$$+Ge\int_0^s \left(\frac{\partial \Psi(x,y,t)}{\partial y}\right)^2 \bar{\theta}_m(x)\mathrm{d}x + Ge\int_0^s \left(\frac{\partial \Psi(x,y,t)}{\partial x}\right)^2 \bar{\theta}_m(x)\mathrm{d}x.$$

The integration of Eq. (20) is based on the same procedure employed for the streamfunction thus using the relations in Eqs. (11) and (12) and the integration by parts. The following equation is obtained

$$\frac{\partial \bar{\Theta}_m(y,t)}{\partial t} - \frac{\partial T_b}{\partial y}\sum_{i=1}^{\infty}\bar{B}_{m,i}\bar{\Psi}_i(y,t) + \sum_{i=1}^{\infty}\sum_{o=0}^{\infty}\bar{C}_{m,i,o}\frac{\partial \bar{\Psi}_i(y,t)}{\partial y}\bar{\Theta}_o(y,t)$$

$$-\sum_{i=1}^{\infty}\sum_{o=0}^{\infty}\bar{D}_{m,i,o}\frac{\partial \bar{\Theta}_o(y,t)}{\partial y}\bar{\Psi}_i(y,t) = -\gamma_m^2\bar{\Theta}_m(y,t) \tag{21}$$

$$+Ge\left[\sum_{i=1}^{\infty}\sum_{o=1}^{\infty}\bar{E}_{m,i,o}\frac{\partial \bar{\Psi}_i(y,t)}{\partial y}\frac{\partial \bar{\Psi}_o(y,t)}{\partial y} + \sum_{i=1}^{\infty}\sum_{o=1}^{\infty}\bar{F}_{m,i,o}\bar{\Psi}_i(y,t)\bar{\Psi}_o(y,t)\right],$$

where the integral transform coefficients are defined as

$$\bar{B}_{m,i} = \int_0^s \frac{\mathrm{d}\bar{\psi}_i(x)}{\mathrm{d}x}\bar{\theta}_m(x)\mathrm{d}x, \quad \bar{C}_{m,i,o} = \int_0^s \bar{\psi}_i(x)\frac{\mathrm{d}\bar{\theta}_o(x)}{\mathrm{d}x}\bar{\theta}_m(x)\mathrm{d}x,$$

$$\bar{D}_{m,i,o} = \int_0^s \frac{\mathrm{d}\bar{\psi}_i(x)}{\mathrm{d}x}\bar{\theta}_o(x)\bar{\theta}_m(x)\mathrm{d}x, \quad \bar{E}_{m,i,o} = \int_0^s \bar{\psi}_i(x)\bar{\psi}_o(x)\bar{\theta}_m(x)\mathrm{d}x, \tag{22}$$

$$\bar{F}_{m,i,o} = \int_0^s \frac{\mathrm{d}\bar{\psi}_i(x)}{\mathrm{d}x}\frac{\mathrm{d}\bar{\psi}_o(x)}{\mathrm{d}x}\bar{\theta}_m(x)\mathrm{d}x.$$

We can now proceed multiplying Eq. (21) by the eigenfunction of the auxiliary problem for the temperature in the y-direction $\tilde{\theta}_n(y)$ of Eqs. (10b) and then integrating over y

$$
\int_0^1 \tilde{\theta}_n(y) \frac{\partial \bar{\Theta}_m(y,t)}{\partial t} \mathrm{d}y - \int_0^1 \tilde{\theta}_n(y) \frac{\partial T_b}{\partial y} \sum_{i=1}^\infty B_{m,i} \bar{\Psi}_i(y,t) \mathrm{d}y
$$

$$
+ \int_0^1 \tilde{\theta}_n(y) \sum_{i=1}^\infty \sum_{o=0}^\infty C_{m,i,o} \frac{\partial \Psi_i(y,t)}{\partial y} \bar{\Theta}_o(y,t) \mathrm{d}y - \int_0^1 \tilde{\theta}_n(y) \sum_{i=1}^\infty \sum_{o=0}^\infty D_{m,i,o} \frac{\partial \bar{\Theta}_o(y,t)}{\partial y} \Psi_i(y,t) \mathrm{d}y
$$

$$
= - \int_0^1 \tilde{\theta}_n(y) \gamma_m^2 \bar{\Theta}_m(y,t) \mathrm{d}y + \int_0^1 \tilde{\theta}_n(y) \frac{\partial^2 \bar{\Theta}_m(y,t)}{\partial y^2} \mathrm{d}y
$$

$$
+ Ge \int_0^1 \tilde{\theta}_n(y) \left[\sum_{i=1}^\infty \sum_{o=1}^\infty E_{m,i,o} \frac{\partial \Psi_i(y,t)}{\partial y} \frac{\partial \Psi_o(y,t)}{\partial y} + \sum_{i=1}^\infty \sum_{o=1}^\infty F_{m,i,o} \Psi_i(y,t) \Psi_o(y,t) \right] \mathrm{d}y.
\tag{23}
$$

What is obtained integrating Eq. (23) with the technique employed for Eq. (21) is

$$
\frac{d\tilde{\bar{\Theta}}_{m,n}(t)}{dt} - \sum_{i=1}^\infty \sum_{j=1}^\infty B_{m,n,i,j} \tilde{\bar{\Psi}}_{i,j}(t) + \sum_{i=1}^\infty \sum_{j=1}^\infty \sum_{o=0}^\infty \sum_{p=0}^\infty \left(C_{m,n,i,j,o,p} - D_{m,n,i,j,o,p} \right) \tilde{\bar{\Psi}}_{i,j}(t) \tilde{\bar{\Theta}}_{o,p}(t)
$$

$$
= -(\gamma_m^2 + \sigma_n^2) \tilde{\bar{\Theta}}_{m,n}(t) + Ge \sum_{i=1}^\infty \sum_{j=1}^\infty \sum_{o=1}^\infty \sum_{p=1}^\infty \left(E_{m,n,i,j,o,p} + F_{m,n,i,j,o,p} \right) \tilde{\bar{\Psi}}_{i,j}(t) \tilde{\bar{\Psi}}_{o,p}(t),
\tag{24}
$$

where the integral transform coefficients are defined as

$$
B_{m,n,i,j} = \bar{B}_{m,i} \tilde{B}_{n,j} = \bar{B}_{m,i} \int_0^1 \frac{\partial T_b}{\partial y} \tilde{\psi}_j(y) \, \tilde{\theta}_n(y) \mathrm{d}y,
$$

$$
C_{m,n,i,j,o,p} = \bar{C}_{m,i,o} \tilde{C}_{n,j,p} = \bar{C}_{m,i,o} \int_0^1 \frac{\mathrm{d}\tilde{\psi}_j(y)}{\mathrm{d}y} \tilde{\theta}_p(y) \tilde{\theta}_n(y) \mathrm{d}y,
$$

$$
D_{m,n,i,j,o,p} = \bar{D}_{m,i,o} \tilde{D}_{n,j,p} = \bar{D}_{m,i,o} \int_0^1 \tilde{\psi}_j(y) \frac{\mathrm{d}\tilde{\theta}_p(y)}{\mathrm{d}y} \tilde{\theta}_n(y) \mathrm{d}y,
\tag{25}
$$

$$
E_{m,n,i,j,o,p} = \bar{E}_{m,i,o} \tilde{E}_{n,j,p} = \bar{E}_{m,i,o} \int_0^1 \frac{\mathrm{d}\tilde{\psi}_j(y)}{\mathrm{d}y} \frac{\mathrm{d}\tilde{\psi}_p(y)}{\mathrm{d}y} \tilde{\theta}_n(y) \mathrm{d}y,
$$

$$
F_{m,n,i,j,o,p} = \bar{F}_{m,i,o} \tilde{F}_{n,j,p} = \bar{F}_{m,i,o} \int_0^1 \tilde{\psi}_j(y) \, \tilde{\psi}_p(y) \tilde{\theta}_n(y) \mathrm{d}y.
$$

In order to complete the initial value problem, we now integral transform the perturbed contribution of the initial values in Eqs. (4), namely

$$
\tilde{\bar{\Psi}}_{i,j}(0) = f_{i,j}, \quad \tilde{\bar{\Theta}}_{m,n}(0) = g_{m,n},
\tag{26}
$$

where

$$
f_{i,j} = \int_0^s \bar{\psi}_i(x) \, \mathrm{d}x \int_0^1 \tilde{\psi}_j(y) \, \Psi_0(x,y) \, \mathrm{d}y,
$$

$$
g_{m,n} = \int_0^s \bar{\theta}_m(x) \, \mathrm{d}x \int_0^1 \tilde{\theta}_n(y) \, \Theta_0(x,y) \, \mathrm{d}y.
\tag{27}
$$
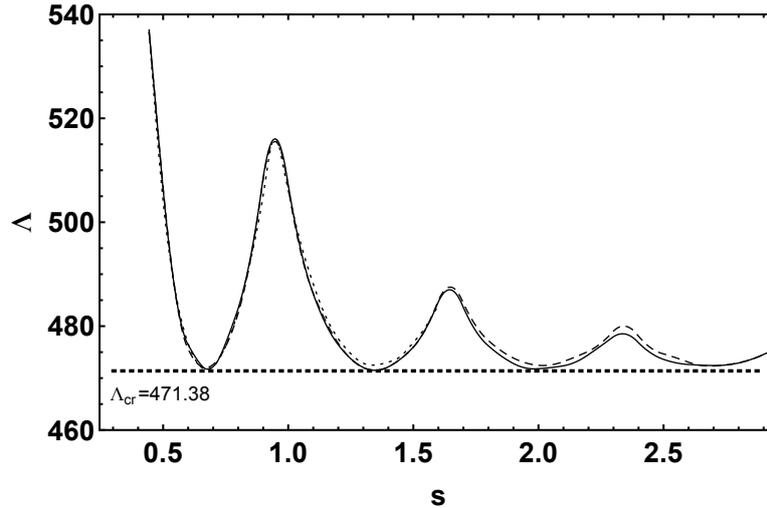
Figure 2. Critical values of Λ as a function of the aspect ratio s for $Ge \to 0$, $R = 0$ and different values of the number of equations employed n
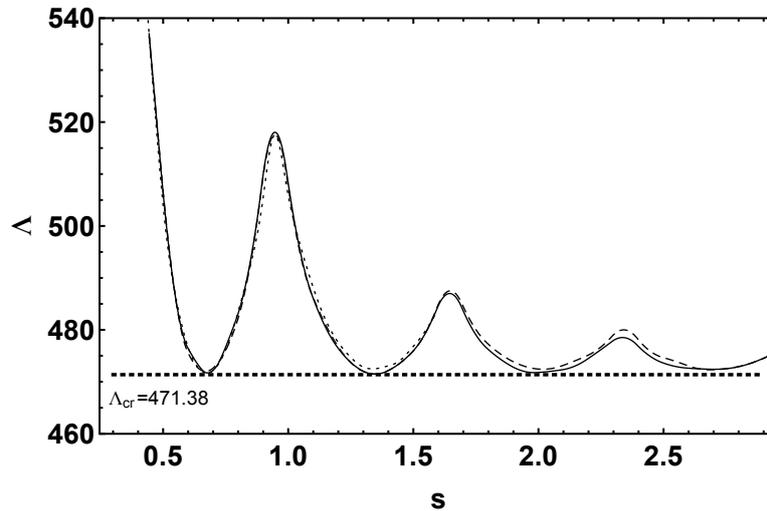
Figure 3. Critical values of Λ as a function of the aspect ratio s for $Ge = 1$, $R = 0$ and different values of the number of equations employed n

Discussion of the results and concluding remarks

The task of this investigation is comparing the results obtained by the nonlinear analysis with the values obtained by the linear stability analysis. We start switching off the contribution of the Darcy-Bènard-like instability source: whenever the temperature gap between the horizontal boundaries is negligible, *viz.* $R = 0$, the relative buoyancy force contribution is absent. With $R = 0$ the linear stability analysis responds a threshold value for the governing parameter $\Lambda_{cr} = 471.38$ and a threshold value for the wavenumber $k_{cr} = 4.6752$, [Nield and Barletta (2010)]. The subscript *cr* stands for critical value. Figures (2) and (3) show the neutral stability curves obtained fixing $R = 0$ for different values of *Ge* as functions of the aspect ratio s. The different curves reported in the frames refer to different values of the number of equations employed to model the problem. The eigenfunction expansion has, in fact, to be truncated at some point and n is the number of equations obtained with the different choices of the truncation point. In Fig. 2 and Fig. 3 the dotted lines refer to those truncation points that produce a number of equations $n = 30$, the dashed lines refer to $n = 60$ and the continuous lines to $n = 90$. Figures (2) and (3) prove that the present nonlinear analysis reproduce exactly the same
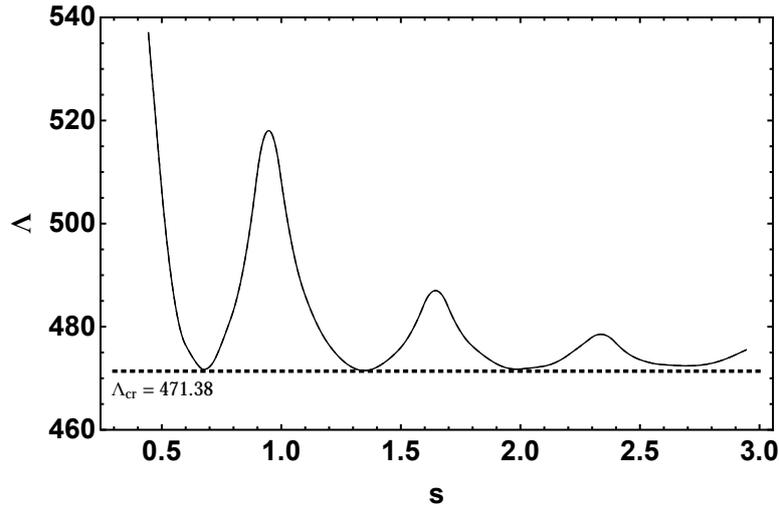
Figure 4. Critical values of Λ as a function of the aspect ratio s for a given value of the number of equations employed, $n = 90$, and two different values of the pair (Ge, R): $(Ge \to 0, R = 0)$, continuous line, and $(Ge = 1, R = 0)$, dashed line
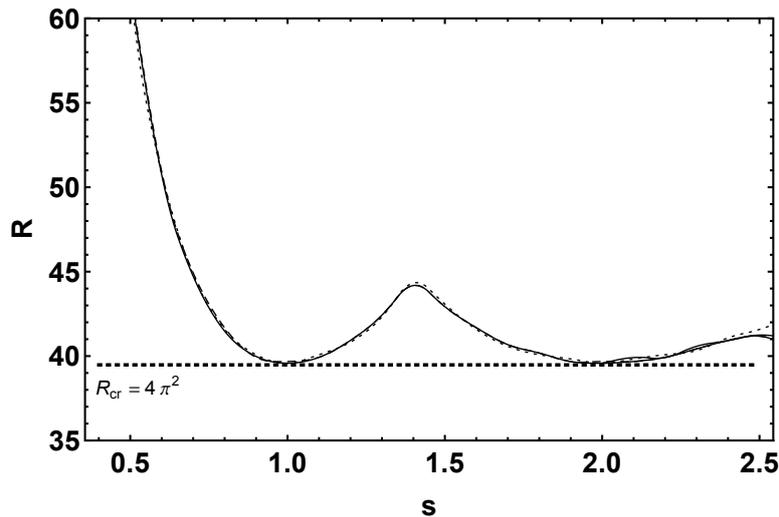
Figure 5. Critical values of R as a function of the aspect ratio s for $Ge \to 0$, $\Lambda = 10^{-4}$ and different values of the number of equations employed n

minimum (highlighted by the horizontal dashed line $\Lambda_{cr} = 471.38$) obtained by the linear stability analysis. The value of the aspect ratio, relative to second minimum of Λ, results to be equal to the wavelength corresponding to the critical wavenumber value $k_{cr} = 4.6752$ obtained by the linear stability analysis. The difference between Fig. 2 and Fig. 3 lies on the values of the Gebhart number assumed: Fig. 2 refers to $Ge \to 0$ and Fig. 3 is refers to $Ge = 1$. The limit $Ge \to 0$ is compatible with finite values of Λ if we consider, together with the limit $Ge \to 0$, a fast basic flow, *viz.* $Pe \gg 1$. Figure (4) shows a comparison between the neutral stability curves obtained with $R = 0$, $n = 90$ and two different values of the Gebhart number: $Ge \to 0$, the continuous line, and $Ge = 1$, the dashed line. This figure shows how the system results not to be sensitive to the magnitude of the Gebhart number. We proceed switching off the viscous dissipation and looking for the Darcy-Bénard-like instability. In order to neglect the viscous dissipation contribution the limits $Ge \to 0$ and $\Lambda \to 0$ are considered. In this case the linear stability analysis responds a threshold value for the critical parameter $R_{cr} = 4\pi^2$ and a wavenumber $k_{cr} = \pi$, [Nield and Barletta (2010)]. Figure (5) shows the neutral stability curves for $Ge \to 0$ and $\Lambda = 10^{-4}$. Once again the dotted line refer to those truncation points that produce a

number of equations $n = 30$, the dashed line refer to $n = 60$ and the continuous line to $n = 90$. These curves show that the present nonlinear analysis reproduces exactly the same minimum (highlighted by the horizontal dashed line $R_{cr} = 4\pi^2$) found by the linear stability analysis. Moreover, the critical aspect ratio value relative to the second minimum in Fig. 5 is equals to the wavelength corresponding to the critical wavenumber value $k_{cr} = \pi$ found by the linear stability analysis. The nonlinear stability analysis here proposed is preliminary investigation of the problem presented. A partial investigation of the parametric range is indeed presented. Nonetheless it is a fairly good starting point for the investigation of nonlinear thermal instabilities. We may conclude that, in the parametric range here studied, the setup investigated does not present subcritical instabilities and the nonlinear stability analysis does not highlight a change behaviour with respect to the linear analysis.

References

Nield, D. A., Bejan, A. (2013), Convection in Porous Media, 4th edition. Springer, New York.

Nield, D.A., (2007), The modeling of viscous dissipation in a saturated porous medium. *ASME Journal of Heat Transfer*, **129**, 1459–1463.

Storesletten, L., Barletta, A., (2009), Linear instability of mixed convection of cold water in a porous layer induced by viscous dissipation. *International Journal of Thermal Sciences*, **48**, 655–664.

Barletta, A., Celli, M., Rees, D. A. R. (2009), Darcy-Forchheimer flow with viscous dissipation in a horizontal porous layer: onset of convective instabilities. Stability analysis of natural convection in porous cavities through

Nield, D. A., Barletta, A., (2010), Extended Oberbeck-Boussinesq approximation study of convective instabilities in a porous layer with horizontal flow and bottom heating. *International Journal of Heat and Mass Transfer*, **53**, 577–585.

Cotta, R. M., (1998), The integral transform method in thermal & fluids sciences & engineering. Begell House, New York.

Pontes, J., de B. Alves, L. S., Cotta, R. M., (2002), Stability analysis of natural convection in porous cavities through integral transforms. *International Journal of Heat and Mass Transfer*, **45**, 1185–1195.

Numerical Simulation of Drops Impacting on Textured Surfaces

Y.X. Wang, *S. Chen

School of Aerospace Engineering and Applied Mechanics, Tongji University, China.

*Corresponding author: schen_tju@tongji.edu.cn

Abstract

In the present study drops impacting on textured surfaces were investigated using many-body dissipative particle dynamics (MDPD). A novel linear-jointed solid/fluid interaction with short-range repulsive and long-range attractive forces was used to generate different wettability and a simple but efficient numerical method was introduced to measure the contact angle. A systematic study was carried out to obtain the relationship between initial impact velocity and spreading properties on surfaces with different wettability induced by chemistry and topology. The simulated results showed that the drop is easily rebounded at lower $\emptyset_S$ with high impact velocity and only spreading occurs at low impact velocity or larger $\emptyset_S$. Referring to the triple-phase contact line, the time-based retraction phase was divided into two periods, and it was analyzed from an energy transition and dissipation viewpoint. It is expected to provide simulation details for the water-repellency surfaces design.

Keywords: Drop, Impact, Contact angle, Surface roughness, Many-body dissipative particle dynamics

Introduction

For years, surface wetability, especially the superhydrophobicity, has been the focus of surface science. Superhydrophobicity can cause drop to roll very easily off solid surfaces or bounce back upon impacting, just like the lotus leaf which can make rain drops roll off in ball easily (Lotus effect) [Barthlott and Neinhuis (1997)]. Generally speaking, superhydrophobicity means that the contact angle is larger than 150° while the contact angle hysteresis less than 10°, which confers to drops a high mobility on these surfaces. Unfortunately, for smooth and flat surfaces the possible highest contact angle is less than 120º if without special processing. In nature, there are many plants and animals showing superhydrophobicity besides the lotus leaf, such as the antifogging mosquito eyes [Gao et al. (2007)] and legs of a water strider [Gao and Jiang (2004)] and feathers of many birds. Recently, scientists decoded the mechanism of them and found that the microtextured and nanotextured roughness contributes significantly to the quality of the water-repellency property. This breakthrough has attracted lots of researches to investigate drop wetting states, such as the Cassie [Cassie and Baxter (1944)] and Wenzel [Wenzel (1936)] state (Figure 1 Left and Right), and dynamic behavior on textured surfaces. The Cassie state is often described as "air trapping" or composite surface which means the liquid bridges between surface protrusions and no longer penetrates the interspace where it is filled with air (Figure 1). Cassie and Baxter considered the contact angle on heterogeneous surfaces composed of two different materials (solid surface and air), in which $\emptyset_S$ and $1 - \emptyset_S$ ($\emptyset_S = W/L$, Figure 1) are the fractional areas of the wetted solid/liquid and liquid/air interfaces, respectively. Based on this viewpoint, they gave the Cassie-Baxter theoretical formula, $\cos\theta_C = \emptyset_S \cos\theta_Y + (1 - \emptyset_S) \cos 180°$. This formula shows that the trapped air always drives the Cassie angle, θ_C, from Young angle to a larger contact angle (Figure 2). Many researches show great interests on this kind of

drop state. Singh et al. [Singhet al. (2013)] draped the nanotextured surface with a single-layer graphene sheet to avoid the water intruding into the textured surface features and keep the drop in an ideal Cassie state. Their experiment showed that the seperated drop on the graphene sheet could hardly be pinned to the substrate and the contact angle hysteresis was reduced significantly. Also, Kim et al. [Kim et al. (2012)] did a systematic research about drop impact characteristics on multiscaled rough surfaces and found only on the nanoscaled surface the drop could rebound completely. From the static and dynamic contact angle measurement results they found the nanostructures showed superhydrophobic properties and were close to the nonwetted state (Cassie state). The theoretical analysis and physical experiments show that the Cassie state can provide a smaller angle hysteresis as well as a larger contact angle, thence, in many industrial applications such as self-cleaning surfaces people show more interests in the Cassie state than the Wenzel state. However, in experiments it's not easy to hold Cassie state when there exists higher impacting velocity of drop, which can cause a Cassie-to-Wenzel transition. However the ideal Cassie state without transition could be investigated by numerical method. In this paper, the drop spreading dynamics under different wettability and impact velocities in ideal Cassie state was studied by using a particle-based numerical method (MDPD).
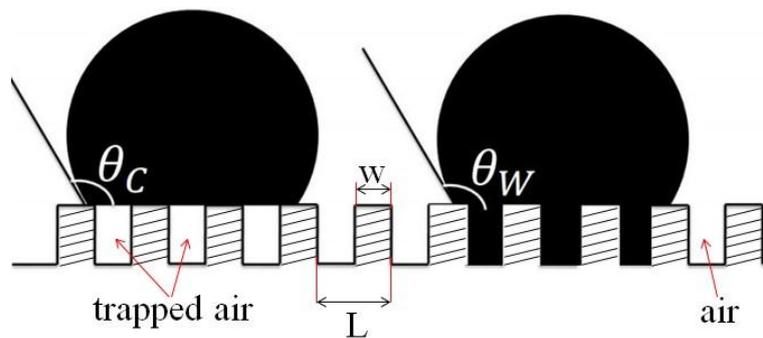
Figure 1 Drop states on textured surface
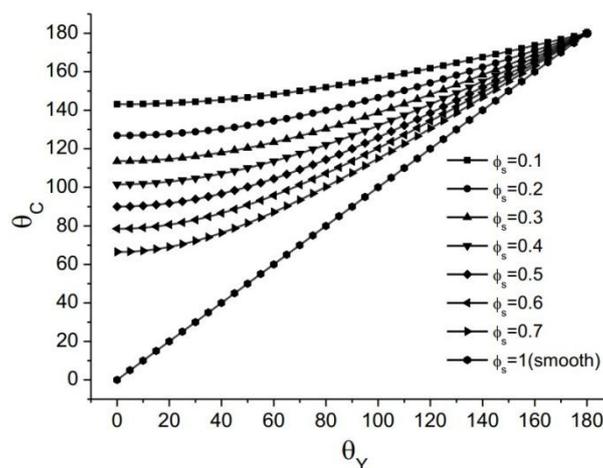
Figure 2 Theoretical Cassie angles on different rough surfaces

Computational method

MDPD

MDPD [Warren (2003); Trofimov et al. (2005)] is a modified numerical method of DPD [Hoogerbrugge and Koelman (1992); Groot and Warren (1997)] which includes the van der Waals loop in the EOS so as to make the DPD suitable for simulations of fluid systems with free surfaces. In the present work, we employ the MDPD scheme reported by Warren and in this scheme both the random and dissipative forces are kept the same with classical DPD while conservative force is revised as EQ.(1):

$$\mathbf{F}_{ij}^C = A\omega_c(r_{ij})\mathbf{e}_{ij} + B(\bar{\rho_i} + \bar{\rho_j})\omega_d(r_{ij})\mathbf{e}_{ij} \qquad (1)$$

in which $A\omega_c(r_{ij})\mathbf{e}_{ij}$ is the attractive part by setting A<0 and the weight function $\omega_c(r_{ij})$ is defined as the classical DPD, $\omega_c(r_{ij}) = 1 - r_{ij}/r_c$. $B(\bar{\rho_i} + \bar{\rho_j})\omega_d(r_{ij})\mathbf{e}_{ij}$ is the repulsive part which depends on a weighted average of the local density by defining $\rho_i(r_{ij})$ as follows:

$$\rho_i(r_{ij}) = \sum_{i\neq j} \omega_\rho(r_{ij}) \qquad (2)$$

$$\omega_\rho(r_{ij}) = 15/(2\pi r_d^3)(1 - r_{ij}/r_d)^2 \qquad (3)$$

The weight functions ω_d in EQ.(1) are in similar definitions with $\omega_c(r_{ij})$, $\omega_d(r_{ij}) = 1 - r_{ij}/r_d$, but with different cutoff distances: $r_c = 1$ and $r_d = 0.75$.

Boundary Condition

To generate various hydrophobic and hydrophilic wetting behaviors, the solid/fluid conservative force is modeled by combining short-range repulsive and long-range attractive forces. Here, three linear weight functions are simply joined together and this is some different from the bell-shaped weight function of smoothed particle hydrodynamics.

$$F_{sl}^C = \begin{cases} F_b(1 - r_{ij}/r_a) & , \; r_{ij} \leq r_b \\ F_a/(r_b - r_a)(r_b - r_{ij}) & , \; r_b < r_{ij} \leq r_a \\ F_a/(r_{sl} - r_a)(r_{sl} - r_{ij}) & , \; r_a < r_{ij} \leq r_{sl} \end{cases} \qquad (4)$$

F_{sl}^C is the conservative force between solid and fluid particles which depends on their distance r_{ij} .The parameters F_a and F_b in EQ.(4) determine the strength of the attractive and repulsive interactions, r_{sl} is the wall-fluid interaction range and the two sub-ranges r_a and r_b are the positions of the maximum attractive force and the vanishing of the repulsive force (Figure 3). To simplify the simulation, F_b, r_{sl}, r_a and r_b are fixed and only F_a is changed to generate different wettability. Figure 4a shows a gently deposited drop on flat surfaces with different F_a, a function between F_a and the static contact angle (or the so-called Young angle) can be obtained. In this paper only the case that the fluid is hydrophobic on the flat solid surface is considered, so F_a is fixed at 70 and the related contact angle is around 118°.
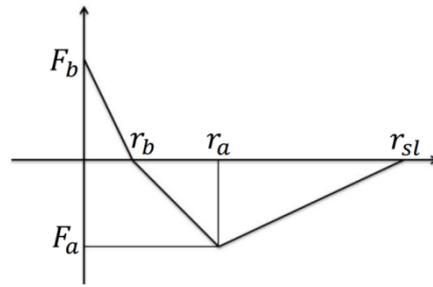
Figure 3 Solid/fluid interaction function curve
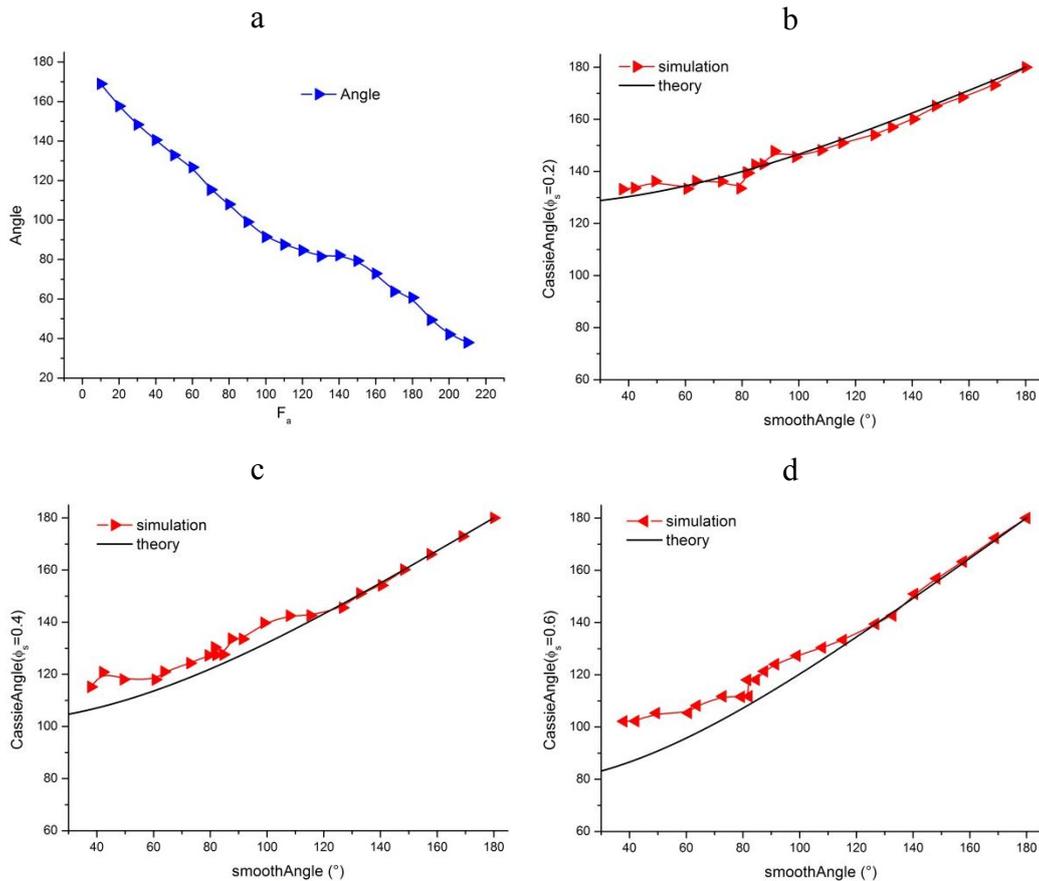
Figure 4 The contrast of simulated contact angles and Cassie theoretical values

Avoiding penetrating is another important problem and also for getting the ideal Cassie state the bounce-back reflections boundary condition is employed at the top of the pillars and air cushions which separate the drop from the solid substrate. Here the particles which penetrate the wall are forced back to the position of the last time step without changing their velocities. By these settings, any desired wettability can be obtained and also the algorithm is very robust. Furthermore the MDPD contact angle and the theoretically predicted angle (see the Cassie theory formula) are compared and a good agreement of them are shown in Figure 4b, 4c and 4d which correspond to $\emptyset_S = 0.2$, $\emptyset_S = 0.4$ and $\emptyset_S = 0.6$.

Measurement of the contact angle

Different from the fitting circle method [Koishi et al. (2009)], a simple but efficient numerical method by using geometrical computation is engaged to measure the contact angle between the solid surface and drop. Gently deposited drop is used to get the static contact angle. When the drop is stable on the solid surface, the position of the particles which enter the thin layer near the solid surface (upon the top of the pillars) are recorded, then, the difference between the maximum and the minimal values in the X direction is calculated as the length L of the contact line. Then, for all fluid particles, the maximum value in the Y direction is considered as the height H of the contact height. At last, a concise geometrical formula as follow is used to obtain the contact angle θ.

$$\theta = 2\tan^{-1}(2H/L) \tag{6}$$

Results and discussion

Time-Resolved Impact

Figure 5 shows snapshots of drops impacting on textured surfaces with $\emptyset_S = 0.2, 0.6$ and smooth surface ($\emptyset_S = 1$) at different velocities 0.1 and 10. The images in the first row show drops just before first contacting the surfaces. Before impacting, the drops are almost spherical with diameter D_0. After impacting, the drops spread until they reach a maximal spreading diameter, D_{max} (second row). The third row shows the minimal spreading diameter after the maximum spreading and in the last row most drops are in a stable state excepting the drop in second column which bounces back into the air. At high velocity (V=10), the drops are strongly deformed into a flat film much thinner than drop diameter D_0 in the middle region and gibbous rim emerging at the edge makes the drops look like a pancake, the similar result was observed by some other simulation [Eggers et al. (2010)] and physical experiments [Deng et al. (2013); Kim et al. (2012)]. In the retraction phase (third row and fourth row), the drop shape and state (rebound or pinning) depend on both the roughness of the surfaces and the deposited velocities. When deposited at high velocity, the drops elongate on the textured surfaces, but rebounding from the surface only happens at low $\emptyset_S$ ($\emptyset_S = 0.2$) and on the smooth surface the drop almost keeps its shape throughout the process. On the contrast, when the drops are deposited gently, the contact lengths are almost unchanged after the maximum spreading. We suppose that the kinetic energy of the drop is insufficient for the drop to move far and take shape of a pancake.

$\emptyset_S = 0.2$		$\emptyset_S = 0.6$		$\emptyset_S = 1$	
$V = 0.1$	$V = 10$	$V = 0.1$	$V = 10$	$V = 0.1$	$V = 10$
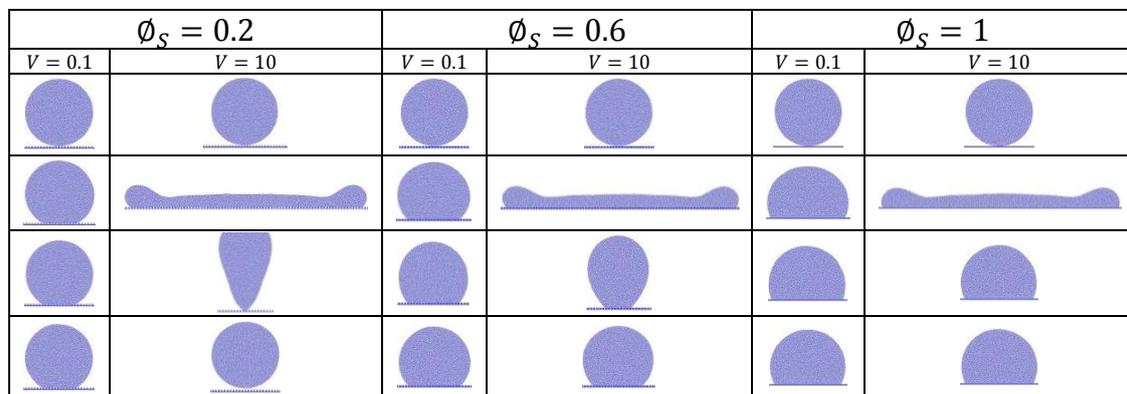
Figure 5 Time-resolved drop dynamics at different impact velocities

Time-Based Spreading Dynamics

The spreading diameter of the drop on the solid surfaces reflects the spreading and retraction dynamics and the energy dissipation of the impacting drop. For all surfaces, spreading diameters increase at first, then reach a maximum, and at last decreases. The maximal spreading diameter of the drop increases with increasing deposited V and spreading is always faster than retraction. Figure 6 shows that the temporal evolution of the spreading diameter depends on impact velocity significantly. Our simulation shows a good agreement with Deng's experiment [Deng et al. (2013)] in this period. Figure 7 shows two special snapshots in the drop spreading just before and after the drop reaching its maximum spreading diameter at V=10 and $\emptyset_S = 0.2$, the related dynamical contact angles of them can be seen as the advancing contact angle θ_A and receding contact angle θ_R [Deng et al. (2013)], the small difference between them indicates a small contact angle hysteresis. According to Deng' treatment, we also separate the retraction phase into two periods which can be divided at the time of their minimal spreading diameters respectively. In the first retraction period, the spreading diameters decrease with time monotonously, but there are some different phenomena between different textured surfaces: for low $\emptyset_S$ ($\emptyset_S = 0.2$), the drops start to rebound from low impact velocity (V=3) and from relative high velocity (V=6) for $\emptyset_S = 0.4$, but no rebounding occurs at $\emptyset_S = 0.6$ and smooth surface. In the second retraction period, the pinning drops undergo a slight fluctuation around their final spreading diameters respectively, that is because when the drops meet their minimal spreading diameter, the drops elongate (Figure 5, third row) and transform the retraction kinetic energy into interface energy and makes the drop unstable, after that, energy is dissipated through contact line fluctuating. From a DPD viewpoint, the dissipative force between solid particles and fluid particles can play an important role in this period. As a contrast, drop on the smooth surface ($\emptyset_S = 1$) only experiences the first retraction period, reflecting strong adhesion of the drop on the smooth surface. For the rebounding drops, the contact time (from the first contact with solid surfaces to the rebounding time) also depends on roughness and impact velocities: the higher the speed, the longer the contact time for the same textured surface and the contact time always shorter for the low $\emptyset_S$ under the same impact velocity(Figure 8). From all the four cases we also found that the equilibrated spreading diameters are independent of the impact velocities and larger at the high $\emptyset_S$, including $\emptyset_S = 1$, but not too obviously.

c   d
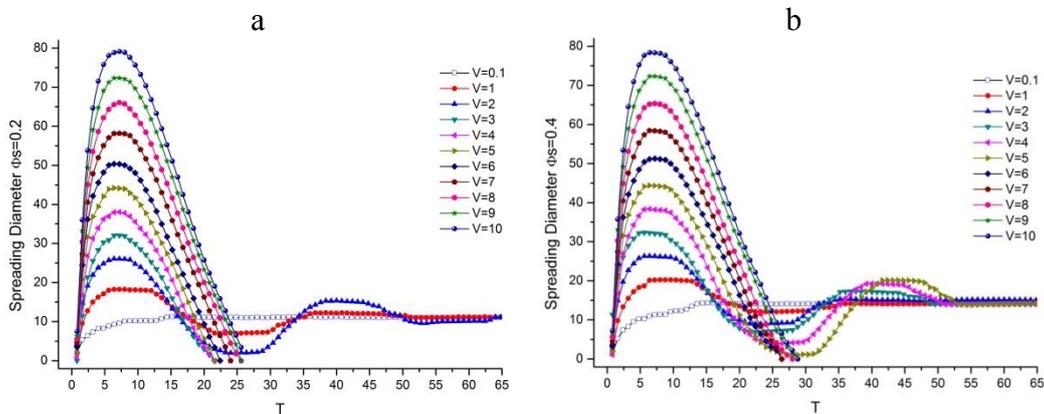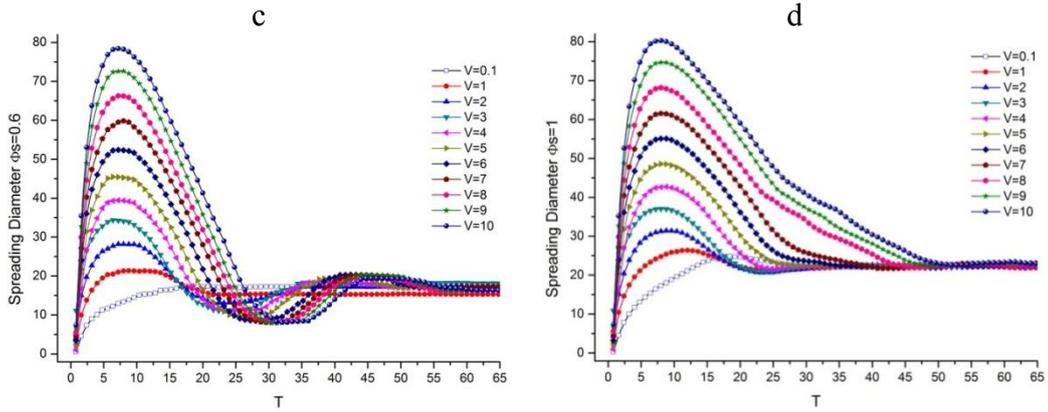
Figure 6 Time-based evolution of the spreading diameter
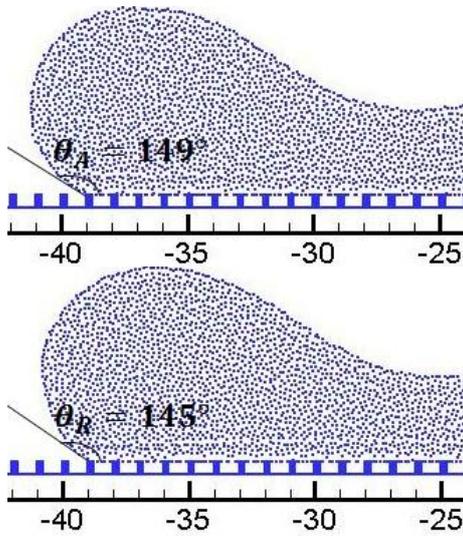
Figure 7 The advancing and receding "pancake"
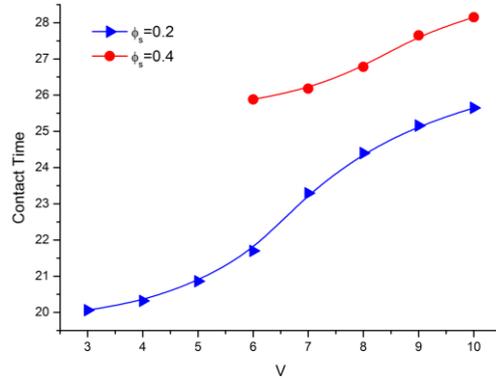
Figure 8 Contact time

Recently, a new viewpoint has been proposed trying to explain the contact line pinning by the so-called effective liquid Hammer pressure [Deng et al. (2009)]. Hammer pressure is caused by the hitting of the drop on the surface and the liquid is compressed which creates a shock wave that adds a vertical component to the velocity of the fluid, then, the shock wave relaxes as soon as it overtakes the moving contact line. In the future we will do more research about Hammer pressure. Figure 9 shows the compression of liquid.
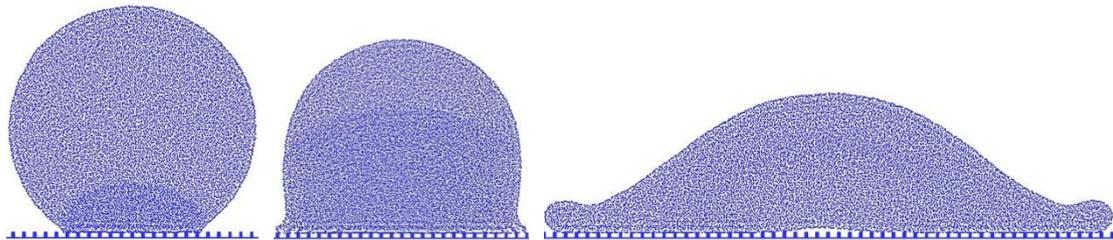
Figure 9 Drop at different moment just after impact ($\emptyset_S = 0.2$, V=10)

Conclusions

In this paper, a particle-based numerical method, MDPD, was built to simulate drop impact on textured surfaces. The simulated values show a good agreement with Cassie-Baxter theory. Also, the linear-jointed solid/fluid interaction and the contact angle measurement work well in the simulation.

After impacting on hydrophobic surfaces, drops undergo rebounding or pinning. The maximum spreading diameters and contact time depend on the different textured surfaces and velocities. At low velocities, the drops always pin on the surfaces and the maximum spreading diameters are small. When speeding the drops, at low $\emptyset_S$ the drop will rebound easily but pin at high $\emptyset_S$. The retraction phase shows two well-separated periods, i.e. a monotonous decreasing one and a fluctuant one. In the first period, energy stored in the deformation of the surface is transformed back into kinetic energy and this phase is inertia-dominated. The existence of a fluctuant period shows that the transforming between kinetic energy and the interface energy propel the movement of the contact line, which leads a dissipation of the total energy and make the drop stable at the end. At last, some evidences about Hammer pressure have been found and this may open a door for the farther research about the mechanism of drop spreading dynamics.

References

Barthlott, W. and Neinhuis, C. (1997) Purity of the sacred lotus, or escape from contamination in biological surfaces *Planta* 202, 1

Gao, X.F., Yan, X., Yao, X., Xu, L., Zhang, K., Zhang, J.H. and Yang, B. (2007) The dry-style antifogging properties of mosquitocompound eyes and artificial analogues prepared by soft lithography*Adv. Mater.* 19, 2213

Gao, X. and Jiang, L.(2004) Biophysics: water-repellent legs of water striders*Nature*432(7013),36

Wenzel,R.N.(1936) Resistance of solid surfaces to wetting by water,*Industrial & Engineering Chemistry* 28(8), 988

Cassie,A.B.D. and Baxter,S.(1944) Wettability of porous surfaces,*Transactions of the Faraday Society* 40,546

Singh, E., Thomas, A. V., Mukherjee, R., Mi, X., Houshmand, F., Peles, Y., Shi, Y.F. and Koratkar, N. (2013) Graphene Drape Minimizes the Pinning and Hysteresis of Water Drops on Nanotextured Rough Surfaces. *ACS nano*7(4), 3512

Kim,H., Lee,C., Kim,M.H.and Kim,J.(2012) Drop Impact Characteristics and Structure Effects of Hydrophobic Surfaces with Micro- and/or Nanoscaled Structures,*Langmuir* 28, 11250

Deng,X., Schellenberger,F., Papadopoulos,P., Vollmer,D. and Butt,H.J. (2013) Liquid Drops Impacting Superamphiphobic Coatings,*Langmuir* 29, 7847

Eggers,J., Fontelos,M.A., Josserand,C., and Zaleski,S. (2010) Drop dynamics after impact on a solid wall: Theory and simulations,*Physics of Fluids* 22, 062101

Warren, P.B.(2003) Vapour-liquid coexistence in many-body dissipative particle dynamics,*Phys. Rev. E* 68 066702

Trofimov, S.Y., Nies, E.L.F. and Michels, M.A.J. (2005) Constant-pressure simulations with dissipative particle dynamics,*The Journal of chemical physics* 123 144102

Hoogerbrugge,P.J. and Koelman, J.M.V.A. (1992) Simulating Microscopic Hydrodynamic Phenomena with Dissipative Particle Dynamics,*Europhys. Lett.* 19 155

Groot,R.D. and Warren, P.B. (1997) Dissipative particle dynamics: Bridging the gap between atomistic and mesoscopic simulation,*The Journal of chemical physics* 107 4423

Koishi, T., Yasuoka, K., Fujikawa, S., Ebisuzaki, T. and Zeng, X. C. (2009) Coexistence and transition between Cassie and Wenzel state on pillared hydrophobic surface,*Proceedings of the National Academy of Sciences* 106 8435

Deng, T., Varanasi, K.K., Hsu, M., Bhate, N., Keimel, C., Stein, J. and Blohm, M. (2009) Nonwetting of impinging droplets on textured surfaces, *Applied Physics Letters* 94, 133109

Adaptive polygon scaled boundary finite element method for elastodynamics

*Z.H. Zhang [1]and Z.J. Yang[2]

[1]College of Civil and Environmental Engineering, Ningbo University, China
[2]College of Civil Engineering and Architecture, Zhejiang University, China

*Corresponding author: zhangzihua@nbu.edu.cn

Abstract

An adaptive polygon scaled boundary finite element method (APSBFEM) is developed for elastodynamic problems. Flexible polygon meshes are generated from background Delaunay triangular meshes and used to calculate structure's dynamic responses. In each time step, a posteriori-type energy error estimator is employed to locate the polygon subdomains with exceeding spatial discretization error, then edge midpoints of the corresponding triangles are inserted into the background. A new Delaunay triangular mesh and a polygon mesh are regenerated successively. The state variables, including displacement, velocity and acceleration are mapped from the old polygon mesh to the new one by a simple algorithm. A benchmark elastodynamic problem is modeled to validate the developed method. The results show that the adaptive meshes are capable of capturing the steep stress gradient areas, and the dynamic responses agree well with those from the adaptive finite element method and the general polygon scaled boundary finite element method using fine meshes.

Keywords: Adaptive, Scaled Boundary Finite Element Method, Polygon, Elastodynamics, Energy error estimator

Introduction

Spatial discretization error inevitably exists in numerical methods, especially for dynamic problems. In order to limit the discretization error within an acceptable level, adaptive methods can be used to refine the steep stress gradient areas with exceeding error automatically by means of error estimators and remeshing procedures. Adaptive finite element methods (AFEM) [Zeng and Wiberg (1992)] have been developed to seek appropriate spatial discretization with the least computational cost, but the remeshing procedures are complicated and time-consuming, especially for large scale problems. Furthermore, mesh mapping after remeshing to transfer state variables from the old mesh to the new one is approximate in FEM, leading to high accumulative errors in subsequent time steps.

The scaled boundary finite element method (SBFEM) [Song and Wolf (1997)] is a semi-analytical method combining the advantages of the finite element method (FEM) and the boundary element method (BEM). The domain consists of a small number of large-sized subdomains and only the subdomain boundaries need to be discretized. The modeled dimensions are reduced by one as the BEM, but no fundamental solutions are needed. Consequently, the FEM's flexibility and the BEM's simplicity in pre-processing and remeshing are both retained.

Polygon elements are widely used in FEM and have two attractive features. First, polygon elements are flexible in meshing domains with complex geometries such as polycrystal. Second, polygon elements generally have superior accuracy because of their high order shape functions. Recently, a versatile procedure is developed to generate polygon mesh from Delaunay triangulation and applied to static and dynamic crack propagation modeling [Ooi et al. (2012; 2013)]. It is demonstrated that the polygon scaled boundary finite element method (PSBFEM) is good at dealing with domains with complicated geometries and singularities, not only in pre-processing but also in remeshing, while the high accuracy of SBFEM is retained.

Combining the polygon subdomains with a simple remeshing procedure, a novel adaptive polygon SBFEM (APSBFEM) for elastodynamics is developed. This paper is organized as follows: Section 2 discusses the SBFEM and its solutions in time domain briefly. Section 3 presents a posteriori

energy error estimator. A simple adaptive method is described in Section 4, including the remeshing procedure and the mesh mapping algorithm, and a flowchart is given out as well. A benchmark elastodynamics problem is modeled and discussed in Section 5. Conclusions are stated in Section 6.

Methodology

The scaled boundary finite element method

A domain consists of 3 polygon subdomains is described in Figure 1(a). Figure 1(b) shows the details of Subdomain 1. The subdomain is represented by scaling a defining curve S relative to a scaling center and the entire subdomain boundary has to be visible from the scale center. A normalized radial coordinate ξ is defined, varying from zero at the scaling center and unit value on S. A circumferential coordinate η is defined around the defining curve S. A curve similar to S defined by $\xi=0.5$ is shown in Figure 1(b). The coordinates ξ and η form a local coordinate system used in all the subdomains and simple transformation equations between the local and global Cartesian coordinates can be established through:

$$x = x_0 + \xi\left(\frac{x_1 + x_2 - 2x_0}{2} + \frac{(x_2 - x_1)\eta}{2}\right) \tag{1}$$

$$y = y_0 + \xi\left(\frac{y_1 + y_2 - 2y_0}{2} + \frac{(y_2 - y_1)\eta}{2}\right) \tag{2}$$

where (x_1, y_1) and (x_2, y_2) are nodal coordinates of an element on the boundary and (x_0, y_0) are the coordinates of the scaling center.

The displacements of any point (ξ, η) in a subdomain are calculated by

$$\mathbf{u}(\xi, \eta) = \mathbf{N}(\eta)\mathbf{u}(\xi) \tag{3}$$

where $\mathbf{u}(\xi)$ are the displacements along the radial lines and are analytical with respect to ξ. $\mathbf{N}(\eta)$ are the shape functions in the circumferential direction which are the same as used in FEM.

The stresses in a subdomain are calculated by

$$\boldsymbol{\sigma}(\xi, \eta) = \mathbf{D}\mathbf{B}^1(\eta)\mathbf{u}(\xi),_\xi + \frac{1}{\xi}\mathbf{D}\mathbf{B}^2(\eta)\mathbf{u}(\xi) \tag{4}$$

where $\mathbf{B}^1(\eta)$ and $\mathbf{B}^2(\eta)$ are coefficient matrixes, and $\mathbf{D}$ is the elasticity matrix.

Solutions in time domain

In elastodynamics, the equilibrium condition of a subdomain without body loads can be formulated using the virtual work statement [Yang et al. (2011)]

$$\int_{V_s} \delta\boldsymbol{\varepsilon}^T \boldsymbol{\sigma} dV - \int_{S_s} \delta\mathbf{u}^T \mathbf{t} ds + \int_{V_s} \delta\mathbf{u}^T \rho\ddot{\mathbf{u}} dV = 0 \tag{5}$$

where $\delta\boldsymbol{\varepsilon}$ is the virtual strain vector, $\delta\mathbf{u}$ the virtual displacement vector, $\boldsymbol{\sigma}$ the stress vector, $\mathbf{u}$ the displacement vector, $\ddot{\mathbf{u}}$ the acceleration vector, ρ the material density, $\mathbf{t}$ the traction on the boundary, V_s the volume and S_s the boundary of the subdomain.

(a) Subdomaining of a domain
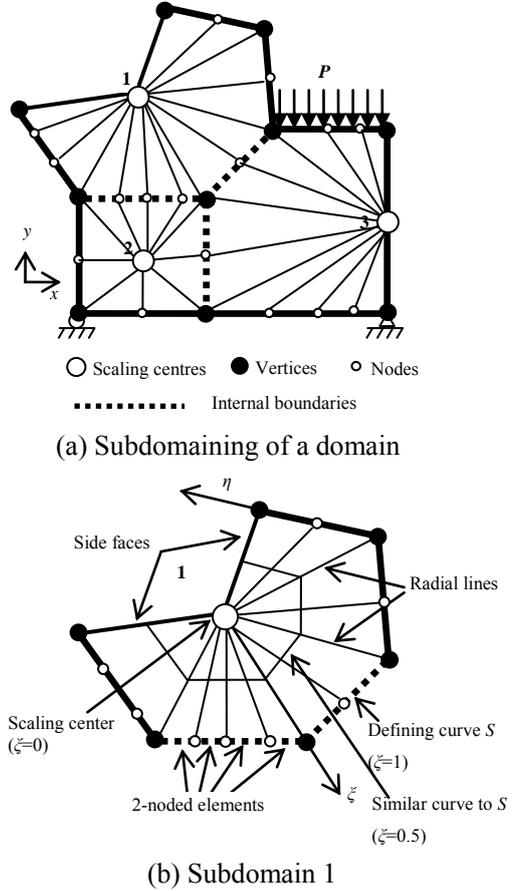
(b) Subdomain 1

Figure 1. Concept of SBFEM

The dynamic equilibrium equation of a subdomain is derived as

$$\mathbf{M}_s\ddot{\mathbf{u}}_\mathbf{b} + \mathbf{K}_s\mathbf{u}_\mathbf{b} = \mathbf{p}_s \tag{6}$$

where $\mathbf{u}_\mathbf{b}$ is the displacement vector and $\ddot{\mathbf{u}}_\mathbf{b}$ is the acceleration vector on the subdomain boundary, $\mathbf{p}_s$ the subdomain load vector, $\mathbf{K}_s$ the subdomain stiffness matrix and $\mathbf{M}_s$ the subdomain mass matrix.

The nodal displacement vector on the subdomain boundary is

$$\mathbf{u}_\mathbf{b} = \mathbf{\Phi c} \tag{7}$$

where $\mathbf{\Phi} = \{\mathbf{\phi}_1, \mathbf{\phi}_2, \ldots, \mathbf{\phi}_N\}$ is a matrix in which the vectors $\mathbf{\phi}_i$ are obtained from solving a standard eigen problem and $\mathbf{c} = \{c_1, c_2, \ldots, c_N\}^T$ are constants (modal participation factors) dependent on boundary conditions, and N is the number of degrees of freedom (DOFs) of the subdomain.

Time integration scheme

Assembling Eq. (6) for all subdomains leads to the global equation system

$$\mathbf{M\ddot{U}} + \mathbf{KU} = \mathbf{P} \tag{8}$$

where $\mathbf{M}$ and $\mathbf{K}$ are the assembled global mass and stiffness matrices, $\mathbf{P}$ the global load vector, $\mathbf{U}_n$ and $\ddot{\mathbf{U}}_n$ the nodal displacement and acceleration vectors respectively. The Newmark integration scheme is used to solve Eq. (8) in this study and $\beta = 0.25$ and $\gamma = 0.5$ are used with unconditional stability.

The subdomain displacement field is then obtained as

$$\mathbf{u}(\xi,\eta) = \mathbf{N}(\eta)\sum_{i=1}^{N} c_i\xi^{\lambda_i}\mathbf{\phi}_i \tag{9}$$

where λ_i ($i=1$-N) are eigen values from solving a standard eigen problem.

The stress field in the subdomain is then calculated from Eq. (4)

$$\mathbf{\sigma}(\xi,\eta) = \mathbf{DB}^1(\eta)\left(\sum_{i=1}^{N} c_i\lambda_i\xi^{\lambda_i-1}\mathbf{\phi}_i\right) + \mathbf{DB}^2(\eta)\left(\sum_{i=1}^{N} c_i\xi^{\lambda_i-1}\mathbf{\phi}_i\right) \tag{10}$$

Differentiating Eq. (9) with respect to time, the velocities and accelerations at any point in a subdomain are obtained as

$$\dot{\mathbf{u}}(\xi,\eta) = \mathbf{N}(\eta)\sum_{i=1}^{N} \dot{c}_i\xi^{\lambda_i}\mathbf{\phi}_i \tag{11}$$

$$\ddot{\mathbf{u}}(\xi,\eta) = \mathbf{N}(\eta)\sum_{i=1}^{N} \ddot{c}_i\xi^{\lambda_i}\mathbf{\phi}_i \tag{12}$$

The constants $\dot{c}_i$ and $\ddot{c}_i$ are calculated from the nodal velocities and accelerations on the subdomain boundary $\dot{\mathbf{u}}_\mathbf{b}$ and $\ddot{\mathbf{u}}_\mathbf{b}$ which are subsets of $\dot{\mathbf{U}}$ and $\ddot{\mathbf{U}}$, respectively, by differentiating Eq. (7) with respect to time

$$\dot{\mathbf{c}} = \mathbf{\Phi}^{-1}\dot{\mathbf{u}}_\mathbf{b} \tag{13}$$

$$\ddot{\mathbf{c}} = \mathbf{\Phi}^{-1}\ddot{\mathbf{u}}_\mathbf{b} \tag{14}$$

where $\dot{\mathbf{c}} = \{\dot{c}_1, \dot{c}_2, \ldots, \dot{c}_N\}^T$ and $\ddot{\mathbf{c}} = \{\ddot{c}_1, \ddot{c}_2, \ldots, \ddot{c}_N\}^T$.

Posteriori energy error estimator

For elastodynamic problems, the energy norm of the total energy can be estimated by

$$\|\mathbf{u}\| = \left(\|\mathbf{u}\|_k^2 + \|\mathbf{u}\|_s^2 \right)^{1/2} \tag{15}$$

where $\|\mathbf{u}\|_k$ and $\|\mathbf{u}\|_s$ are the energy norm of the kinetic energy and the strain energy respectively. *NS* is the number of subdomains.

Based on SBFEM, the strain energy norm can be estimated by [Zhang el al. (2011)]

$$\|\mathbf{u}\|_s \approx \left(\sum_{s=1}^{NS} \sum_{i=1}^{N} \sum_{i=j}^{N} \frac{c_i c_j}{\lambda_i + \lambda_j} \int_{S_s} \boldsymbol{\sigma}_i^*(\eta)^T \mathbf{D}^{-1} \boldsymbol{\sigma}_j^*(\eta) \mid J \mid d\eta \right)^{1/2} \tag{16}$$

where $\boldsymbol{\sigma}_i^*$ is the recovered stresses of *i*th mode at the boundary nodes, obtained by nodal average method for linear element here.

And the kinetic energy norm can also be evaluated semi-analytically

$$\|\mathbf{u}\|_k = \left(\sum_{s=1}^{NS} \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{\rho \dot{c}_i \dot{c}_j}{\lambda_i + \lambda_j + 2} \int_{S_s} \left(\dot{\mathbf{u}}_i(\eta) \right)^T \dot{\mathbf{u}}_j(\eta) |J| d\eta \right)^{1/2} \tag{17}$$

where $\dot{\mathbf{u}}_i(\eta)$ is the velocity vector of *i*th mode along the subdomain boundary.

Substituting Eqs. (16) and (17) into Eq. (15) yields

$$\|\mathbf{u}\| = \left(\sum_{s=1}^{NS} \sum_{i=1}^{N} \sum_{j=1}^{N} \left(\frac{\rho \dot{c}_i \dot{c}_j}{\lambda_i + \lambda_j + 2} \int_{S_s} \left(\dot{\mathbf{u}}_i(\eta) \right)^T \dot{\mathbf{u}}_i(\eta) |J| d\eta + \frac{c_i c_j}{\lambda_i + \lambda_j} \int_{S_s} \boldsymbol{\sigma}_i^*(\eta)^T \mathbf{D}^{-1} \boldsymbol{\sigma}_j^*(\eta) \mid J \mid d\eta \right) \right)^{1/2} \tag{18}$$

The domain energy error can be evaluated as

$$\|\mathbf{e}\| \approx \left(\sum_{s=1}^{NS} \|\mathbf{e}\|_s^2 \right)^{1/2} \tag{19}$$

where

$$\|\mathbf{e}\|_s \approx \sum_{i=1}^{N} \sum_{i=j}^{N} \frac{c_i c_j}{\lambda_i + \lambda_j} \int_{S_s} \mathbf{e}_{\sigma i}^*(\eta)^T \mathbf{D}^{-1} \mathbf{e}_{\sigma j}^*(\eta) \mid J \mid d\eta \tag{20}$$

is the energy error of a single subdomain. $\mathbf{e}_{\sigma i}^*(\eta)$ is the *i*th modal stress error on the boundary and calculated by

$$\mathbf{e}_{\sigma i}^*(\eta) = \mathbf{N}(\eta)\boldsymbol{\sigma}_i^* - \mathbf{D}(\lambda_i \mathbf{B}^1(\eta) + \mathbf{B}^2(\eta))\boldsymbol{\varphi}_i \tag{21}$$

The dynamic energy error estimator is defined as

$$\delta = \frac{\|\mathbf{e}\|}{\|\mathbf{u}\|} \times 100\% \tag{22}$$

Adaptive procedure

Remeshing

Assuming the optimized mesh is obtained when each subdomain contributes equally to the domain energy error. The average limit of the subdomain error is defined as

$$\left\| \mathbf{e} \right\|_s^{\mathrm{lim}} = \bar{\delta} \left(\frac{\left\| \mathbf{u} \right\|^2}{NS} \right)^{1/2} \qquad (23)$$

where $\bar{\delta}$ is the target error estimator of the domain.

A parameter θ is used to identify the subdomains need to be refined

$$\theta = \frac{\left\| \mathbf{e} \right\|_s}{\left\| \mathbf{e} \right\|_s^{\mathrm{lim}}} \qquad (24)$$

The polygon mesh of SBFEM is generated from a Delaunay triangular mesh by locating the scaling center at the common point of a patch of triangles and taking the centroids of these elements as the vertices of the subdomain. The readers are referred to [Ooi et al. (2012)] for details. In each time step, the following mesh refinement procedure is applied to all subdomains with $\theta>1$, as illustrated in Fig. 2:

(a) Triangular background mesh and polygon subdomain

(b) Refinement of triangular background mesh

(c) Regenerating a new polygon mesh

Figure 2. Remeshing procedure

(i) Locate the polygon subdomain(s) with exceeding error, i.e. $\theta>1$, in the old mesh (Fig. 2(a));
(ii) Find the corresponding triangles of the polygon that need to be refined in the triangular background mesh, and add midpoints on the triangles' boundaries, so that each triangle is split into four quarters (Fig. 2(b));
(iii) Regenerate a Delaunay triangular mesh and a new polygon mesh (Fig. 2(c)). Consequently, the size of the polygon subdomain is scaled to half after one time refinement.

Considering the number of subdomains is small, the time on seeking the polygon is short. Since the remeshing procedure is actually implemented in the triangular background mesh and the topology of SBFEM is generated directly, it is more convenient and efficient than the remeshing procedure [Zhang el al. (2011)] carried out in the SBFEM mesh.

Mesh mapping

Once a new polygon mesh is obtained, nodal state variables, including displacement, velocity and acceleration, need to be transferred from the old mesh to the new one as initial conditions of the following time step. In SBFEM, these variables at any point within a subdomain or on its boundaries are directly calculated by Eqs. (9), (11) and (12). Specifically, for a point located at coordinates (x_A, y_A) in the new mesh after remeshing, the polygon subdomain in the old mesh within which the point (x_A, y_A) is located is first found. The coordinates (x_A, y_A) are then easily transformed

to SBFEM coordinates (ξ_A, η_A) by Eqs. (1) and (2) in this polygon subdomain. The displacements, velocities and accelerations in the new mesh at Point A can be calculated as

$$\mathbf{u}_A = \mathbf{u}(\xi_A, \eta_A) = \mathbf{N}(\eta_A) \sum_{i=1}^{N} c_i \xi_A^{\lambda_i} \boldsymbol{\varphi}_i \quad (25)$$

$$\dot{\mathbf{u}}_A = \dot{\mathbf{u}}(\xi_A, \eta_A) = \mathbf{N}(\eta_A) \sum_{i=1}^{N} \dot{c}_i \xi_A^{\lambda_i} \boldsymbol{\varphi}_i \quad (26)$$

$$\ddot{\mathbf{u}}_A = \ddot{\mathbf{u}}(\xi_A, \eta_A) = \mathbf{N}(\eta_A) \sum_{i=1}^{N} \ddot{c}_i \xi_A^{\lambda_i} \boldsymbol{\varphi}_i \quad (27)$$

The flow chart

Fig. 3 illustrates the flow chart of the presented method. A parent mesh consisting of relatively large-sized polygon subdomains is generated from a Delaunay triangular mesh and a target error estimator $\overline{\delta}$ is input first. At time step n, the state variables $\mathbf{U}_n, \dot{\mathbf{U}}_n, \ddot{\mathbf{U}}_n$ are solved by the Newmark integration scheme using the old mesh at the end of time step (n-1) and the error estimator δ is calculated. If δ exceeds $\overline{\delta}$, the adaptive procedure is triggered and the triangular background mesh is refined and a new polygon mesh is regenerated. The nodal state variables are then mapped from the old polygon mesh to the new one as the initial conditions. This iteration is repeated until the target is satisfied.

Numerical example

An L-shaped domain subjected to a triangular blast loading was analyzed by the adaptive method. The dimensions and material properties with SI units are shown in Fig. 4. The dynamic responses were calculated in a time period of (0.0, 8.0s) with a constant time increment Δt=0.1s. The element size of the triangular background mesh is 25 and totally 23 polygon subdomains are generated, as shown in Fig. 5. The target error is set as 10%.

Fig. 6(a) ~ Fig. 6(f) describe the evolution process of adaptive meshes with horizontal stresses. At the beginning of loading, steep stress gradient areas appear around the left boundary of the domain and the polygon subdomains herein are refined, whereas large-sized subdomains are used in the right area with no stress (Fig. 6(a)). With the stress wave propagates, more and more polygon subdomains are refined (Fig. 6(b) and Fig. 6(c)). In Fig. 6(d), the stress wave rebounds from the right boundary
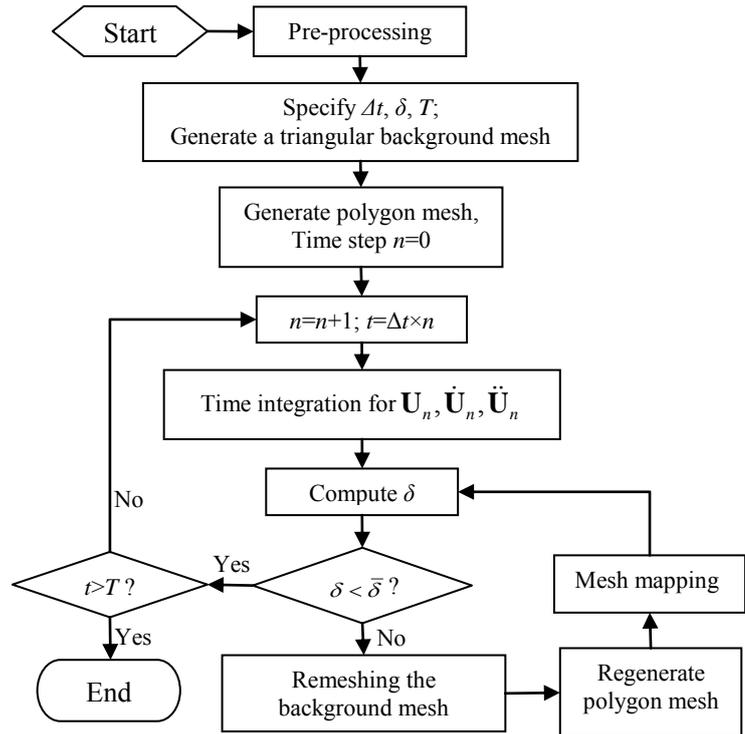
Figure 3. Flow chart of the presented method
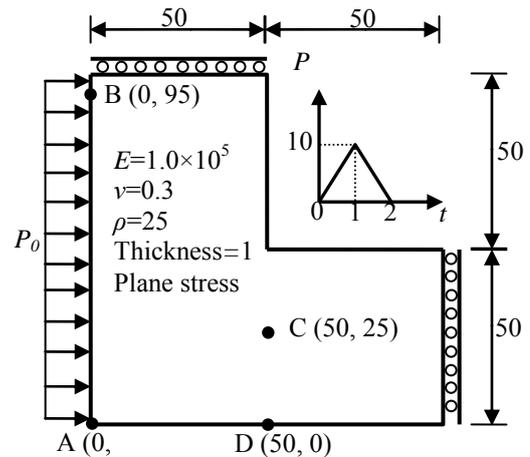
Figure 4. An L-shaped domain
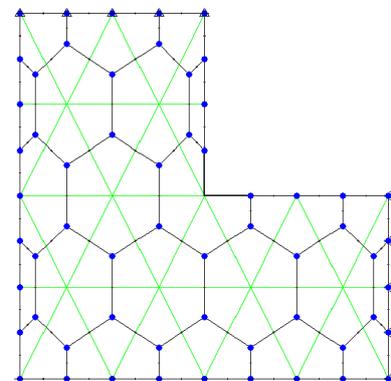
Figure 5. Parent mesh (mesh 1, DOFs=278)

and polygon subdomains near the bottom begins to be refined. Afterwards, the stress distribution becomes complicated and the steep stress gradient areas are mainly around the re-entrant corner, which has strong singularity, and very fine mesh are used, as shown in Fig. 6(e) and Fig. 6(f).

(a). t=0.4s, DOFs=562

(b). t=1.2s, DOFs=572

(c). t=2.0s, DOFs=630

(d). t=2.8s, DOFs=692

(e). t=4.0s, DOFs=1350

(f). t=7.0s, DOFs=1610
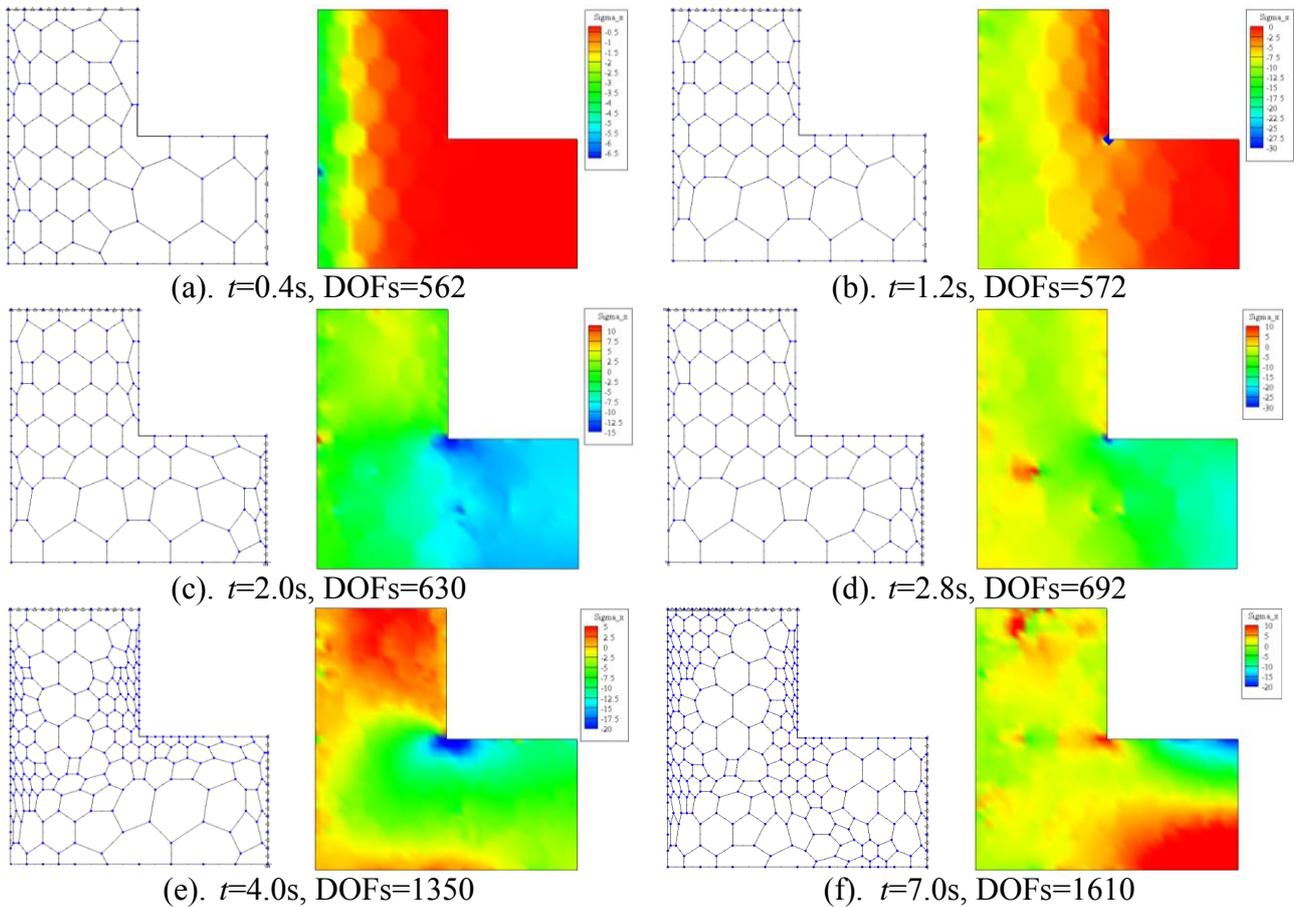
Figure 6. Adaptive meshes and evolution of horizontal stress contours

The structural dynamic responses are given out in Fig. (a) ~ Fig. (d). The results are calculated by PSBFEM based on an uniform fine mesh as shown in Fig. 7 (PSBFEM, mesh 1), an uniform coarse mesh as shown in Fig. 5 (PSBFEM, mesh 2), the presented method (APSBFEM) and AFEM proposed by [Zeng and Wiberg (1992)], respectively. It can be seen that not only displacements but also stresses of the presented method agree better with the results of AFEM than the results of PSBFEM based on uniform coarse mesh.

Fig. 12 records the energy error of PSBFEM using coarse mesh and the presented method. It can be seen that the value of energy error fluctuates between 6%-16% based on PSBFEM, while the error is limited under the target of 10%. Fig. 13 records the degrees of freedom (DOFs) used in the adaptive meshes.
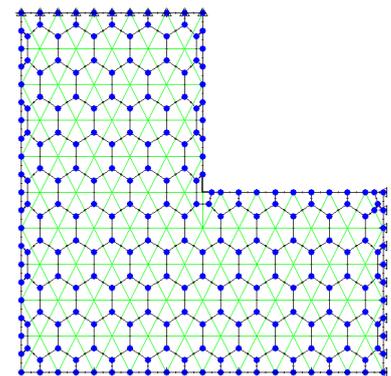
Figure 7. Fine mesh (mesh 2, DOFs=1172)

Conclusions

An adaptive method based on polygon scaled boundary finite elements for general elastodynamic problems is developed in this study. The original and adaptive polygon meshes are generated from a triangular background mesh which is created by the Delaunay algorithm, thus the presented method

is suitable for problems with complex boundaries and cracks. The refinement is actually conducted in the background mesh so that the remeshing procedure is very convenient and straightforward. The semi-analytical energy error estimator and the simple mesh mapping algorithm endows the presented method with a good precision. It is expected to extend this method to more complicated problems, such as crack propagation.
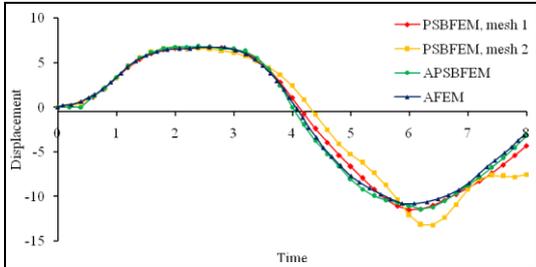
Figure 8. Horizontal displacement at A
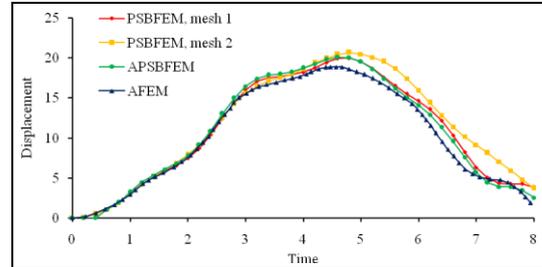
Figure 9. Horizontal displacement at B
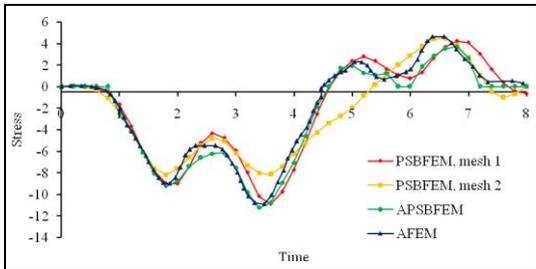
Figure 10. Horizontal normal stress at C
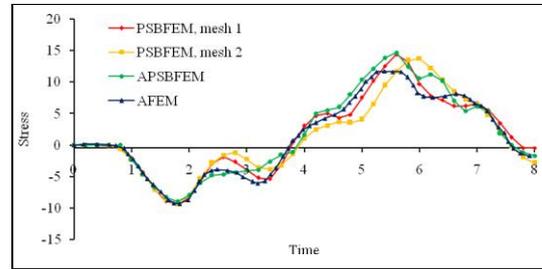
Figure 11. Horizontal normal stress at D
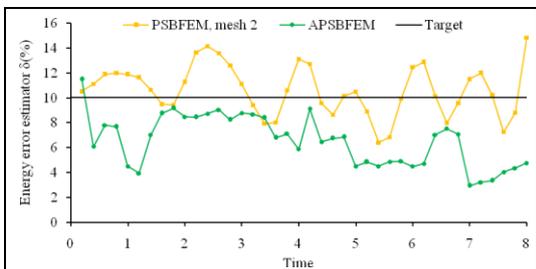
Figure 12. History of the energy error estimator
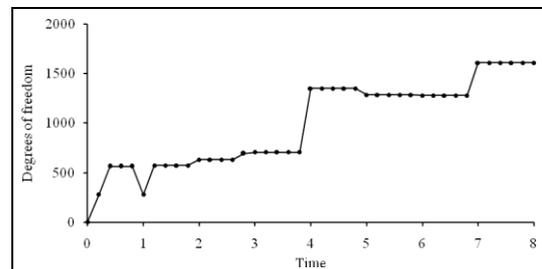
Figure 13. History of degrees of freedom

References

Ooi, E. T., Shi, M., Song, C. M., Tin-Loi, F., and Yang, Z. J. (2013) Dynamic crack propagation simulation with scaled boundary polygon elements and automatic remeshing technique. *Engineering Fracture Mechanics*, **106**, 1-21.

Ooi, E. T., Song, C. M., Tin‑Loi, F., and Yang, Z. J. (2012) Polygon scaled boundary finite elements for crack propagation modelling. *International Journal for Numerical Methods in Engineering*, **91**, 319-342.

Song, C. M. and Wolf, J. P. (1997) The scaled boundary finite-element method—alias consistent infinitesimal finite-element cell method—for elastodynamics, *Computer Methods in applied mechanics and engineering* **3**, 329-355.

Yang, Z. J., Zhang, Z. H., Liu, G. H., and Ooi, E. T. (2011). An *h*-hierarchical adaptive scaled boundary finite element method for elastodynamics. *Computers and Structures* **89**, 1417-1429.

Zeng, L. F. and Wiberg, N. –E. (1992) Spatial mesh adaptation in semidiscrete finite element analysis of linear elastodynamic problems, *Computational Mechanics* **9**, 315-332.

Zhang, Z. H., Yang, Z. J., Liu, G. H., and Hu, Y. J. (2011) An adaptive scaled boundary finite element method by subdividing subdomains for elastodynamic problems. Science China Technological Sciences **54**, 101-110.

Ballistic Impacts of a Full-Metal Jacketed (FMJ) Bullet on a Validated Finite Element (FE) Model of Helmet-Cushion-Head

* Kwong Ming Tse[1], Long Bin Tan[1], Bin Yang[2], Vincent Beng Chye Tan[1], Siak Piang Lim[1, 3], * Heow Pueh Lee[1, 3]

[1] Department of Mechanical Engineering, National University of Singapore
9 Engineering Drive 1, Singapore 117576
[2] College of Automobile and Traffic Engineering, Nanjing Forestry University
159 LongPan Rd, Nanjing, People's Republic of China 210037
[3] National University of Singapore (Suzhou) Research Institute
377 Lin Quan Street, Suzhou Industrial Park, Jiang Su
People's Republic of China 215123

*Corresponding author: tsekm@nus.edu.sg or tsekm.research@yahoo.com (Tse, K.M.);
mpeleehp@nus.edu.sg (Lee, H.P.)

Abstract

In order to determine the severity of the head injuries sustained from ballistic impact orientation and to investigate the effectiveness of the cushioned combat helmet in protecting the head from ballistic impact, series of ballistic impact simulations (frontal, lateral, rear and top) of FMJ bullet on a subject-specific FE head model, which are based on National Institute of Justice (NIJ) test standard. Two different designs of helmet interior cushion, namely the strap-netting system and the Oregon Aero (OA) foam, are adopted in this study. In general, the head experiences highest G in rear impacts among all impact orientations. The FE simulations also show that the use of OA foams helps to reduce frontal impact G forces and thus offers better protection from all various impact orientations. The OA cellular foams are more effective in limiting the transmission of force by being able to absorb more energy, via plateau characteristic prior to foam densification, compared to the stiffer linear elastic front cushion of strap-netted helmet. The simulations also showed both the helmets passed the NIJ requirement, WSTC and FMVSS criterion.

Keywords: Ballistic impact, helmet, subject-specific head model, head injury, cushion, head acceleration

1. Introduction

Advanced combat helmets (ACH) protects military personnel from sustaining traumatic brain injuries (TBI) due to blunt and ballistic impacts in both peace and war times. It is particularly important that these helmets are capable and effective in their function. Since 1970s, tremendous efforts had been spent on research of head protective helmets using finite element method (FEM) which serves as a cost-effective alternative to experiments [Khalil (1973; 1974); Van Hoof et al. (1999; 2001); Baumgartner and Willinger (2003); Aare and Kleiven (2007); Tham et al. (2008); Lee and Gong (2010); Yang and Dai (2010)] . For example, Khalil et al. (1974) performed low-velocity ballistic impacts using a very simplified head-helmet finite element (FE) model and validated against corresponding experiments. Van Hoof et al. (1999, 2001) found that the helmet interior exhibited large deformation exceeding the gap between inner helmet shell and head in experimental and numerical ballistic impact studies. Another study simulating ballistic impact on military helmet was by Baumgartner and Willinger (2003) who predicted skull fracture without traumatic brain injury (TBI). More recently, Aare and Kleiven (2007) studied the effects of helmet shell stiffness and impact orientation on a FE head model during a ballistic impact, while Yang and Dai (2010) focused on evaluation of the rear effect with different impact orientation. Of late, Tan et al. (2012) had performed both experimental tests and FE simulations on helmeted Hybrid III headform using

spherical projectile and found that foam cushioning system would help to reduce the head acceleration.

In order to determine the severity of the head injuries sustained from ballistic impact and to investigate the effectiveness of the cushioned combat helmet in protecting the head from ballistic impact, series of ballistic impact simulations (frontal, lateral, rear and top) of FMJ bullet on a subject-specific FEHM, which are based on National Institute of Justice (NIJ) test standard, were performed for a duration of 4ms using the explicit code in Abaqus v6.10 (SIMULIA, RI, USA). Similar to Tan et al. (2012), the interior cushioning systems included in this current study were namely strap-netting system (in Helmet 1) and Oregon Aero (OA) interior foam cushioning system (in Helmet 2). It should be noted that this subject-specific FEHM used in this study has been validated against the ICP and relative displacement data of three cadaveric experiments [Tse et al. (2014)].

2. Methods and Materials

2.1 Model Development and Model Description

The subject-specific FE model of human head and brain was reconstructed from computed tomography (CT) and magnetic resonance imaging (MRI) images. More details on the head model could be found in Tse et al. (2014)'s study. As for the advanced combat helmet (ACH) model, it was reconstructed from axial CT images while the two interior cushioning systems (OA foam and strap-netting) were drawn from scratch. The 9mm full-metal jacketed (FMJ) bullet, which geometrical details could be found in Tham et al. (2008)'s study, was used in the NIJ ballistic simulations. The entire assembly of the helmet-cushion-head model was shown in Figure 1. It should be noted that a preloading step was implemented prior to the actual ballistic impact step so that the two interior cushioning systems fit well with both the head and helmet models.
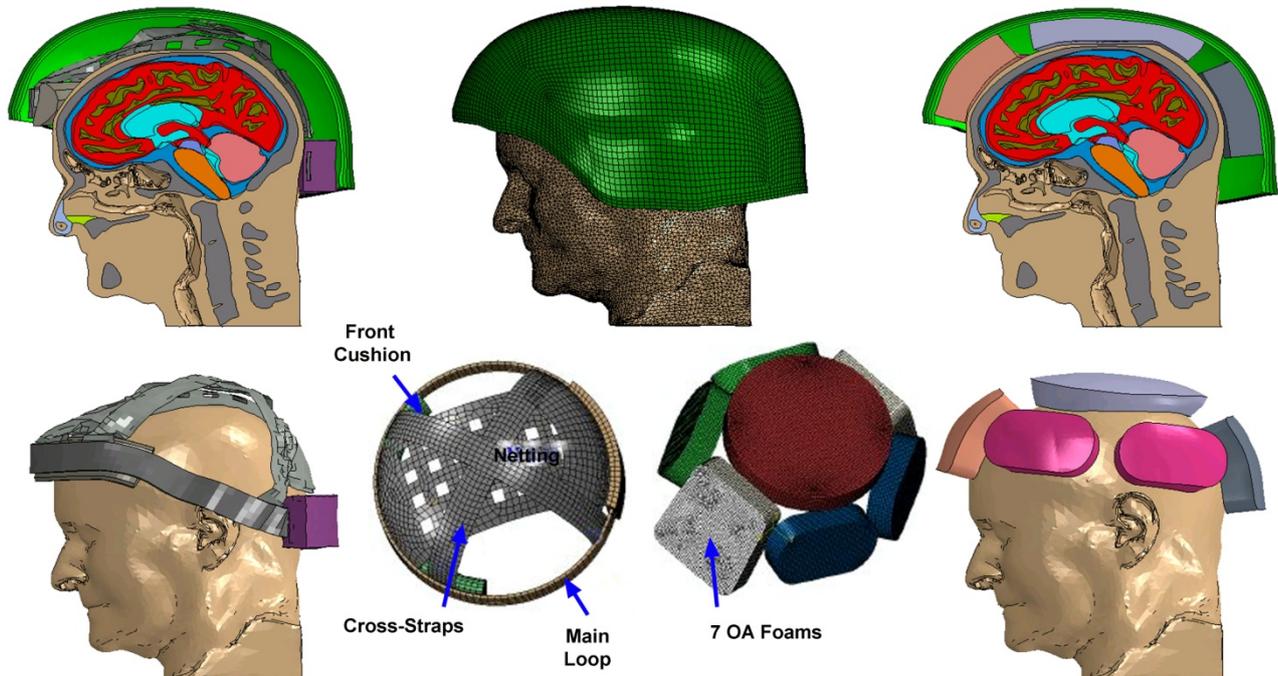
Figure 1: The two configurations of the helmet-interior cuhsion-head assembly.

2.2 Material Properties

For the head model, all the skeletal and cartilaginous tissues were modeled as linear elastic, isotropic materials, while the brain tissues were modeled with viscoelastic material properties [Tse et al. (2014)]. The helmet laminates adopted linear elastic but anisotropic material properties [Tan et

al. (2012)]. As for the two interior cushioning systems, their material properties were obtained from the in-house experiments in our previous work [Tan et al. (2012)]. The FMJ bullet, which is made of brass and lead, has its mechanical properties shown in Table 1.

Table 1: Material properties of both the intracranial and extracranial components used in the models.

Components		Material Properties	Poisson's ratio, υ	Density, ρ (kg/mm^3)
		Young's Modulus, E (MPa) / Shear Modulus, G (MPa) $G(t) = G_\infty + (G_0 - G_\infty)e^{-\beta t}$		
Head	Brainstem	$G_0 = 0.0225$ MPa, $G_\infty = 0.0045$ MPa, $\beta = 80$ s^{-1}	0.4996	1.06E-06
	Cerebral Peduncle	$G_0 = 0.0225$ MPa, $G_\infty = 0.0045$ MPa, $\beta = 80$ s^{-1}	0.4996	1.06E-06
	Cerebrum	$G_0 = 0.528$ MPa, $G_\infty = 0.168$ MPa, $\beta = 35$ s^{-1}	0.48	1.14E-06
	Cerebellum	$G_0 = 0.528$ MPa, $G_\infty = 0.168$ MPa, $\beta = 35$ s^{-1}	0.48	1.14E-06
	CSF	$E = 1.314$	0.4999	1.04E-06
	Gray Matter	$G_0 = 0.034$ MPa, $G_\infty = 0.0064$ MPa, $\beta = 700$ s^{-1}	0.4996	1.04E-06
	Lateral Cartilage	$E = 30$	0.45	1.50E-06
	Septum Cartilage	$E = 9$	0.32	1.50E-06
	Bone	$E = 8000$	0.22	1.21E-06
	Soft Tissues	$E = 16.7$	0.46	1.04E-06
	Tooth	$E = 2070$	0.3	2.25E-06
	Ventricles	$E = 1.314$	0.4999	1.04E-06
	White Matter	$G_0 = 0.041$ MPa, $G_\infty = 0.0078$ MPa, $\beta = 700$ s^{-1}	0.4996	1.04E-06

		E_{11} (MPa)	E_{22} (MPa)	E_{33} (MPa)	G_{12} (MPa)	G_{13}/G_{23} (MPa)	υ_{12}	$\upsilon_{13}/\upsilon_{23}$	ρ (kg/mm^3)
ACH	Helmet Shells	18000	18000	4500	770	2600	0.25	0.33	1230

		Young's Modulus / Shear Modulus		Poisson's ratio	Density
Interior Cushion Systems	Cross Straps (Helmet 1)	$E = 60$		0.25	400
	Front Cushion (Helmet 1)	$E = 18$		0.25	200
	Main Loop (Helmet 1)	$E = 60$		0.25	400
	Netting (Helmet 1)	$E = 60$		0.25	400
	Rear Cushion (Helmet 1)	$E = 18$		0.25	200
	OA Foams (Helmet 2)	Direct compression data from experiment			164
Projectile	Cartridge Brass	$E = 110000$		0.375	8520
	Lead Core	$G = 200$		-	11840

2.3 Failure Modeling of Helmet and FMJ Bullet

This preliminary study modeled both the helmet property degradation and the inter-laminar failure using surface traction criteria [Tan et al. (2012)]. Additionally, the Hashin Fabric Criterion was used to model the fabric-reinforced aramid laminates of the helmet shell which takes the bi-directional strength of the fibers into account [Tan et al. (2012)]. As for the FMJ bullet, Johnson Cook plasticity hardening and damage initiation criterion was used to model the exterior cartridge brass material [Johnson and Cook (1983)], whilst the Mie-Grüneisen hydrodynamic equation of state material model was used to model the lead core [Abaqus (2013)] (Table 2).

Table 2: Material constants in failure modeling of FMJ bullet.

Parts	Material Constants For Failure Modeling							
	Constants in Johnson-Cook Strain Rate Hardening							
	A	**B**	**n**	**M**	T_m **(K)**	T_{trans} **(K)**	**C**	ε_0 **(s⁻¹)**
Cartridge Brass	112	505	0.42	1.68	1189	373	0.009	1
	Constants in Johnson-Cook Damage Initiation Criterion							
	d_1	d_2	d_3	d_4	d_5	T_m **(K)**	T_{trans} **(K)**	ε_0 **(s⁻¹)**
	0.54	4.89	3.03	0.014	1.12	1189	373	1
Lead Core	**Constants in Mie-Grüneisen hydrodynamic equation of state** **(Linear Us-Up Hugoniot form)**						**Specific Heat Capacity** $(J \cdot kg^{-1} \cdot K^{-1})$	
	c_0 **(cm/µs)**		**s**		Γ_0		150	
	0.2006		1.429		2.60			

2.4 Boundary Conditions

All the contact conditions imposed between the intracranial interfaces were taken from Tse et al. (2014). Table 3 shows the required boundary conditions of NIJ-STD-0106.00 that were applied at the base of the helmet-cushion-head assembly except for top impact which could be treated as if the military personnel proning on the ground while the fragment or bullet hits at the top of the helmet. As for initial condition, an initial velocity of 358m·s⁻¹ was prescribed to the entire FMJ bullet [Aare and Kleiven (2007)], for each of the impact orientation and helmet cushions configuration.

Table 3: Boundary conditions for the NIJ ballistic impact simulations.

Impact Orientation	Displacement Constraints at the Base of the Helmet-Cushion-Head Assembly
Front	$U_2=0$; $U_3=0$; $\theta_1=0$; $\theta_3=0$
Side	$U_1=0$; $U_3=0$; $\theta_2=0$; $\theta_3=0$
Rear	$U_2=0$; $U_3=0$; $\theta_1=0$; $\theta_3=0$
Top	$U_1=0$; $U_2=0$; $U_3=0$; $\theta_3=0$

3. Results and Discussion

Table 4 showed the maximum values of the helmet strain, dynamic deflection as well as depth of helmet dent. The impact energy was partially absorbed by the helmet through deflection and deformation of the helmet shells, whilst majority of it was absorbed by the interior cushions. It could be seen in Table 4 that Helmet 1 generally deflected more than Helmet 2, except for lateral impacts. However, it should be noted that the projectile stroke on the rim of Helmet 1 with the presence of the underlying interior main loop preventing subsequent deflection, unlike the lateral impact of Helmet 2.

Table 4: Helmet parameters of the head with the two helmet configurations.

	Helmet 1 (with strap-netting)							
	Front	**Locations**	**Side**	**Locations**	**Rear**	**Locations**	**Top**	**Locations**
Max. Helmet Strain	0.195	Impact Site	0.164	Impact Site	0.039	Impact Site	0.120	Top Left of Helmet
Max. Dynamic Deflection (mm)	7.121	Right Helmet Rim	13.261	Right Helmet Rim	9.466	Left Helmet Rim	24.955	Left Helmet Rim
Max. Depth of Helmet Dent (mm)	10.452	Impact Site	12.597	Impact Site	12.628	Impact Site	13.624	Impact Site
Contact Between Helmet Shell & Head	No	-	No	-	No	-	No	-
	Helmet 2 (with OA foam padding)							
	Front	**Locations**	**Side**	**Locations**	**Rear**	**Locations**	**Top**	**Locations**
Max. Helmet Strain	0.074	Impact Site	0.105	Impact Site (2nd Outermost Layer)	0.037	Impact Site	0.0482	Posterior Top Right of Helmet
Max. Dynamic Deflection (mm)	5.911	Impact Site	28.569	Impact Site	6.665	Left Helmet Rim	16.846	Rear Helmet Rim
Max. Depth of Helmet Dent (mm)	10.817	Impact Site	15.185	Impact Site	9.665	Impact Site	12.566	Impact Site
Contact Between Helmet Shell & Head	No	-	No	-	No	-	No	-

Figure 2 showed the impact sequences of all the various orientations on the two helmets. Similar phenomenon had been observed in both helmets, as shown in Figure 2. It was noted that, in all eight impact orientation, there was no penetration of FMJ projectile into the helmets. This indicated that both the helmets had successfully deflected all the FMJ projectiles travelling at the speed of 358m·s⁻¹ and met the NIJ requirement. In general, permanent dents of 9-15mm on the helmet exterior were observed at various sites of impact. The "crater" or spatial extent of the impression is around 60mm in diameter, and a bulge could be seen at the backplane of the helmet.
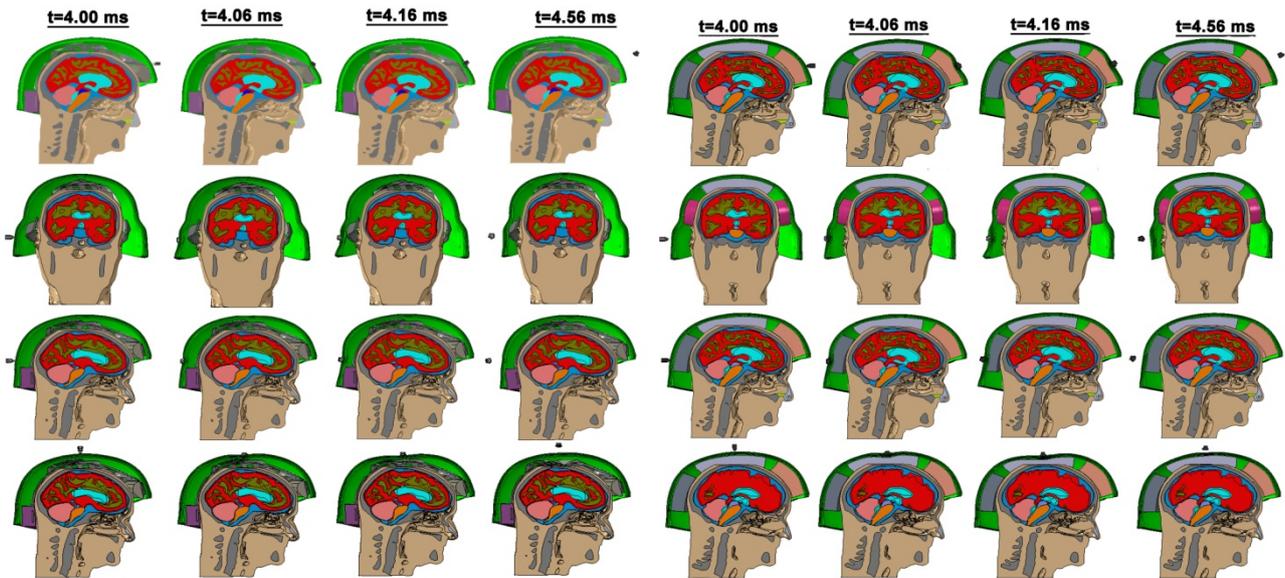
Figure 2: Impact sequence of various impact orientations for Helmet 1 (Left) and Helmet 2 (Right).

In this preliminary study, acceleration at the centre of the head was chosen as the parameter for analyses and for gauging the severity of TBI sustained (Figure 3). It was noted that the peak acceleration was found to be relatively high for both front and rear impacts for Helmet 1 (strap-netting) but were significantly reduced when equipped with Helmet 2 (OA foam). As for the remaining impact orientations (lateral and top impacts), Helmet 2 did not help much in reducing the peak head acceleration. Nevertheless, there was a general reduction in peak skull stresses in Helmet 2 (with OA foam padding) as compared to Helmet 1 (with strap-netting), with the percentage of reduction up to 44.94%, 0.07%, 109.21% and 8.39% for frontal, lateral, rear and top impacts. This showed that OA cellular foams were more effective as interior cushions as they limited the transmission of force by being able to absorb more energy, via plateau characteristic prior to foam densification, compared to the stiffer linear elastic front cushion of Helmet 1.
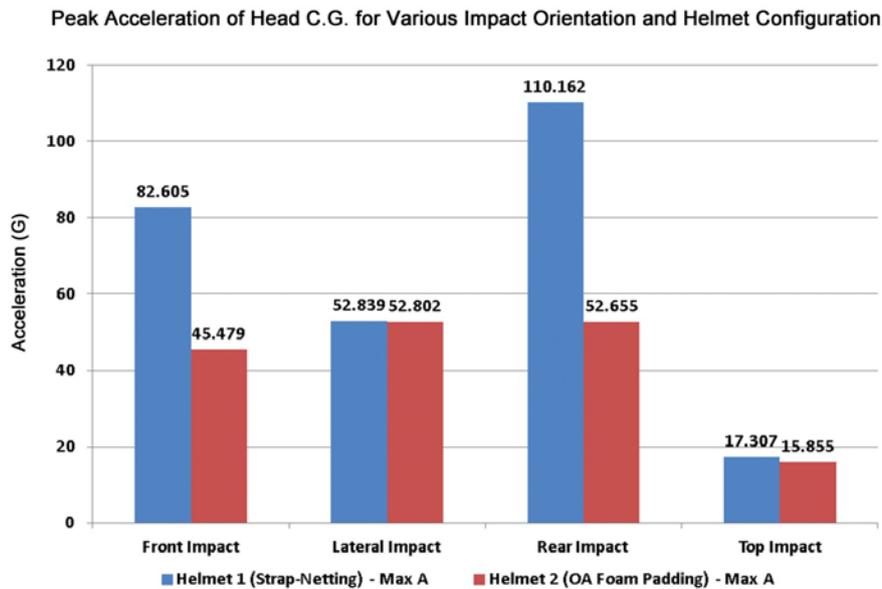
Figure 3: Peak acceleration at the C.G. of the head for various impact orientation and helmet liner configuration.

Comparison of the head acceleration obtained for various impact orientation indicated that the rear impact resulted in highest acceleration value of up to 110G when equipped with Helmet 1 (with strap-netting), followed by frontal and lateral impacts, whilst the lateral impact were most severe for Helmet 2 (with OA foam padding), followed by rear and front impacts. Top impacts were the least severe among all the impact orientation due to the nature of the boundary condition for the military personnel in the prone position. The peak acceleration values obtained from the simulations were also compared with the established injury criteria such as Wayne State Tolerance Curve (WSTC) and Federal Motor Vehicles Safety Standards (FMVSS) 218 criterion (Figure 4). It could be concluded from Figure 4 that all the impacts with both helmets passed the criteria of the WSTC and FMVSS.
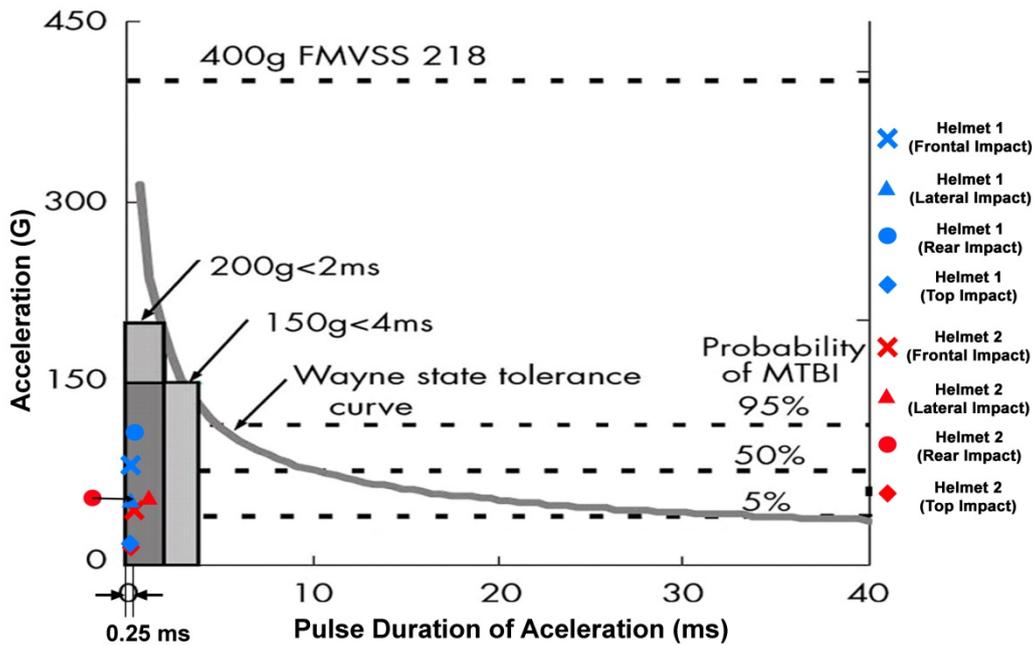
Figure 4: Acceleration responses for the helmets with two interior cushion designs in various impact directions, in relation to other published criteria Modified from [Shewchenko et al. (2005)].

4. Conclusion

In this study, ballistic analysis using FEM had been carried out to evaluate the performance of the ACH as well as the effectiveness of its interior cushioning systems, in protect both military personnel and civilians from traumatic head injury. Rear impacts gave rise to highest head acceleration while the top impacts were the least severe among all the impact orientation. The use of OA foams helped to reduce impact G forces and thus offered better protection from all various impact orientations. The simulations also showed both the helmets passed the NIJ requirement, WSTC and FMVSS criterion. However, it is still too early to arrive at any concrete conclusion for the severity of impact orientation since the human tolerance for different impact orientation was different [Allsop and Kennett (2002)] and the probability curves were based on automotive safety standards. More investigations on criterion for ballistic impact would be needed in the future.

Conflict of interest
None.

Acknowledgements

References

Aare, M. and Kleiven, S. (2007). Evaluation of head response to ballistic helmet impacts using the finite element method, *International Journal of Impact Engineering* **34** (3), 596-608.

Abaqus Analysis User's Manual Version 6.13. Dassault Systèmes Simulia Corp.

Allsop, D. and Kennett, K. (2002). *Skull and facial bone trauma*. In: Nahum, A.M., Melvin, J. (Eds.), *Accidental Injury: Biomechanics And Prevention*. 2nd Edition. Springer-Verlag, pp. 577.

Baumgartner, D. and Willinger, R. (2003). Finite element modelling of human head injuries caused by ballistic projectiles. In: *Procceeding of NATO Research & Technology Organisation (RTO) Specialist Meeting*, Koblenz, Germany, 2003

Johnson, G.R. and Cook, W.H. (1983). A constitutive model and data for metals subjected to large strains, high strain rates and high temperature. In: *Procceeding of the 7th International Symposium on Ballistics*, 1983, 541-547.

Khalil, T.B. (1973). *Impact on a model head-helmet system*. University of California.

Khalil, T.B., Goldsmith, W. and Sackman, J.L. (1974). Impact on a model head-helmet system. *International Journal of Mechanical Sciences* **16** (9), 609-625.

Lee, H.P. and Gong, S.W. (2010). Finite element analysis for the evaluation of protective functions of helmets against ballistic impact. *Computer Methods in Biomechanics and Biomedical Engineering* **13** (5), 537-550.

National Institute of Justice (NIJ) Standard, NILECJ-STD-0106.00: Ballistic Helmets. 1975. ed. Justice, U.S.D.o. (Washington, DC, US).

Shewchenko, N., Withnall, C., Keown, M., Gittens, R. and Dvorak, J. (2005). Heading in football. Part 1: development of biomechanical methods to investigate head response. *Br J Sports Med* **39** Suppl 1, i10-25.

Tan, L.B., Tse, K.M., Lee, H.P., Tan, V.B.C. and Lim, S.P. (2012). Performance of an advanced combat helmet with different interior cushioning systems in ballistic impact: Experiments and finite element simulations. *International Journal of Impact Engineering* **50** (0), 99-112.

Tham, C.Y., Tan, V.B.C. and Lee, H.P. (2008). Ballistic impact of a KEVLAR® helmet: Experiment and simulations. *International Journal of Impact Engineering* **35** (5), 304-318.

Tse, K.M., Tan, L.B., Lee, S.J., Lim, S.P. and Lee, H.P. (2014). Development and validation of two subject-specific finite element models of human head against three cadaveric experiments. *International Journal for Numerical Methods in Biomedical Engineering* **30** (3), 397-415.

Van Hoof, J., Cronin, D.S., Worswick, M.J., Williams, K.V. and Nandlall, D. (2001). Numerical head and composite helmet models to predict blunt trauma. In: *Procceeding of 19th International Symposium on Ballistics, Interlaken, Switzerland*, 7-11 May 2001 2001

Van Hoof, J., Deutekom, M.J., Worswick, M.J. and Bolduc, M. (1999). Experimental and numerical analysis of the ballistic response of composite helmet materials. In: *Procceeding of 18th International Symposium on Ballistics*, San Antonio, TX, USA, 15-19 November 1999 1999

Yang, J. and Dai, J. (2010). Simulation-based assessment of rear effect to ballistic helmet impact. *Computer-Aided Design and Applications* **7** (1), 59-73.

Eliminating the Pressure-Velocity Coupling from the Incompressible Navier-Stokes Equations Using Integral Transforms

Daniel J. N. M. Chalhub, Leandro A. Sphaier, and Leonardo S. de B. Alves[*]

Laboratory of Theoretical and Applied Mechanics - LMTA
Department of Mechanical Engineering - PGMEC
Universidade Federal Fluminense
Niteroi, Rio de Janeiro, Brazil

*Corresponding author: leonardo.alves@mec.uff.br

Abstract

The present work proposes two methodologies using the Integral Transform Technique to solve the Poisson equation arising from the incompressible Navier-Stokes equations. The solution of this Poisson equation is very common in the formulations based on pressure-correction and are the main bottleneck of these approaches. The new formulation proposed in this work will allow the elimination of the pressure-velocity decomposition and also eliminate the sub-iterations of the usual pressure-correction methods. The results show a comparison in performance of both proposed approaches.

Keywords: Incompressible Navier-Stokes, Pressure Correction, Integral Transformation, Poisson Equation

x, y, z	Classical cartesian coordinates
L	Domain dimension in x direction
H	Domain dimension in y direction
t	Time
v	Vectorial velocity
u	Velocity component in x direction
v	Velocity component in y direction
f	Vectorial body force (in acceleration dimensions)
f_x	Body force component in x direction
f_y	Body force component in y direction
p	Pressure
$\bar{p}$	Transformed pressure for CITT solution
ρ	Fluid density
μ	Dynamic viscosity
ν	Kinematic viscosity
n	Index for CITT
n_{max}	Truncation order for CITT
i, q, p	Indices for the position in mesh for x direction
i_{max}	Maximum mesh divisions in x direction
j, r, s	Indices for the position in mesh for y direction
j_{max}	Maximum mesh divisions in y direction
l	Index for the time step
X_n	Eigenfunctions for the CITT solution
λ_n	Eigenvalues for the CITT solution

N_n	Norms of the eigenvalue problem
y', y''	Dummy integration variables
c_1, c_2	Integration constants
V	Domain volume
Λ	General function
ϵ	Absolute error

Introduction

In numerical simulations of incompressible flows, the main difficulty is the velocity and pressure coupling by the incompressibility constraint. The projection methods were developed to overcome this problem. These methods can be classified into three classes [Germond et al (2006)], namely pressure-correction methods, velocity-correction methods, and consistent splitting methods.
The most attractive feature of projection methods is that, at each time step, one only needs to solve a sequence of decoupled elliptic equations for the velocity and the pressure, making it very efficient for large scale numerical simulations. Although projection methods are widely used, many authors already drew attention to the fact that the decomposition used in this method is intrinsically second order accurate [Munz et al. (2003), Guermond et al. (2006)], preventing any approximation order higher than this. Pressure-correction schemes are time-marching techniques composed of two sub-steps for each time step: the pressure is treated explicitly or ignored in the first sub-step and is corrected in the second one. The linear momentum equations play the major role in determining the velocity components. Thus, it is left to the continuity equation to determine pressure, even if this variable does not appear explicitly in this equation. The most common methodology to determine an equation for the pressure-correction is to combine both equations by taking the divergence of the momentum equations and substituting the continuity equation where necessary, effectively generating a Poisson type equation for pressure. This makes it possible to obtain an equation for the pressure-correction, using the continuity equation. At this point, it is worth to highlight that the pressure-correction method is an iterative strategy which generate more accurate values at each iteration. The pressure-correction equation is an extrapolation to improve mass conservation at each iteration.

This procedure requires sub-iteration per time step, which is the major computational cost because at each sub-iteration, a Poisson equation for pressure must be solved. One could use Multigrid for the solution of the poisson equation to speed up the process, however, the sub-iteration are still required.

In the realm of analytical methods, the Integral Transform Technique, also known as the Classical Integral Transform Technique (CITT) [Mikhailov and Ozisik (1984)], has been playing a big role. It deals with expansions of the sought solution in terms of infinite orthogonal basis of eigenfunctions, keeping the solution process always within a continuous domain. The resulting system is generally composed of a set of uncoupled differential equations which can be solved analytically. However, a truncation error is involved since the infinite series must be truncated to obtain numerical results. This error decreases as the number of summation terms (truncation order) is increased, and the solution converges to a final value. Due to the series representation nature of the Integral Transform Technique, the estimated error can be easily obtained, which results in better global error control of the solution. The disadvantage associated with this approach is the need of a more elaborate analytical manipulation. This effort can be greatly minimized with the use of symbolic computation [Wolfram (2003)].

The present work proposes two methodologies using the Integral Transform Technique to solve the Poisson equation arising from the incompressible Navier-Stokes equations. The two proposed methodologies are: The single transformation and the double transformation. The new proposed formulation will allow the elimination of the pressure-velocity decomposition and also eliminate the sub-iterations of the usual pressure-correction methods.

Just a few works were concerned in mixing the Integral Transform Technique with other discrete schemes. Among these works, one could mention the work [Chalhub et al (2013)] that a new methodology for solving unsteady convective heat transfer problems via the generalized integral

transform technique was developed. The proposed scheme was based on writing the unknown potential in term of eigenfunction expansions; however, rather than transforming advection terms, an upwind approximation is used prior to the integral transformation. In the works [Guedes et al. (1994a; 1994b)], the authors analyzed the unsteady forced convection in laminar flow between parallel plates. This problem is solved using a hybrid scheme that combines the Generalized Integral Transform Technique with second-order finite differences. At [Cotta and Gerk (1994)], the integral transform method is employed in conjunction with second-order-accurate explicit finite-differences schemes, to handle accurately a class of parabolic-hyperbolic problems. In the work [Castelloes and Cotta (2006)], the solution is obtained using partial integral transformation strategy to solve the problem and the work of [Naveira et al. (2009)] showed a hybrid numeric-analytical solution for unsteady forced laminar convection between parallel plates.

Problem Formulation

In order to solve a fluid flow problem, the conservation laws of physics are needed: the mass conservation, also known as the continuity equation and the momentum conservation. In addition to these equations, a constitutive equation is required, which the Newton's law of viscosity will be used. The flow is also considered to be incompressible, in other words, the fluid has a constant density.

Combining these equations [Kundu (1990)], one can arrive at the incompressible Navier-Stokes equations[1]:

$$\frac{\partial v}{\partial t} + v \cdot \nabla v = -\frac{1}{\rho}\nabla p + \nu \nabla^2 v + f \quad \text{for} \quad x \in V \quad \text{and} \quad t \ge 0 \tag{1}$$

$$\nabla \cdot v = 0 \quad \text{for} \quad x \in V \tag{2}$$

in which equation (1) is the momentum conservation equation and equation (2) is the mass conservation equation, also called the incompressibility constraint.

Based on the projection methods for incompressible flows [Guermond et al. (2006)], to obtain the Poisson equation for pressure, one should apply the divergence operator on equation (1) and use the continuity equation (2):

$$\frac{1}{\rho}\nabla^2 p = \nabla \cdot f - \nabla v : (\nabla v)^{\mathrm{T}} \quad \text{for} \quad x \in V \tag{3}$$

where ρ is the fluid density, f is the body force vector, v is the velocity vector, p is the pressure and V is a general domain volume.

Then, the continuity could be replaced at the Navier-Stokes equations (1) by the Poisson equation (3), resulting in the following system to be solved:

$$\frac{\partial v}{\partial t} + v \cdot \nabla v = -\frac{1}{\rho}\nabla p + \nu \nabla^2 v + f \quad \text{for} \quad x \in V \quad \text{and} \quad t \ge 0 \tag{4}$$

$$\frac{1}{\rho}\nabla^2 p = \nabla \cdot f - \nabla v : (\nabla v)^{\mathrm{T}} \quad \text{for} \quad x \in V \tag{5}$$

In this work, normal zero gradients for pressure at the boundaries will be used.

$$(\nabla p \cdot n)_{\partial V} = 0 \tag{6}$$

where ∂V is the boundary of the general domain volume.

The problem can be simplified for cartesian domain:

$$\frac{\partial^2 p(x,y,t)}{\partial x^2} + \frac{\partial^2 p(x,y,t)}{\partial y^2} = \rho[h(x,y,t) - g(x,y,t)] \tag{7}$$

$$\left(\frac{\partial p(x,y,t)}{\partial y}\right)_{y=0} = 0 \quad \left(\frac{\partial p(x,y,t)}{\partial y}\right)_{y=H} = 0 \tag{8}$$

---

[1]Some authors refer to the Navier-Stokes equations as just the momentum equation.

$$\left(\frac{\partial p(x,y,t)}{\partial x}\right)_{x=0} = 0 \qquad \left(\frac{\partial p(x,y,t)}{\partial x}\right)_{x=L} = 0 \tag{9}$$

where:

$$g(x,y,t) = \left(\frac{\partial u}{\partial x}\right)^2 + 2\frac{\partial u}{\partial y}\frac{\partial v}{\partial x} + \left(\frac{\partial v}{\partial y}\right)^2 \tag{10}$$

$$h(x,y,t) = \frac{\partial f_x}{\partial x} + \frac{\partial f_y}{\partial y} \tag{11}$$

The main goal of this work is to develop the integral transformation technique to solve the Poisson equation (7) showed above.

Single Transformation (ST)

In order to establish the transformation pair, the pressure field is written as function of an orthogonal eigenfunctions obtained from the following auxiliary eigenvalue problem known as the Helmholtz classical problem [Mikhailov and Osizik (1984)], where $X_n(x)$ are the eigenfunctions and λ_n are the eigenvalues.

$$\frac{d^2 X_n(x)}{dx^2} + \lambda_n^2 X_n(x) = 0 \tag{12}$$

$$X'(0) = 0 \qquad\qquad X'(L) = 0 \tag{13}$$

which has the following solution:

$$X_n(x) = \cos\left(x\lambda_n\right) \tag{14}$$

$$\lambda_n = \frac{\pi n}{L} \qquad \text{for} \qquad n = 1,2,3,... \tag{15}$$

It should be noted that for these boundary conditions, one needs also to account for non-trivial solutions corresponding to $\lambda_0 = 0$.

$$X_0 = 1 \tag{16}$$

$$\lambda_0 = 0 \tag{17}$$

Now, the transformation pair can be defined:

$$\text{Transformation} \Rightarrow \overline{p}_n(y,t) = \int_0^L p(x,y,t)\,X(x)_n\,dx \tag{18}$$

$$\text{Inversion} \Rightarrow p(x,y,t) = p_0(y,t) + \sum_{n=1}^{\infty}\frac{X_n(x)\overline{p}_n(y,t)}{N_n} \tag{19}$$

where the norm N_n is defined by:

$$N_n = \int_0^L X_n^2\,dx = \frac{L}{2} \qquad \text{for} \qquad n \neq 0 \tag{20}$$

The final solution is given by two portions of the pressure: the average pressure in x direction p_{avg} and the modified pressure p_{mod}:

$$p(x,y,t) = p_{\text{avg}}(y,t) + p_{\text{mod}}(x,y,t) \tag{21}$$

where p_{avg} comes from the solution of the eigenproblem when $\lambda = 0$ and p_{mod} comes from the solution when $\lambda \neq 0$, in other words:

$$p_{\text{avg}}(y,t) = p_0(y,t) \tag{22}$$

$$p_{\text{mod}}(x,y,t) = \sum_{n=1}^{\infty} \frac{X_n(x)\overline{p}_n(y,t)}{N_n} \tag{23}$$

Solution for p_{mod} $(\lambda \neq 0)$

The integral transformation of the governing differential equation is derived by applying the operator $\int_0^L (\bullet) X_n \, dx$ on equation (7), obtaining the following transformed Poisson equation:

$$\frac{\partial^2 \overline{p}_n(y,t)}{\partial y^2} - \lambda_n^2 \overline{p}_n(y,t) = \rho \overline{h}_n(y,t) - \rho \overline{g}_n(y,t) \tag{24}$$

$$\left(\frac{\partial \overline{p}_n(y,t)}{\partial y} \right)_{y=0} = 0 \tag{25}$$

$$\left(\frac{\partial \overline{p}_n(y,t)}{\partial y} \right)_{y=H} = 0 \tag{26}$$

where the transformation of the parameters g and h are given by:

$$\overline{g}_n(y,t) = \int_0^L g(x,y,t) X_n \, dx \tag{27}$$

$$\overline{h}_n(y,t) = \int_0^L h(x,y,t) X_n \, dx \tag{28}$$

The equation (24) has an analytical solution that is shown bellow:
$$\overline{p}_n(y,t) =$$

$$e^{-y\lambda_n} \left(-\frac{1}{2} \cosh(y\lambda_n) \operatorname{csch}(H\lambda_n) e^{(H+y)\lambda_n} \int_0^H \frac{\rho e^{-y\lambda_n}}{\lambda_n} (\overline{h}_n(y,t) - \overline{g}_n(y,t)) \, dy + \right.$$

$$\frac{1}{4} (e^{2y\lambda_n} + 1)(\coth(H\lambda_n) - 1) \int_0^H \frac{\rho e^{y\lambda_n}}{\lambda_n} (\overline{g}_n(y,t) - \overline{h}_n(y,t)) \, dy + \tag{29}$$

$$\left. \frac{1}{2} \int_0^y \frac{\rho e^{y'\lambda_n}}{\lambda_n} (\overline{g}_n(y',t) - \overline{h}_n(y',t)) \, dy' + \frac{1}{2} e^{2y\lambda_n} \int_0^y \frac{\rho e^{-y'\lambda_n}}{\lambda_n} (\overline{h}_n(y',t) - \overline{g}_n(y',t)) \, dy' \right)$$

To find the actual solution for modified pressure p_{mod}, the inversion formula is used, equation (22). By observing equations (29), (27) and (28), one can notice that there are integrals of the discrete variables u, v, f_x and f_y. In order to compute these integrals, the following integral separation is proposed:

$$\int_0^L \Lambda(u,v,f_x,f_y) \, dx = \sum_{q=1}^{i_{\max}} \int_{x_{q-1}}^{x_q} \Lambda(u,v,f_x,f_y) \, dx \tag{30}$$

where Λ is a general function of u, v, f_x and f_y.

Then, to compute the integrals analytically, a Taylor expansion is used to expand the variables u, v, f_x and f_y in each subdomain:

Solution for p_{avg} ($\lambda = 0$)

In order to obtain the transformed differential equation for $\lambda = 0$, a very similar process is done. The transformed equation is given bellow:

$$\frac{\partial^2 \overline{p}_0(y,t)}{\partial y^2} = \rho \overline{h}_0(y,t) - \rho \overline{g}_0(y,t) \tag{31}$$

$$\left(\frac{\partial \overline{p}_0(y,t)}{\partial y} \right)_{y=0} = 0 \tag{32}$$

$$\left(\frac{\partial \overline{p}_0(y,t)}{\partial y} \right)_{y=H} = 0 \tag{33}$$

where:

$$\overline{g}_0(y,t) = \int_0^L g(x,y,t)\,dx \tag{34}$$

$$\overline{h}_0(y,t) = \int_0^L h(x,y,t)\,dx \tag{35}$$

The previous equation admits analytical solution of the following form:

$$\overline{p}_0(y,t) = \int_0^y \int_0^{y''} \rho\left(\overline{h}_0(y',t) - \overline{g}_0(y',t) \right) dy'\,dy'' + c_1 y + c_2 \tag{36}$$

applying the boundary conditions, one arrives to the following system of equations:

$$c_1 = 0 \tag{37}$$

$$0 = \int_0^H \rho\left(\overline{h}_0(y,t) - \overline{g}_0(y,t) \right) dy + c_1 \tag{38}$$

which tells that c_1 must be zero and the integral also must be zero:

$$\int_0^H \rho\left(\overline{h}_0(y,t) - \overline{g}_0(y,t) \right) dy = 0 \tag{39}$$

Equation (9) is know as the Poisson-Neumann compatibility condition [Abdallah (1987; 1988); Pozrikidis (2001)]. Knowing these information, the solution of the transformed differential equation is achieved:

$$\overline{p}_0(y,t) = \int_0^y \int_0^{y''} \rho\left(-\overline{g}_0(y',t) + \overline{h}_0(y',t) \right) dy'\,dy'' \tag{40}$$

Then the same integral separation (equation (19)) and Taylor series expansions are used to derive analytically the coefficients $\overline{h}_0$ and $\overline{g}_0$:

Double Transformation (DT)

In a very similar manner from previous formulation, one first establishes the transformation pair. In order to obtain that for this approach, two eigenvalue problems are defined. The eigenvalue problem associated with the x direction is given by: Eigenvalue problem associated with the problem in x direction

$$\frac{d^2 X_n(x)}{dx^2} + \lambda_n^2 X_n(x) = 0 \tag{41}$$

$$X'(0) = 0 \qquad X'(L) = 0 \tag{42}$$

which has the following solution:

$$X_n(x) = \cos\left(x\lambda_n \right) \tag{43}$$

$$\lambda_n = \frac{\pi n}{L} \qquad \text{for} \qquad n = 1, 2, 3, \dots \tag{44}$$

The problem associated with the y direction is given by:

$$\frac{d^2 Y_n(y)}{dy^2} + \beta_m^2 Y_n(y) = 0 \tag{45}$$

$$Y'(0) = 0 \qquad\qquad Y'(H) = 0 \tag{46}$$

which has the following solution:

$$Y_n(x) = \cos\left(y\beta_m\right) \tag{47}$$

$$\beta_m = \frac{\pi m}{H} \qquad \text{for} \qquad m = 1, 2, 3, \dots \tag{48}$$

The transformation pair can be defined:

$$\text{Transformation} \Rightarrow \bar{\bar{p}}_{n,m}(t) = \int_0^H \int_0^L p(x,y,t) X(x)_n Y(x)_m \, dx \, dy \tag{49}$$

$$\text{Inversion} \Rightarrow p(x,y,t) = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \frac{\bar{\bar{p}}_n(y,t) X_n(x) Y_m(y)}{Nx_n \, Ny_m} \tag{50}$$

where the norms Nx_n and Ny_m are defined by:

$$Nx_n = \int_0^L X_n^2 \, dx = \frac{L}{2} \qquad \text{for} \qquad n \neq 0 \tag{51}$$

$$Ny_m = \int_0^H Y_m^2 \, dx = \frac{H}{2} \qquad \text{for} \qquad m \neq 0 \tag{52}$$

it is also known that $Nx_0 = L$ and $Ny_0 = H$

Applying the operator $\int_0^H \int_0^L (\bullet) X_n Y_m \, dx \, dy$ on the Poisson equation, the following transformed Poisson equation (algebraic) is obtained:

$$-(\lambda_n^2 + \beta_m^2) \bar{\bar{p}}_{n,m}(y,t) = \rho \bar{\bar{h}}_{n,m}(y,t) - \bar{\bar{g}}_{n,m}(y,t) \tag{53}$$

where the transformation of the parameters g and h are given by:

$$\bar{\bar{g}}_{n,m}(t) = \int_0^H \int_0^L g(x,y,t) X_n Y_m \, dx \, dy \tag{54}$$

$$\bar{\bar{h}}_{n,m}(t) = \int_0^H \int_0^L h(x,y,t) X_n Y_m \, dx \, dy \tag{55}$$

Which has a direct solution shown bellow:

$$\bar{\bar{p}}_{n,m}(t) = \frac{\rho(\bar{\bar{g}}_{n,m} - \bar{\bar{h}}_{n,m})}{\lambda_n^2 + \beta_m^2} \tag{56}$$

In order to obtain the final solution, the inversion formula needs to be applied, obtaining:

$$p(x,y,t) = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \frac{\rho(\bar{\bar{g}}_{n,m} - \bar{\bar{h}}_{n,m})}{\lambda_n^2 + \beta_m^2} \frac{X_n(x) Y_m(y)}{Nx_n \, Ny_m} \tag{57}$$

which can also be rewritten in the form:

$$p(x,y,t) = \frac{\overline{\overline{p}}_{0,0}}{\mathrm{Nx}_0\,\mathrm{Ny}_0} + \sum_{m=1}^{\infty} \frac{\rho(\overline{\overline{g}}_{0,m} - \overline{\overline{h}}_{0,m})}{\beta_m^2}\,\frac{Y_m(y)}{\mathrm{Nx}_0\,\mathrm{Ny}_m} + \sum_{n=1}^{\infty} \frac{\rho(\overline{\overline{g}}_{n,0} - \overline{\overline{h}}_{n,0})}{\lambda_n^2}\,\frac{X_n(x)}{\mathrm{Nx}_n\,\mathrm{Ny}_0} +$$
$$+ \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} \frac{\rho(\overline{\overline{g}}_{n,m} - \overline{\overline{h}}_{n,m})}{\lambda_n^2 + \beta_m^2}\,\frac{X_n(x)Y_m(y)}{\mathrm{Nx}_n\,\mathrm{Ny}_m} \tag{58}$$

where $\overline{\overline{p}}_{0,0}$ is an arbitrary constant, which will be considered to be zero.

The greatest advantage of this approach is that it requires a lot less analytical effort and the final solution is more simple and compact. But the final solution has a double summation that can increase computational cost. In order to minimize this cost, one can use a reordering scheme, switching from the double summation to a single one.

$$\sum_{n=1}^{\infty}\sum_{m=1}^{\infty} \Rightarrow \sum_{k=1}^{\infty} \tag{59}$$

This can be done knowing the sum terms with higher magnitude and putting them in the beginning of the sum. This is achieved by taking (n,m) pairs that promote lowers $(\lambda_n^2 + \beta_m^2)$. By doing that, one can arrive at the expression:

$$p(x,y,t) = \sum_{m=1}^{\infty} \frac{\rho(\overline{\overline{g}}_{0,m} - \overline{\overline{h}}_{0,m})}{\beta_m^2}\,\frac{Y_m(y)}{\mathrm{Nx}_0\,\mathrm{Ny}_m} + \sum_{n=1}^{\infty} \frac{\rho(\overline{\overline{g}}_{n,0} - \overline{\overline{h}}_{n,0})}{\lambda_n^2}\,\frac{X_n(x)}{\mathrm{Nx}_n\,\mathrm{Ny}_0} +$$
$$+ \sum_{k=1}^{\infty} \frac{\rho(\overline{\overline{g}}_{n(k),m(k)} - \overline{\overline{h}}_{n(k),m(k)})}{\lambda_{n(k)}^2 + \beta_{m(k)}^2}\,\frac{X_{n(k)}(x)Y_{m(k)}(y)}{\mathrm{Nx}_{n(k)}\,\mathrm{Ny}_{m(k)}} \tag{60}$$

Discrete Derivatives

In order to solve the pressure problem, the discrete derivatives of u, v, f_x and f_y must be calculated. In this work, a second order central differencing scheme is used inside the domain and second-order the backward/forward (depending of the boundary) differencing scheme is used at the boundaries.

Results

For all cases presented in this chapter, $L=1$, $H=1$, and $\rho=1$ were used. The chosen source term of the Poisson equation (4) satisfies the compatibility condition (9) and it is of the following form:

$$\left[\rho(h(x,y,t) - g(x,y,t))\right]_{i,j} = \rho\left(x_i^4 + x_i^3 + x_i^2 + x_i - \frac{77}{60}\right)\left(y_j^4 + y_j^3 + y_j^2 + y_j\right) \tag{61}$$

A comparison of computational cost is done for both techniques presented in this work. In order to compare the CITT performance, a fixed mesh is used and many truncation orders for the summations are computed, so only the CITT error is captured. The CITT error is calculated using the following formula:

$$\epsilon_{i,j}^{\mathrm{CITT}}(n_{\max}) = \mathrm{abs}[\,p_{i,j}(n_{\max}) - p_{i,j}(n_{\max}+5)\,] \tag{62}$$

The mesh error for the x-mesh and the y-mesh is calculated using the following formulations respectively:

$$\epsilon_{i,1024}^{x}(n_{\max}) = \mathrm{abs}[\,p_{i,1024}(n_{\max}) - p_{2i,1024}(n_{\max})\,] \tag{63}$$

$$\epsilon_{1024,j}^{y}(n_{\max}) = \mathrm{abs}[\,p_{1024,j}(n_{\max}) - p_{1024,2j}(n_{\max})\,] \tag{64}$$

The codes were compiled and ran using GFORTRAN and the flags -O3 and -fopenmp and in a 8 core CPU machine.

In order to illustrate the convergence of the solution with the variation of Δx, figure 1 shows a graphic of the maximum absolute error with the variation of the mesh size Δx. As one can observe the convergence order is about 2, which was expected since all approximation made in the mathematical formulation were of this order.

Figure 2 shows the convergence of the absolute error with Δy. Although it seems to have a higher convergence order for the poorer refined meshes, the order stabilizes at 2 when more refined meshes are implemented.
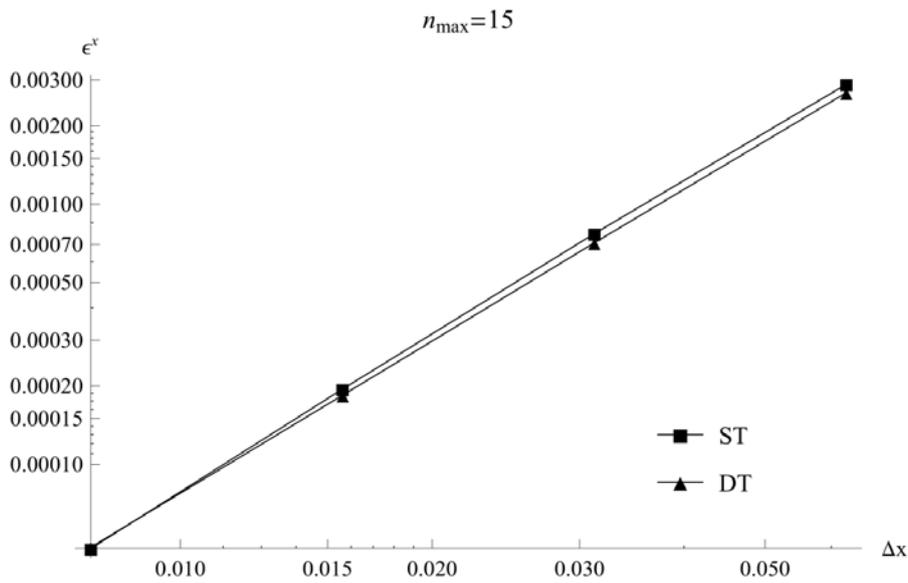
Figure 1. Convergence for the mesh in x-direction, $n_{max} = 15$ and using $j_{max} = 1024$.
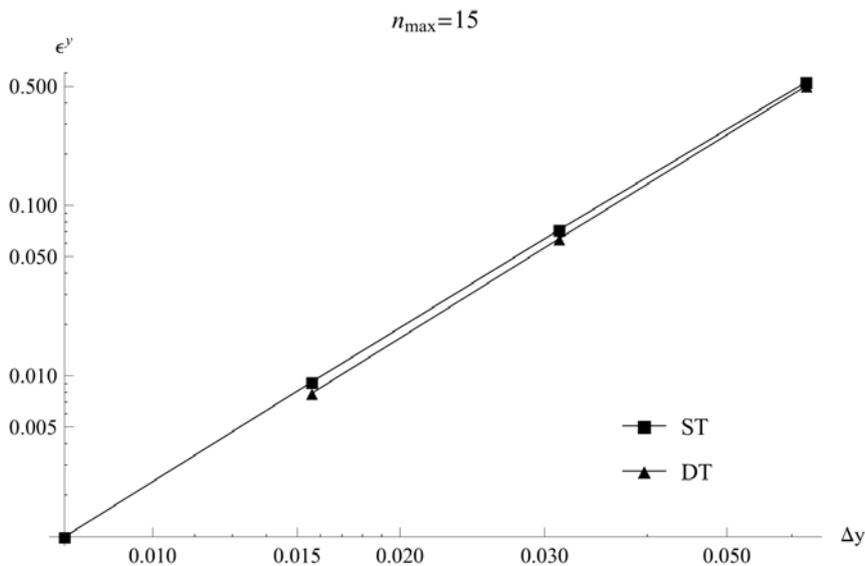
Figure 2. Convergence for the mesh in y-direction, $n_{max} = 15$ and using $i_{max} = 1024$.

In order to illustrate the problem, tables 1 and 2 show the value of the pressure for different points of the domain. It is clear that both methodologies converge to the same value, even though the values are not fully converged with six digits accuracy. One can see that the mesh convergence is very similar for both methodologies.

Table 1. Convergence of the mesh for some points of the domain for $y = 0.25$ and $n_{max} = 30$.

$i_{max} \times j_{max}$	$x = 0.25$		$x = 0.75$	
	ST	DT	ST	DT
16×16	0.0801190	0.0793513	-0.0799973	-0.0809956
32×32	0.0811750	0.0810473	-0.0807687	-0.0811251
64×64	0.0814272	0.0814572	-0.0809571	-0.0811557
128×128	0.0814885	0.0815574	-0.0810034	-0.0811630

Table 2. Convergence of the mesh for some points of the domain for $y = 0.75$ and $n_{max} = 30$.

$i_{max} \times j_{max}$	$x = 0.25$		$x = 0.75$	
	ST	DT	ST	DT
16×16	0.183004	0.186430	-0.185268	-0.181505
32×32	0.186532	0.187220	-0.184883	-0.183862
64×64	0.187368	0.187422	-0.184825	-0.184438
128×128	0.187571	0.187473	-0.184815	-0.184580

Figure 3 shows a comparison of the computational cost of both methodologies for a mesh of 32×32. One can clearly see that the Double Transformation (DT) needs more time to obtain the same error. This effect is due to the introduction of a double summation needed to solve the problem by the DT which requires a bigger effort to compute the solution. Even with the implementation of the reordering scheme, it is not enough to beat the computational cost of the ST. The same effect can be seen on figure 4, which shows the computational cost for a mesh of 128×128.
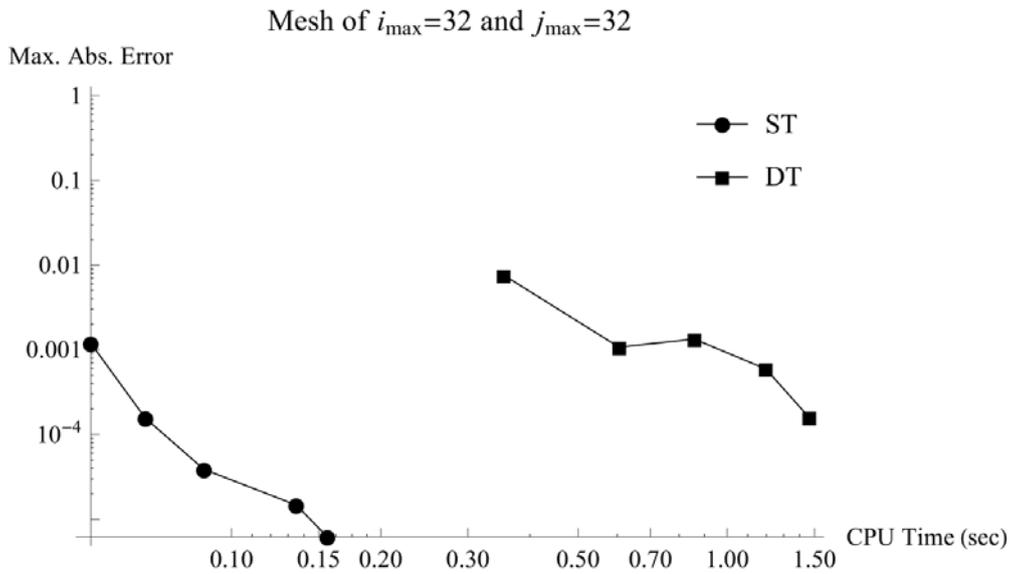
Figure 3. Comparison of the computational cost of CITT using single transformation and CITT using double transformation for a mesh $i_{max} = 32$ and $j_{max} = 32$.
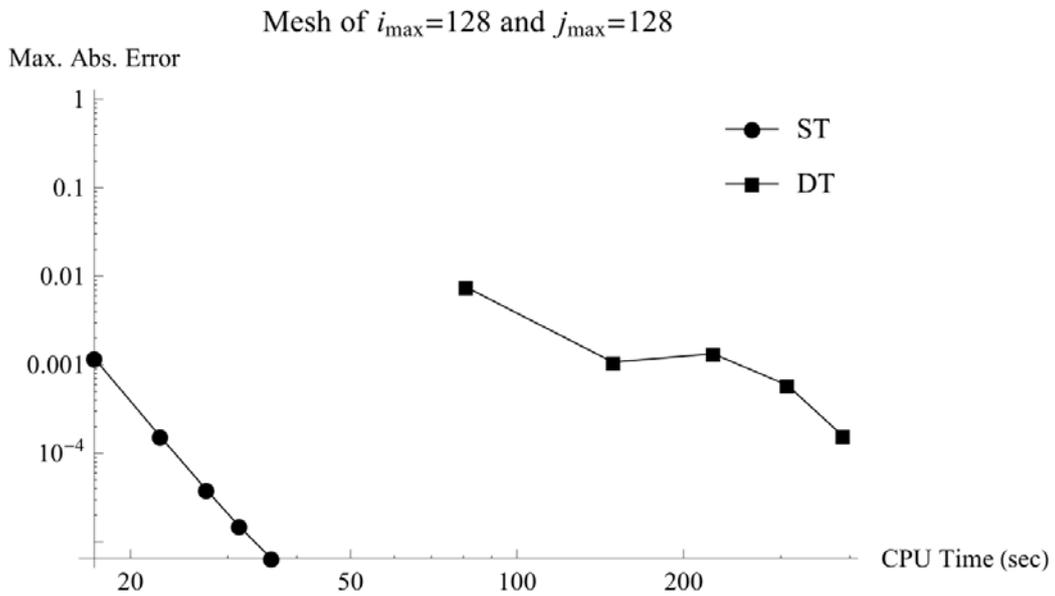
Figure 4. Comparison of the computational cost of CITT using single transformation and CITT using double transformation for a mesh $i_{max} = 128$ and $j_{max} = 128$.

Conclusion

This work presented results of the solution of the Poisson equation arising from the incompressible Navier-Stokes equations. The main motivation of the proposed work is the implementation of this semi-analytical formulation for the Poisson equation in the momentum equation and thus solve it using a numerical technique for initial value problem.

The solution of the Poisson equation using this semi-analytical approach was accomplished using two different schemes: CITT single transformation (ST) and CITT double transformation (DT). Both techniques presented a very similar convergence behavior and results, showing that the formulations proposed are consistent. The comparison between both schemes showed that the double transformation has poorer performance in comparison with the single transformation scheme. Even though the performance of the double transformations was not so good, this formulation has a simpler analytical solution, which might be more interesting when implementing the pressure solution in the momentum equation, and so being possible a fully explicit time marching method for time.

The proposed schemes are very good for smooth pressure fields, but it might have some convergence problems with discontinuous pressure fields, which arise in phase-change problems, due to the Gibbs phenomenon in series truncation. This issue must be further investigated in order to compute the real impact of the implementation of the proposed formulations in these type of problems.

Acknowledgments

References

Guermond, J. L., Minev, P., and Shen, J., 2006. "*An overview of projection methods for incompressible flows*". Comput. Methods Appl. Mech. Engrg., 195, p. 6045.

Munz, C.-D., Roller, S., Klein, R., and Geratz, K., 2003. "*The extension of incompressible flow solvers to the weakly compressible regime*". Computers & Fluids, 32(2), Feb., pp. 173–196.

Mikhailov, M. D., and Ozisik, M. N., 1984. *Unified Analysis and Solutions of Heat and Mass Diffusion*. John Wiley & Sons, New York.

Wolfram, S., 2003. *The Mathematica Book*, 5th ed. Wolfram Media/Cambridge University Press, New York/Champaign, IL.

Chalhub, D. J. M. N., Sphaier, L. A., and de B. Alves, L. S., 2013. "*Integral transform solution of convective heat transfer problems using upwind approximations*". Numerical Heat Transfer - Part B, 63, pp. 167–187.

Guedes, R. O. C., and Ozisik, M. N., 1994. "*Transient heat transfer in simultaneously developing channel flow with step change in inlet temperature*". International Journal Heat and Mass Transfer, 37, pp. 2699–2706.

Guedes, R. O. C., and Ozisik, M. N., 1994. "*Hybrid approach for solving unsteady laminar forced convection inside ducts with periodically varying inlet temperature*". Int. J. Heat and Fluid Flow, 15, pp. 116–121.

Cotta, R. M., and Gerk, J. E. V., 1994. "*Mixed finite-difference/integral transform approach for parabolic-hyperbolic problems in transient forced convection*". Numerical Heat Transfer - Part B, 25, pp. 442–448.

Castellões, F. V., and Cotta, R. M., 2006. "*Analysis of transient and periodic convection in microchannels via integral transforms*". Progress in Computational Fluid Dynamics, 6, pp. 321–326.

Naveira, C. P., Lachi, M., Cotta, R. M., and Padet, J., 2009. "*Hybrid formulation and solution for transient conjugated conduction-external convection*". International Journal of Heat and Mass Transfer, 52(1-2), pp. 112–123.

Kundu, P. K., 1990. *Fluid Mechanics*. Academic Press, San Diego, CA.

Abdallah, S., 1987. "*Numerical solutions for the pressure poisson equation with neumann boundary conditions using a non-staggered grid, i*". Journal of computational physics, 70(1), p. 182–192.

Abdallah, S., and Dreyer, J., 1988. "Dirichlet and neumann boundary conditions for the pressure poisson equation of incompressible flow". Journal for Numerical Methods in Fluids, 8(December 1987), p. 1029–1036.

Pozrikidis, C., 2001. "A note on the regularization of the discrete poisson-neumann problem". Journal of Computational Physics, 172(2), Sept., p. 917–923.

Simple Matrix Method with Nonhomogeneous Space Increments in Finite Difference Method

Sungki Min[1], Hunhee Kim[1], Soo Won Chae[2], and *Junghwa Hong[1]

[1] Department of Control and Instrumentation Engineering, Korea University, Sejong, Republic of Korea.
[2] Department of Mechanical Engineering, Korea University, Seoul, Republic of Korea.

*Corresponding author: hongjh32@korea.ac.kr

Abstract

FDM (finite difference method) has advantages for direct code expansion to numerically solve PDE (partial differential equation). In contrast to finite element method configured with nonhomogeneous element size based on interested region, a constant increment size of element or grid is used in existing FDM. As a result, the calculation near boundary could be inaccurate and incomplete. Particularly, in electrokinetics, except for diffuse layer (about $10\sim10^4$ Angstrom) near solid and fluid interface, calculation of the charge density distribution with having the same and fine space increments at another region is inefficient and could meet the catastrophic memory lack error. In this study, we provide a method how to make the nonhomogeneous space incremental FD grid and to configure PDE having complicated and mixed boundary conditions (e.g. Dirichlet and Neumann) with suggested simple matrices. The suggested FDM using the nonhomogeneous increment of fine grids near the boundaries and interfaces could increase the accuracy of solutions and efficiency of calculations.

Keywords: Finite Difference Method, Partial Differential Equation, Nonhomogeneous Increment, Dirichlet and Neumann Boundary Conditions

Introduction

Solutions of electrokinetic problems are important to understand living bone remodeling processes by electromechanical transduction effects on osteocytes, osteoclasts and osteoblasts. Since this electromechanical transduction postulation about bone remodeling processes requires the existence of ionic interstitial bone fluid flow in bone tissues, the resulting streaming potential is being focused on bone mechanics as a remodeling stimulation on bone cells [Pienkowski and Pollack (1983); Zhang et al. (1997)]. The streaming potential is about coupled phenomenon described by elasticity of bone tissue, fluid mechanics of interstitial fluid flow through canaliculi and lacunae, and electricity of charged ions in bone fluid.

In addition, the streaming potential that is an electrokinetic phenomenon is closely related to electrical charge on the wall of lacunocanalicular flow path. The streaming potential is very sensitively affected by the interface surface electrical potential of the canalicular wall [Ahn and Grodzinsky (2009)]. Since bone is a piezoelectric material [Bassett (1968)], the surface electrical potential of the canalicular wall is changed by the elasticity of bone tissue. As a result, a full analysis including transient behavior of streaming potential can be achieved for the multi-physical study of bone tissue after firstly considering effects of its piezoelectricity on the just boundaries of the lacunocanalicular flow paths on the transient electrokinetics.

To solve electrokinetic problems using FDM, except for diffuse layer (about $10\sim10^4$ Angstrom) near solid and fluid interface, calculation of charge density distribution with having the same space increments at another region is inefficient. In this study, we provide a method how to make the nonhomogeneous space incremental FD grid and to configure PDE having complicated and mixed

boundary conditions with suggested simple matrices. The suggested method is verified using the existing closed solution in the electrokinetics [ref].

Numerical Method

Notation and Definition

For the formulation used in this study, we used the following notations and definitions.

$$f_{i,j,k}^{\tau} \tag{1}$$

where subscript i, j, and k are the space incremental numbers; and superscript τ is the time incremental number. $i = 1 \sim L+1$; $j = 1 \sim M+1$; $k = 1 \sim N+1$; and $\tau = 1 \sim T+1$ where i, j, and k are the positive integer. f is a unknown quantity.

Position

The position is described as x_i, y_j, and z_k in the Cartesian coordinate, r_i, θ_j, and z_k in the cylindrical coordinate, r_i, θ_j, and ϕ_k in the spherical coordinates.

Variable Increment

The increment of the position or time is described by forward, backward and central difference.

$$
\begin{aligned}
(\Delta p)^+ &= p^+ - p^0 \\
(\Delta p)^- &= p^0 - p^- \\
(\Delta p)^0 &= 0.5(p^+ - p^-)
\end{aligned}
\tag{2}
$$

where p means an arbitrary position, time, or value; superscripts $+$, $-$, and 0 mean forward, backward, and central infinitesimal approaches, respectively. In the Cartesian coordinate, time and spatial increments are represented by (3).

$$
\begin{array}{lll}
(\Delta t^{\tau})^+ = t^{\tau+1} - t^{\tau} & (\Delta t^{\tau})^- = t^{\tau} - t^{\tau-1} & (\Delta t^{\tau})^0 = 0.5(t^{\tau+1} - t^{\tau-1}) \\
(\Delta x_i)^+ = x_{i+1} - x_i & (\Delta x_i)^- = x_i - x_{i-1} & (\Delta x_i)^0 = 0.5(x_{i+1} - x_{i-1}) \\
(\Delta y_j)^+ = y_{j+1} - y_j & (\Delta y_j)^- = y_j - y_{j-1} & (\Delta y_j)^0 = 0.5(y_{j+1} - y_{j-1}) \\
(\Delta z_k)^+ = z_{k+1} - z_k & (\Delta z_k)^- = z_k - z_{k-1} & (\Delta z_k)^0 = 0.5(z_{k+1} - z_{k-1})
\end{array}
\tag{3}
$$

First Order Derivatives

The first order derivatives of time and space are described by (4).

$$
\begin{array}{lll}
\left(\dfrac{\Delta f}{\Delta t}\right)^+ = \dfrac{f_{i,j,k}^{\tau+1} - f_{i,j,k}^{\tau}}{t^{\tau+1} - t^{\tau}} & \left(\dfrac{\Delta f}{\Delta t}\right)^- = \dfrac{f_{i,j,k}^{\tau} - f_{i,j,k}^{\tau-1}}{t^{\tau} - t^{\tau-1}} & \left(\dfrac{\Delta f}{\Delta t}\right)^0 = \dfrac{f_{i,j,k}^{\tau+1} - f_{i,j,k}^{\tau-1}}{t^{\tau+1} - t^{\tau-1}} \\[3mm]
\left(\dfrac{\Delta f}{\Delta x}\right)^+ = \dfrac{f_{i+1,j,k}^{\tau} - f_{i,j,k}^{\tau}}{x_{i+1} - x_i} & \left(\dfrac{\Delta f}{\Delta x}\right)^- = \dfrac{f_{i,j,k}^{\tau} - f_{i-1,j,k}^{\tau}}{x_i - x_{i-1}} & \left(\dfrac{\Delta f}{\Delta x}\right)^0 = \dfrac{f_{i+1,j,k}^{\tau} - f_{i-1,j,k}^{\tau}}{x_{i+1} - x_{i-1}} \\[3mm]
\left(\dfrac{\Delta f}{\Delta y}\right)^+ = \dfrac{f_{i,j+1,k}^{\tau} - f_{i,j,k}^{\tau}}{y_{j+1} - y_j} & \left(\dfrac{\Delta f}{\Delta y}\right)^- = \dfrac{f_{i,j,k}^{\tau} - f_{i,j-1,k}^{\tau}}{y_j - y_{j-1}} & \left(\dfrac{\Delta f}{\Delta y}\right)^0 = \dfrac{f_{i,j+1,k}^{\tau} - f_{i,j-1,k}^{\tau}}{y_{j+1} - y_{j-1}} \\[3mm]
\left(\dfrac{\Delta f}{\Delta z}\right)^+ = \dfrac{f_{i,j,k+1}^{\tau} - f_{i,j,k}^{\tau}}{z_{k+1} - z_k} & \left(\dfrac{\Delta f}{\Delta z}\right)^- = \dfrac{f_{i,j,k}^{\tau} - f_{i,j,k-1}^{\tau}}{z_k - z_{k-1}} & \left(\dfrac{\Delta f}{\Delta z}\right)^0 = \dfrac{f_{i,j,k+1}^{\tau} - f_{i,j,k-1}^{\tau}}{z_{k+1} - z_{k-1}}
\end{array}
\tag{4}
$$

Second order derivatives

The second order derivatives are described by (5).

$$\frac{\partial^2 f}{\partial p^2} = \frac{\left(\frac{\Delta f}{\Delta p}\right)^+ - \left(\frac{\Delta f}{\Delta p}\right)^-}{(\Delta p)^0} = \frac{\left(\frac{f^+ - f^0}{p^+ - p^0}\right) - \left(\frac{f^0 - f^-}{p^0 - p^-}\right)}{(\Delta p)^0} = \frac{f^-}{(\Delta p)^0 (\Delta p)^-} - \frac{2f^0}{(\Delta p)^+ (\Delta p)^-} + \frac{f^+}{(\Delta p)^+ (\Delta p)^0} \quad (5)$$

Thus, the second order derivatives of time and space are described by (6).

$$\frac{\partial^2 f}{\partial t^2} = \frac{f_{i,j,k}^{\tau-1}}{(\Delta t^\tau)^0 (\Delta t^\tau)^-} - \frac{2 f_{i,j,k}^{\tau}}{(\Delta t^\tau)^+ (\Delta t^\tau)^-} + \frac{f_{i,j,k}^{\tau+1}}{(\Delta t^\tau)^+ (\Delta t^\tau)^0}$$

$$\frac{\partial^2 f}{\partial x^2} = \frac{f_{i-1,j,k}^{\tau}}{(\Delta x_i)^0 (\Delta x_i)^-} - \frac{2 f_{i,j,k}^{\tau}}{(\Delta x_i)^+ (\Delta x_i)^-} + \frac{f_{i+1,j,k}^{\tau}}{(\Delta x_i)^+ (\Delta x_i)^0}$$

$$\frac{\partial^2 f}{\partial y^2} = \frac{f_{i,j-1,k}^{\tau}}{(\Delta y_j)^0 (\Delta y_j)^-} - \frac{2 f_{i,j,k}^{\tau}}{(\Delta y_j)^+ (\Delta y_j)^-} + \frac{f_{i,j+1,k}^{\tau}}{(\Delta y_j)^+ (\Delta y_j)^0} \qquad (6)$$

$$\frac{\partial^2 f}{\partial z^2} = \frac{f_{i,j,k-1}^{\tau}}{(\Delta z_k)^0 (\Delta z_k)^-} - \frac{2 f_{i,j,k}^{\tau}}{(\Delta z_k)^+ (\Delta z_k)^-} + \frac{f_{i,j,k+1}^{\tau}}{(\Delta z_k)^+ (\Delta z_k)^0}$$

Laplacian

In the Cartesian coordinate, Laplacian of a quantity is described by (7).

$$\nabla^2 f = \frac{f_{i-1,j,k}^{\tau}}{(\Delta x_i)^0 (\Delta x_i)^-} + \frac{f_{i+1,j,k}^{\tau}}{(\Delta x_i)^+ (\Delta x_i)^0} + \frac{f_{i,j-1,k}^{\tau}}{(\Delta y_j)^0 (\Delta y_j)^-} + \frac{f_{i,j+1,k}^{\tau}}{(\Delta y_j)^+ (\Delta y_j)^0} + \frac{f_{i,j,k-1}^{\tau}}{(\Delta z_k)^0 (\Delta z_k)^-} + \frac{f_{i,j,k+1}^{\tau}}{(\Delta z_k)^+ (\Delta z_k)^0}$$

$$- 2 \left(\frac{1}{(\Delta x_i)^+ (\Delta x_i)^-} + \frac{1}{(\Delta y_j)^+ (\Delta y_j)^-} + \frac{1}{(\Delta z_k)^+ (\Delta z_k)^-} \right) f_{i,j,k}^{\tau} \qquad (7)$$

Boundary Conditions

The Boundary conditions are defined on the interface locating at 1, $L + 1$, $M + 1$, and $N + 1$. Dirichlet boundary condition ($\phi = 0$ at interface) could be formed as (8).

x coordinate : $f_{1,j,k}^{\tau} = 0$ or $f_{L+1,j,k}^{\tau} = 0$ or $f_{1,j,k}^{\tau} = f_{L+1,j,k}^{\tau} = 0$

y coordinate : $f_{i,1,k}^{\tau} = 0$ or $f_{i,M+1,k}^{\tau} = 0$ or $f_{i,1,k}^{\tau} = f_{i,M+1,k}^{\tau} = 0$ $\qquad (8)$

z coordinate : $f_{i,j,1}^{\tau} = 0$ or $f_{i,j,N+1}^{\tau} = 0$ or $f_{i,j,1}^{\tau} = f_{i,j,N+1}^{\tau} = 0$

Neumann boundary condition ($\Delta \phi / \Delta x = 0$ at interface) could be formed as (9).

x coordinate : $f_{1,j,k}^{\tau} = f_{2,j,k}^{\tau}$ or $f_{L+1,j,k}^{\tau} = f_{L,j,k}^{\tau}$ or $f_{1,j,k}^{\tau} = f_{2,j,k}^{\tau}$ and $f_{L+1,j,k}^{\tau} = f_{L,j,k}^{\tau}$

y coordinate : $f_{i,1,k}^{\tau} = f_{i,2,k}^{\tau}$ or $f_{i,M+1,k}^{\tau} = f_{i,M,k}^{\tau}$ or $f_{i,1,k}^{\tau} = f_{i,2,k}^{\tau}$ and $f_{i,M+1,k}^{\tau} = f_{i,M,k}^{\tau}$ $\qquad (9)$

z coordinate : $f_{i,j,1}^{\tau} = f_{i,j,2}^{\tau}$ or $f_{i,j,N+1}^{\tau} = f_{i,j,N}^{\tau}$ or $f_{i,j,1}^{\tau} = f_{i,j,2}^{\tau}$ and $f_{i,j,N+1}^{\tau} = f_{i,j,N}^{\tau}$

Formation of Matrix

For example, $\left[\dfrac{\partial^2}{\partial \mathbf{x}^2} \right]$ could be represented by (10) as a matrix form.

$$\left[\frac{\partial^2}{\partial \mathbf{x}^2}\right] = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ \dfrac{1}{(\Delta x_2)^0(\Delta x_2)^-} & \dfrac{-2}{(\Delta x_2)^+(\Delta x_2)^-} & \dfrac{1}{(\Delta x_2)^+(\Delta x_2)^0} & 0 & \cdots & 0 & 0 \\ 0 & \dfrac{1}{(\Delta x_3)^0(\Delta x_3)^-} & \dfrac{-2}{(\Delta x_3)^+(\Delta x_3)^-} & \dfrac{1}{(\Delta x_3)^+(\Delta x_3)^0} & 0 & \cdots & 0 \\ \vdots & & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \dfrac{1}{(\Delta x_{L-1})^0(\Delta x_{L-1})^-} & \dfrac{-2}{(\Delta x_{L-1})^+(\Delta x_{L-1})^-} & \dfrac{1}{(\Delta x_{L-1})^+(\Delta x_{L-1})^0} & 0 \\ 0 & 0 & \cdots & 0 & \dfrac{1}{(\Delta x_L)^0(\Delta x_L)^-} & \dfrac{-2}{(\Delta x_L)^+(\Delta x_L)^-} & \dfrac{1}{(\Delta x_L)^+(\Delta x_L)^0} \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 \end{bmatrix} \quad (10)$$

$$= \begin{bmatrix} \mathbf{0} \\ \left(\dfrac{\partial^2}{\partial \mathbf{x}^2}\right)_{2\sim L} \\ \mathbf{0} \end{bmatrix}$$

The matrix (10) is a square matrix composed of $(L+1) \times (L+1)$. To obtain the elements of the first row, the history at $(\Delta x_1)^-$ is required as well as x_0. However, x_0 is not yet defined. Thus, the elements of the first row are all zero. Since x_{L+2} is not defined for $(\Delta x_{L+1})^+$, the elements of the $(L+1)^{\text{th}}$ column of the matrix are all zero. As a result, the matrix $\left[\dfrac{\partial^2}{\partial \mathbf{x}^2}\right]$ is composed of second order derivatives from the second and to L^{th} rows. In addition, the first and $(L+1)$ rows are formed by boundary conditions.

Formation of Matrix with Boundary Conditions

When a matrix **A** has Neumann boundary condition at node 1, Dirichlet boundary condition at node $L+1$, and the second order derivatives of the matrix **A** is the same as a matrix **B**, (11) can be written.

$$\frac{\partial^2 \mathbf{A}}{\partial x^2} = \mathbf{B} \quad (11)$$

As a matrix form of (11) can be (12)

$$\mathbf{M} \cdot \mathbf{A} = \mathbf{B} \quad (12)$$

where **M** has a following form.

$$M = \underbrace{\begin{bmatrix} \mathbf{0} \\ \left(\dfrac{\partial^2}{\partial \mathbf{x}^2}\right)_{2\sim L} \\ \mathbf{0} \end{bmatrix}}_{\substack{\text{matrix of second order derivaties} \\ \text{for } 2\sim L \text{ rows}}} + \underbrace{\begin{bmatrix} 1 & -1 \\ & \mathbf{0} \\ & & 1 \end{bmatrix}}_{\text{matrix of boundary conditions}} = \begin{bmatrix} 1 & -1 \\ \left(\dfrac{\partial^2}{\partial \mathbf{x}^2}\right)_{2\sim L} \\ & & 1 \end{bmatrix}$$

Therefore, (12) can be represented by (13)

$$\begin{bmatrix} 1 & -1 \\ \left(\dfrac{\partial^2}{\partial \mathbf{x}^2}\right)_{2\sim L} \\ & & 1 \end{bmatrix} \begin{Bmatrix} a_1 \\ a_2 \\ \vdots \\ a_L \\ a_{L+1} \end{Bmatrix} = \begin{Bmatrix} 0 \\ b_2 \\ \vdots \\ b_L \\ 0 \end{Bmatrix} \quad (13)$$

Verification

Verification of theoretical approach was accomplished by comparing with the reference method [Bazant et al. (2004)]. As shown in Fig. 1, an external electrical potential is applied to the isolated ironic fluid.
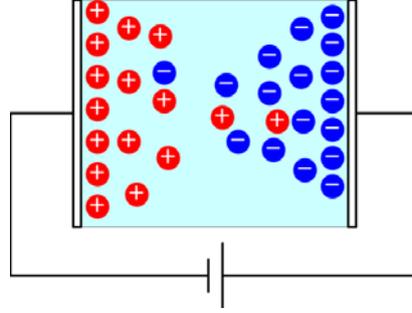
Figure 1. Electrolyte system affected by external electric field which induces the electric potential distribution of $\phi(y)$ with surface potential $\phi_L = -1.0$ V and $\phi_R = +1.0$ V

Analytical Method

The governing equation for the charge density can be represented by (14).

$$\frac{1}{D}\frac{\partial \rho_f}{\partial t} = \nabla^2 \rho_f - \kappa^2 \rho_f + \kappa^2 \varepsilon_f \nabla^2 \phi \tag{14}$$

$$\nabla^2 \psi = -\rho_f / \varepsilon_f$$

ϕ is the external electrical potential generated by external electric field. For the external electrical potential, $\nabla^2 \phi = 0$. Thus, (14) is reduced to Debye-Falkenhagen equation (15).

$$\frac{1}{D}\frac{\partial \rho_f}{\partial t} = \nabla^2 \rho_f - \kappa^2 \rho_f \tag{15}$$

Based on (15), distributions of the electric charge density, ρ_f, change distributions and magnitudes of internal electric potential, ψ.

If ρ_f is a function of the only x-coordinate,

$$\frac{1}{D}\frac{\partial \rho_f}{\partial t} = \frac{\partial^2 \rho_f}{\partial x^2} - \kappa^2 \rho_f \tag{16}$$

where Φ is the total potential that is a summation of the external and internal potentials.

As assumed to be two plates, the one-dimensional boundary condition is represented by (17).

$$\frac{\partial \rho_f}{\partial x} = -\kappa^2 \varepsilon_f \frac{\partial \Phi}{\partial x}\bigg|_{at\ x=\pm h} = -\kappa^2 \varepsilon_f \left(\frac{\partial \psi}{\partial x} + \frac{\partial \phi}{\partial x}\right)\bigg|_{at\ x=\pm h} \tag{17}$$

A general solution of (17) in the Laplace domain (s-domain) can be obtained by (18) [Bazant et al. (2004)].

$$\hat{\rho}_f(x,s) = A(s)\sinh(kx) \tag{18}$$

where $A(s) = \dfrac{-k^2 \varepsilon_f V s^{-1} \text{sech}(kh)}{\tanh(kh) + \lambda_s k + \dfrac{ksh}{\kappa^2 D}\left(1 + \dfrac{\lambda_s}{h}\right)}$, $k^2 = \dfrac{s}{D} + \kappa^2$.

Table 1. Symbols

Symbols	Definitions
ρ_f	charge density
ψ	internal potential due to charge density distribution
ϕ	external potential due to external electric field
D	diffusion coefficient
κ	inverse of Debye length
ε_f	dielectric permittivity
h	half channel height
λ_s	effective thickness for the compact part of the double layer
V	external potential imposed by the external circuit

Numerical Method

The governing equation for the charge density can be discretized by (19).

$$\frac{1}{D}\frac{\rho_f^{t+1} - \rho_f^t}{\Delta t} - \nabla^2 \rho_f^{t+1} + \kappa^2 \rho_f^{t+1} = \kappa^2 \varepsilon_f \nabla^2 \phi^{t+1}$$

$$\left(\frac{1}{D\Delta t} - \nabla^2 + \kappa^2\right)\rho_f^{t+1} = \frac{1}{D\Delta t}\rho_f^t + \kappa^2 \varepsilon_f \nabla^2 \phi^{t+1} \tag{19}$$

When ρ_f is a function of the only x-coordinate,

$$\frac{1}{D}\frac{\rho_f^{t+1} - \rho_f^t}{\Delta t} - \left[\frac{\partial^2}{\partial \mathbf{x}^2}\right]\rho_f^{t+1} + \kappa^2 \rho_f^{t+1} = \kappa^2 \varepsilon_f \left[\frac{\partial^2}{\partial \mathbf{x}^2}\right]\phi^{t+1} \tag{20}$$

Boundary condition

The boundary conditions can be represented by (21).

$$\rho_{f1} - \rho_{f2} = -\kappa^2 \varepsilon_f \left[(\psi_1 - \psi_2) + (\phi_1 - \phi_2)\right]$$

$$\rho_{fL} - \rho_{fL+1} = -\kappa^2 \varepsilon_f \left[(\psi_L - \psi_{L+1}) + (\phi_L - \phi_{L+1})\right] \tag{21}$$

At the interface, the internal potential should be 0 to satisfy $\psi_{L+1} = \psi_1 = 0$. Then, (21) turns out to be (22).

$$\rho_{f1} - \rho_{f2} = -\kappa^2 \varepsilon_f \left[\psi_2 + (\phi_1 - \phi_2)\right]$$

$$\rho_{fL} - \rho_{fL+1} = -\kappa^2 \varepsilon_f \left[\psi_L + (\phi_L - \phi_{L+1})\right] \tag{22}$$

The governing equation (20) can be described by (23).

$$
\left(\frac{1}{D\Delta t} \begin{bmatrix} \mathbf{0} \\ \hline \mathbf{0} & \mathbf{I} & \mathbf{0} \\ \hline \mathbf{0} \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ \hline \left(\dfrac{\partial^2}{\partial \mathbf{x}^2} \right)_{2\sim L} \\ \hline \mathbf{0} \end{bmatrix} + \kappa^2 \begin{bmatrix} \mathbf{0} \\ \hline \mathbf{0} & \mathbf{I} & \mathbf{0} \\ \hline \mathbf{0} \end{bmatrix} \right) \left\{ \begin{array}{c} \rho_{f_1} \\ \hline \rho_{f_2} \\ \vdots \\ \hline \rho_{f_L} \\ \hline \rho_{f_{L+1}} \end{array} \right\}_{j,k}^{t+1} =
$$

$$
\frac{1}{D\Delta t} \begin{bmatrix} \mathbf{0} \\ \hline \mathbf{0} & \mathbf{I} & \mathbf{0} \\ \hline \mathbf{0} \end{bmatrix} \left\{ \begin{array}{c} \rho_{f_1} \\ \rho_{f_2} \\ \vdots \\ \rho_{f_L} \\ \rho_{f_{L+1}} \end{array} \right\}_{j,k}^{t} + \kappa^2 \varepsilon_f \begin{bmatrix} \mathbf{0} \\ \hline \left(\dfrac{\partial^2}{\partial \mathbf{x}^2} \right)_{2\sim L} \\ \hline \mathbf{0} \end{bmatrix} \left\{ \begin{array}{c} \phi_1 \\ \hline \phi_2 \\ \vdots \\ \phi_L \\ \hline \phi_{L+1} \end{array} \right\}_{j,k}^{t+1}
$$

The boundary conditions (22) can be represented by (24).

$$
\begin{bmatrix} 1 & -1 \\ \hline \mathbf{0} \\ \hline & 1 & -1 \end{bmatrix} \left\{ \begin{array}{c} \rho_{f_1} \\ \hline \rho_{f_2} \\ \vdots \\ \hline \rho_{f_L} \\ \hline \rho_{f_{L+1}} \end{array} \right\}_{j,k}^{t+1} = \kappa^2 \varepsilon_f \begin{bmatrix} 0 & -1 \\ \hline \mathbf{0} \\ \hline & 1 & -1 \end{bmatrix} \left\{ \begin{array}{c} \psi_1 \\ \hline \psi_2 \\ \vdots \\ \psi_L \\ \hline \psi_{L+1} \end{array} \right\}_{j,k}^{t+1} + \kappa^2 \varepsilon_f \begin{bmatrix} 1 & -1 \\ \hline \mathbf{0} \\ \hline & 1 & -1 \end{bmatrix} \left\{ \begin{array}{c} \phi_1 \\ \hline \phi_2 \\ \vdots \\ \phi_L \\ \hline \phi_{L+1} \end{array} \right\}_{j,k}^{t+1}
$$
(24)

After combining the boundary conditions to the governing equation, the final form can be obtained as (25).

$$
\therefore \left\{ \begin{array}{c} \rho_{f_1} \\ \hline \rho_{f_2} \\ \vdots \\ \hline \rho_{f_L} \\ \hline \rho_{f_{L+1}} \end{array} \right\}_{j,k}^{t+1} = \left\langle \frac{1}{D\Delta t} \begin{bmatrix} \mathbf{0} \\ \hline \mathbf{0} & \mathbf{I} & \mathbf{0} \\ \hline \mathbf{0} \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ \hline \left(\dfrac{\partial^2}{\partial \mathbf{x}^2} \right)_{2\sim L} \\ \hline \mathbf{0} \end{bmatrix} + \kappa^2 \begin{bmatrix} \mathbf{0} \\ \hline \mathbf{0} & \mathbf{I} & \mathbf{0} \\ \hline \mathbf{0} \end{bmatrix} + \begin{bmatrix} 1 & -1 \\ \hline \mathbf{0} \\ \hline & 1 & -1 \end{bmatrix} \right\rangle^{-1}
$$

$$
\cdot \left\langle \frac{1}{D\Delta t} \begin{bmatrix} \mathbf{0} \\ \hline \mathbf{0} & \mathbf{I} & \mathbf{0} \\ \hline \mathbf{0} \end{bmatrix} \left\{ \begin{array}{c} \rho_{f_1} \\ \rho_{f_2} \\ \vdots \\ \rho_{f_L} \\ \rho_{f_{L+1}} \end{array} \right\}_{j,k}^{t} + \kappa^2 \varepsilon_f \left(\begin{bmatrix} \mathbf{0} \\ \hline \left(\dfrac{\partial^2}{\partial \mathbf{x}^2} \right)_{2\sim L} \\ \hline \mathbf{0} \end{bmatrix} + \begin{bmatrix} 1 & -1 \\ \hline \mathbf{0} \\ \hline & 1 & -1 \end{bmatrix} \right) \left\{ \begin{array}{c} \phi_1 \\ \hline \phi_2 \\ \vdots \\ \phi_L \\ \hline \phi_{L+1} \end{array} \right\}_{j,k}^{t+1} + \kappa^2 \varepsilon_f \begin{bmatrix} 0 & -1 \\ \hline \mathbf{0} \\ \hline & 1 & -1 \end{bmatrix} \left\{ \begin{array}{c} \psi_1 \\ \hline \psi_2 \\ \vdots \\ \psi_L \\ \hline \psi_{L+1} \end{array} \right\}_{j,k}^{t+1} \right\rangle
$$
(25)

Table 2. Properties used in this study

Symbols	Values
D	1.0×10^{-17} [m]
κ	1.29×10^9 [$1/m$]
ε_f	6.94×10^{-10} [C/Vm]
h	1.0×10^{-8} [m]
λ_s	5.0×10^{-10} [m]
V	1.0 [V]

Utilizing the data in Tab. 2, transient behaviors of the charge density in the space are calculated and compared in Fig. 2 from the closed form solution and the numerical method proposed in this study. In general, the results from the analytic and numerical methods are in agreement for each time as shown in Tab. 3. In transient analyses, the behavior from the numerical method has a time-delay of 0.01 sec than that from the analysis. However, two behaviors from the numerical method and analysis in steady state are almost identical in time and space.

Table 3. Defined variances[*] for comparisons of numerical solutions on analytic results

Time [sec]	Values [dimensionless]
0.1	4.5042×10^{-2}
0.2	4.1764×10^{-2}
0.5	3.4787×10^{-2}
1.0	2.5344×10^{-2}
Steady State	2.3630×10^{-3}

[*]Defined variance: $s^2 = \dfrac{1}{n-1} \displaystyle\sum_{j=1}^{n} \left(\dfrac{x_{numerical\,j} - x_{analytic\,j}}{x_{analytic\,j}} \right)^2$
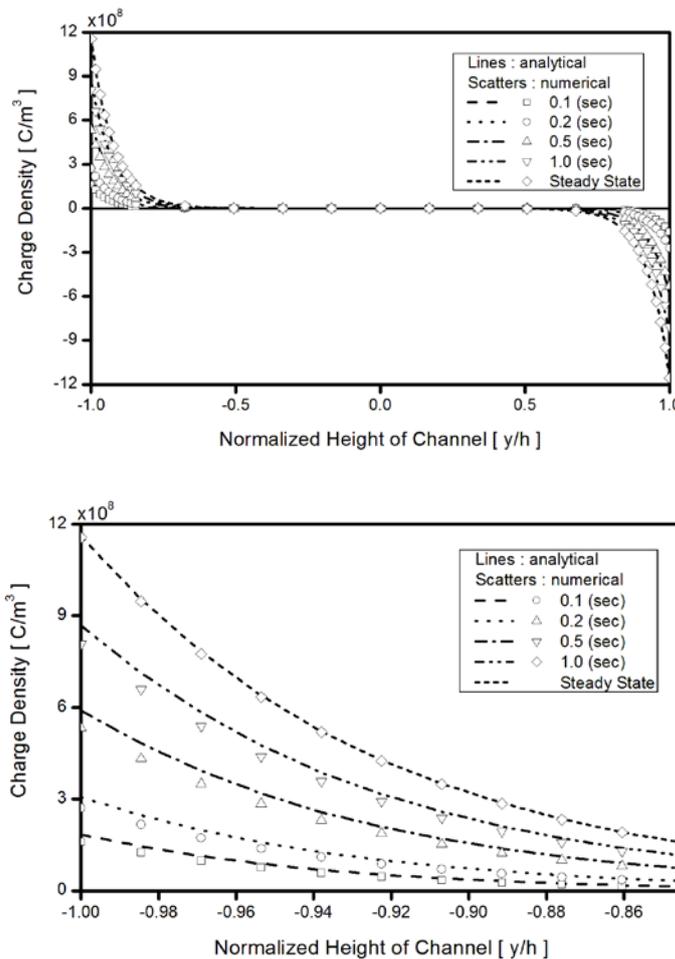
Figure 2. Comparisons of the analytic solutions on the numerical results (upper : total charge distribution and lower : magnified left side)

Conclusion

In electrokinetics, accurate prediction of the temporal and spatial charge density distributions near the ironic fluid and piezoelectric solid interface become important from the multi-physical point of view. Since the interested region for the changes of charge density distribution is the order of 10 Angstrom just near the interface, a very fine FD grid is required. When a constant fine space increment is applied for electrokinetic problems to obtain an accurate solution using FDM, the calculation process is inefficient.

In this study, FDM using the nonhomogeneous space increment is formulated and applied to an electrokinetic problem. A very fine FD grid is used for the interested diffuse layer just near the interface. At the same time, coarse FD grids are applied to the other ironic fluid path that is the order of several nanometers. In addition, a governing equation matrix is combined with a boundary condition matrix to construct an integrated equation matrix. As a result, PDE problems having mixed boundary conditions could be easily formulated and numerically solved by the proposed FDM using the simple matrix method.

In addition, the formulated FDM using the simple matrix method becomes a fully implicit form for space. Therefore, effects of amount of space increment on solution procedures could be minimized. The proposed method could be useful for obtaining transient solutions in electrokinetic problems particularly for bone remodeling, which electrical potentials are being changed temporally and spatially by the piezoelectricity of bone matrix.

References

Ahn, A. C., and Grodzinsky, A. J. (2009) Relevance of collagen piezoelectricity to "Wolff's Law": A critical review, *Medical Engineering and Physics* **31**, 733-74.

Bazant, M. J., Thornton, K., and Ajdari, A. (2004) Diffuse-charge dynamics in electrochemical systems, *Physical Review E* **70**, 021506–1–54.

Bassett, C. A. (1968), Biologic significance of piezoelectricity, *Calcified Tissue Research* **1**, 252-272.

Pienkowski, D., and Pollack, S. R. (1983) The origin of stress-generated potentials in fluid-saturated bone, *Journal of Orthpaedic Research* **1**, 30-41.

Zhang, D., Cowin, S. C., Weinbaum, S. (1997) Electrical signal transmission and gap junction regulation in a bone cell network: a cable model for an osteon, *Annals of Biomedical Engineering* **25**, 357-374.

Accurate Transient Response Analysis of Non-Classically Damped

Systems

Li Li, *Yujin Hu, and Ling Ling

School of Mechanical Science and Engineering, Huazhong University of Science and Technology, China.

*Corresponding author: yjhu@mail.hust.edu.cn

Abstract

It is difficult, or even unnecessary, to obtain all the modes of a large-scaled model. Thus, the modal truncation error is generally introduced and the quality of the responses may be adversely affected. Based on the Neumann series and the FFT technique, an accurate modal superposition method is presented to calculate the transient response of non-classically damped systems. The presented method maintains original-space without having to involve the state-space formula. The method is convergent if and only if all the complex modes whose resonant frequencies are less than the maximal sampling frequency of the FFT must be available. Finally, the applicability of the method is investigated using a simple numerical example with non-classical damping.

Keywords: Transient response analysis, Non-classically damped systems, FFT, Modal truncation error, Mode superposition method, Krylov subspace

Introduction

The purpose of a transient response analysis is to calculate the behaviour of a structural or mechanical system subjected to a time-varying forces. The inclusion of damping in the dynamic analyses of structural or mechanical systems has become an integral part of many design methodologies, including predicting vibration levels, transient responses, transmissibility and design problems dominated by energy dissipation.

In general, two different methods are used for the transient response analysis: direct transient response method (DTRM) and modal transient response method. The DTRM calculate dynamic responses by performing a direct numerical integration on complete coupled equations of motion at discrete times, typically with a fixed time step. In the most likely case for many engineering applications, the DTRM should be implemented for many time steps and large-scaled problems. Under such circumstance, it may be more effective by using reduced basis technique. Often the modes are used as the reduced basis vectors (known as the mode superposition method). The mode superposition method allows us to treat the equations of motion as a reduced-order form so that the step-by-step solution is less costly. It is should be noted that the quality of the responses depends on the number of modes involved. Although the accuracy of the calculated responses can be improved by increasing the number of modes, the convergence rate is very slow. Note that the eigenvalue solution is very computationally expensive, or even impossible, especially for large-scaled problems. Many approaches (Huang et al., 1997; Palmeri and Lombardo, 2011; D'Aveni and Muscolino, 2001; Besselink et al., 2013; Qu, 2007) were presented to deal with the modal truncation problems. However, these correction

approaches are only restricted to the case of undamped or proportionally damped systems. In general, proportional damping means that energy dissipation is almost uniformly distributed throughout the mechanical system. There is no any physical reason why the proportional damping must be satisfied. In practice, mechanical systems with significantly different levels of energy dissipation are frequently encountered in dynamic designs. As can be seen from experimental data, physical system produces complex modes and therefore no physical system is strictly proportionally damped system. To this end, the concern of this study is the non-classically damped system. Several approximation techniques are developed to efficiently calculate the responses of non-classically damped systems. Among these approximation methods, the most common method is so-called the proportional approximation method (PAM), which is simply to ignore the off-diagonal (coupling) elements of the transformed damping matrix. It was shown that light damping, diagonal dominance of the transformed damping matrix and good separation of the normal modes (these conditions are once believed to produce small errors) are not sufficient conditions for the accuracy of the PAM. Although the PAM is a powerful approximate method, the results of the PAM are not always with acceptable accuracy. The accurate responses may be obtained by using the complex modal superposition method. Although these correction approaches used in undamped systems can be extended to non-proportionally damped systems based on the state-space formula, these state-space approaches are usually time-consuming since its size is two times the size of the original space and lack the physical insight provided by the superposition of the modes of the equation of motion in physical space. Note that the complex modes may be also efficiently calculated using the computational methods in the original space (Fischer, 2000; Holz et al., 2004; Adhikari, 2011; Rajakumar, 1993; Lee et al., 1998). Some of these original space based approaches have been programmed in famous softwares [see, e.g., Nastran (Komzsik, 2001)]. Some works (Li et al., 2014c; Li et al., 2014b; Li et al., 2013) were therefore developed to eliminate the complex modal truncation error of the frequency responses of damped systems without having to involve the state-space formula. It is shown (Li et al., 2014c) that the complex modal truncation error can be exactly expressed as a power-series expansion in terms of the available modes and system matrices and a hybrid expansion method (HEM) is presented to calculate the frequency responses of non-classically damped systems. Complex modes can be also recently shown to transform any viscously damped system with N DOF into N independent second-order equations [see, e.g., (Kawano et al., 2013; Morzfeld et al., 2011; Ma et al., 2010) for details].

This paper presents an accurate method to calculate the transient response of non-classically damped systems based on the Neumann series and the FFT technique. The method maintains original-space without having to involve the state-space formula so that it is efficient in computational effort and storage capacity. The method is convergent if and only if all the complex modes whose resonant frequencies are less than the maximal sampling frequency of the FFT must be available. Finally, it will be shown by a numerical example that, the proposed method can show a good agreement with the exact responses.

Transient response analysis

The equation of motion of an N DOF linear damped system with zero initial condition appears as the following matrix form

$$\mathbf{M}\ddot{\mathbf{x}}(t) + \mathbf{C}\dot{\mathbf{x}}(t) + \mathbf{K}\mathbf{x}(t) = \mathbf{f}(t) \tag{1}$$

Here $\mathbf{M}$, $\mathbf{C}$ and $\mathbf{K}$ are real mass, viscous damping and stiffness matrices, respectively. $\mathbf{x}(t)$ $\dot{\mathbf{x}}(t)$ $\ddot{\mathbf{x}}(t)$ and $\mathbf{f}(t)$ are displacement, velocity, acceleration and force, respectively. In this paper, assume that $\mathbf{K}$ is a positive definite symmetric matrix, $\mathbf{M}$ and $\mathbf{C}$ are symmetric matrices.

The time domain equation of motion may be cast into a frequency domain form by using the Fourier transform technique

$$\left(-\omega^2\mathbf{M} + i\omega\mathbf{C} + \mathbf{K}\right)\mathbf{X}(\omega) = \mathbf{F}(\omega) \tag{2}$$

where

$$\mathbf{F}(\omega) = F\left[\mathbf{f}(t)\right] \text{ and } \mathbf{X}(\omega) = F\left[\mathbf{x}(t)\right] \tag{3}$$

Here $F[\ ]$ denote the Fourier transform and ω is the circular (angle) frequency. The form can be given under the assumption that the complex input forcing can be interpolated by trigonometric polynomials. In practice, we usually need to find the frequency spectra of excitation by using the Fourier transform.

The transient response can be then obtained by using the inverse Fourier transform, that is

$$\mathbf{x}(t) = F^{-1}\left[\mathbf{X}(\omega)\right] \tag{4}$$

In practice, a general analytical loading function cannot be easily obtained. It means that the discrete Fourier transform and inverse discrete Fourier transform algorithms given below should be used.

$$\begin{cases} \mathbf{X}_k = \sum_{n=0}^{N_{FT}-1} \mathbf{x}_n \exp\left(\dfrac{-2\pi ikn}{N_{FT}}\right) & \forall k = 0,1,2,\dots,N_{FT}-1 \\[4mm] \mathbf{x}_n = \dfrac{1}{N_{FT}} \sum_{k=0}^{N_{FT}-1} \mathbf{X}_k \exp\left(\dfrac{2\pi ikn}{N_{FT}}\right) & \forall n = 0,1,2,\dots,N_{FT}-1 \end{cases} \tag{5}$$

Here N_{FT} is a number of sample points. $\mathbf{x}_n$ are the elements of discrete time displacements and $\mathbf{X}_k$ are the elements of frequency spectrums of the discrete time series $\{\mathbf{x}_n\}$. Note the forcing samples should be obtained using the discrete Fourier transform [Equation (3)] and the transient response [Equation (4)] can be obtained by using inverse discrete Fourier transform once the frequency spectrums are calculated by solving Equation (2). The inverse Fourier transform procedure is defined using a positive sign in the exponential term and can be efficiently calculated using the *inverse fast Fourier transform* (IFFT) algorithm, which has been developed into a mature technology applied

successfully to calculate the displacement both in the frequency and time domain [for detail discussions on this aspect can be found in (Barkanov et al., 2003; Brigham, 1988; Duhamel and Vetterli, 1990)].

Accurate calculation of frequency spectrums

The eigenvalue problem can be written in matrix form as

$$\left(\lambda_j^2 \mathbf{M} + \lambda_j \mathbf{C} + \mathbf{K} \right) \boldsymbol{\varphi}_j = \mathbf{0} \quad \forall \, j = 1,\, 2,\, \ldots, 2N \tag{6}$$

Here λ_j and $\boldsymbol{\varphi}_j$ denote the jth eigenvalue and eigenvector (complex mode shape). Suppose these eigenvalues are ordered following increasing magnitude of imaginary parts. Assume these eigenvalues are distinct, the frequency spectrums can be calculated using the complex mode superposition method as

$$\mathbf{X}(\omega) = \sum_{j=1}^{2N} \frac{\boldsymbol{\varphi}_j^T \mathbf{F}(\omega) \boldsymbol{\varphi}_j}{\left(i\omega - \lambda_j \right) \theta_j} \ \text{ with } \ \boldsymbol{\varphi}_i^T \left(2\lambda_i \mathbf{M} + \mathbf{C} \right) \boldsymbol{\varphi}_i = \theta_j \tag{7}$$

Note the parameter θ_j can be chosen to be unity by normalizing eigenvectors. The method requires that all the modes should be available to obtain an exact response. Often only a few lower modes are considered in practical analysis. Suppose the lower L pairs of complex modes are available, the frequency spectrums are usually calculated in the following way with a modal truncation error involved

$$\mathbf{X}(\omega) = \mathbf{X}_{\text{MDM}}(\omega) + \mathbf{E}_{\text{MDM}}(\omega) \tag{8}$$

in which

$$\mathbf{X}_{\text{MDM}}(\omega) = \sum_{j=1}^{2L} \frac{\boldsymbol{\varphi}_j^T \mathbf{F}(\omega) \boldsymbol{\varphi}_j}{\theta_j \left(i\omega - \lambda_j \right)} \tag{9}$$

$$\mathbf{E}_{\text{MDM}}(\omega) = \sum_{j=2L+1}^{2N} \frac{\boldsymbol{\varphi}_j^T \mathbf{F}(\omega) \boldsymbol{\varphi}_j}{\theta_j \left(i\omega - \lambda_j \right)} \tag{10}$$

Equation (10) is known as the mode displacement method (MDM). Although the MDM is an efficient approximate method, the results of the MDM are not always with acceptable accuracy since the modal truncation error given by Equation (11) is introduced. Sometimes, the MDM will lead to misleading results [see, e.g., in (Li et al., 2014c; Li et al., 2014a)]. Recently, some reduced basis techniques were developed for the frequency response analysis [see e.g., (Hetmaniuk et al., 2012; Freund, 2003; Bai and Su, 2005; Hetmaniuk et al., 2013; Rumpler et al., 2014) for details].

Next, we present an accurate mode superposition method to calculate the frequency spectrums.

Theorem. *Suppose the damped system (1) only has distinct eigenvalues and the sample frequencies ω satisfy the convergence condition*

$$\left| \omega \right| < \left| \lambda_{2L+1} \right| \tag{11}$$

Then the frequency spectrums can be given by

$$\mathbf{X}(\omega) = \tilde{\mathbf{X}}(\omega) + \tilde{\mathbf{E}}(\omega) \tag{12}$$

where

$$\tilde{\mathbf{X}}(\omega) = \sum_{j=1}^{2L} \frac{\boldsymbol{\varphi}_j^T \mathbf{F}(\omega) \boldsymbol{\varphi}_j}{\theta_j (\mathrm{i}\omega - \lambda_j)} + \sum_{r=1}^{h} (\mathrm{i}\omega)^{r-1} \left[\mathbf{E}_r(\omega) + \sum_{j=1}^{2L} \frac{\boldsymbol{\varphi}_j^T \mathbf{F}(\omega) \boldsymbol{\varphi}_j}{\theta_j \lambda_j^r} \right] \tag{13}$$

$$\tilde{\mathbf{E}}(\omega) = \sum_{j=2L+1}^{2N} \left(\frac{\mathrm{i}\omega}{\lambda_j} \right)^h \frac{\boldsymbol{\varphi}_j^T \mathbf{F}(\omega) \boldsymbol{\varphi}_j}{\theta_j (\mathrm{i}\omega - \lambda_j)} \tag{14}$$

with

$$\begin{cases} \mathbf{E}_1(\omega) = \mathbf{K}^{-1}\mathbf{F}(\omega) \\ \mathbf{E}_2(\omega) = -\mathbf{K}^{-1}\left[\mathbf{C}\mathbf{E}_1(\omega)\right] \\ \mathbf{E}_r(\omega) = -\mathbf{K}^{-1}\left[\mathbf{C}\mathbf{E}_{r-1}(\omega) + \mathbf{M}\mathbf{E}_{r-2}(\omega)\right] \quad \forall r > 2 \end{cases} \tag{15}$$

Proof. The frequency responses can be exactly expressed as the lower available modes and system matrices in terms of Neumann series expansion (Li et al., 2014c)

$$\mathbf{X}(\omega) = \sum_{j=1}^{2L} \frac{\boldsymbol{\varphi}_j^T \mathbf{F}(\omega) \boldsymbol{\varphi}_j}{\theta_j (\mathrm{i}\omega - \lambda_j)} + \sum_{r=1}^{\infty} (\mathrm{i}\omega)^{r-1} \left[\mathbf{E}_r(\omega) + \sum_{j=1}^{2L} \frac{\boldsymbol{\varphi}_j^T \mathbf{F}(\omega) \boldsymbol{\varphi}_j}{\theta_j \lambda_j^r} \right] \tag{16}$$

In practice, a few terms in power-series expressed by the second term on the right-hand of Equation (17) are considered for suitable accuracy requirements. Suppose the first *h* power-series terms are retained, the frequency response can be expressed as

$$\mathbf{X}(\omega) \approx \tilde{\mathbf{X}}(\omega) = \sum_{j=1}^{2L} \frac{\boldsymbol{\varphi}_j^T \mathbf{F}(\omega) \boldsymbol{\varphi}_j}{\theta_j (\mathrm{i}\omega - \lambda_j)} + \sum_{r=1}^{h} (\mathrm{i}\omega)^{r-1} \left[\mathbf{E}_r(\omega) + \sum_{j=1}^{2L} \frac{\boldsymbol{\varphi}_j^T \mathbf{F}(\omega) \boldsymbol{\varphi}_j}{\theta_j \lambda_j^r} \right] \tag{17}$$

The series-truncation error of Equation (18) is introduced and given by

$$\tilde{\mathbf{E}}(\omega) = \mathbf{X}(\omega) - \tilde{\mathbf{X}}(\omega) = \sum_{r=h+1}^{\infty} (\mathrm{i}\omega)^{r-1} \left[\mathbf{E}_r(\omega) + \sum_{j=1}^{2L} \frac{\boldsymbol{\varphi}_j^T \mathbf{F}(\omega) \boldsymbol{\varphi}_j}{\theta_j \lambda_j^r} \right] \tag{18}$$

Here we introduce an important property between system matrices and complex modes (Li et al., 2014c)

$$\mathbf{E}_r(\omega) + \sum_{j=1}^{2L} \frac{\boldsymbol{\varphi}_j^T \mathbf{F}(\omega) \boldsymbol{\varphi}_j}{\theta_j \lambda_j^r} = -\sum_{j=2L+1}^{2N} \frac{\boldsymbol{\varphi}_j^T \mathbf{F}(\omega) \boldsymbol{\varphi}_j}{\theta_j \lambda_j^r} \tag{19}$$

Substituting Equation (20) into Equation (19) yields

$$\tilde{\mathbf{E}}(\omega) = -\sum_{j=2L+1}^{2N} \sum_{r=h+1}^{\infty} \left(\frac{\mathrm{i}\omega}{\lambda_j} \right)^r \frac{\boldsymbol{\varphi}_j^T \mathbf{F}(\omega) \boldsymbol{\varphi}_j}{\mathrm{i}\omega\theta_j} \tag{20}$$

If the convergence condition (12) is satisfied, one obtains

$$|\omega| < |\lambda_j| \quad \forall j = 2L+1,\ 2L+2,\ ...,\ 2N \tag{21}$$

Then the series-truncation error can be given by Equation (15) by using the theory of the geometric sequence, and the theorem is proved.

Remark 1. The method is convergent if and only if all the complex modes whose eigenvalues are less than the maximal sampling frequency of the FFT are available. Equation (14) can be used in the dynamic analysis of practical problems since the number h of correct terms is usually very small. Since the power-series expansion is truncated, the series-truncation error given by Equation (15) is introduced. When the convergence condition is satisfied, the errors can be decreased with the number h is increased. By comparing Equations (11) and (15), it is clearly shown that the frequency response obtained by Equation (14) can improve the accuracy of the response calculated by MDM if the sample frequency is at 0-$|\lambda_{L+1}|$ rad/s.

Remark 2. One of the most robust approaches to obtain these vectors $\mathbf{E}_r(\omega)$ (known as the Krylov vectors) is to firstly compute a matrix factorization (e.g., $\mathbf{LDL}^T$ factorization) of the sparse stiffness matrix $\mathbf{K}$. Note it only need to be obtained once for different sample frequencies. Then vectors $\mathbf{E}_r(\omega)$ can be determined by an iteration process in terms of forward and backward substitutions. In general, the computational cost of forward and backward substitutions is much smaller than the matrix factorization.

Criteria. The number h for any simple frequency ω can be determined if the following inequality is satisfied.

$$\|\mathbf{S}(h)\| + \|\mathbf{S}(h-1)\| < \varepsilon \quad \forall h > 1 \tag{22}$$

in which

$$\mathbf{S}(h) = (\mathrm{i}\omega)^{h-1}\left[\mathbf{E}_h(\omega) + \sum_{j=1}^{2L} \frac{\boldsymbol{\varphi}_j^T \mathbf{F}(\omega)\boldsymbol{\varphi}_j}{\theta_j \lambda_j^h} \right] \tag{23}$$

Here the parameter ε is the given accuracy tolerance.

Numerical Example

To illustrate this new method, we consider a simple there DOF damped system with the mass, damping and stiffness matrices given by

$$\mathbf{M} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 10^2 & 0 \\ 0 & 0 & 10^4 \end{pmatrix},\ \mathbf{C} = 10^3 \times \begin{pmatrix} 5 & 0 & -1 \\ 0 & 1 & -1 \\ -1 & -1 & 3 \end{pmatrix},\ \mathbf{K} = 10^8 \times \begin{pmatrix} 2 & 0 & -2 \\ 0 & 2 & -2 \\ -2 & -2 & 9 \end{pmatrix}$$

The complex eigenvalues are obtained as: $-0.2475\pm222.46\mathrm{i}$, $-5.0526\pm1421.4\mathrm{i}$ and $-2499.8\pm13920\mathrm{i}$. The excitation point is located at the first DOF and the force is shown in Figure 1. The time step is chosen as $\Delta t=0.005$ seconds, which means that the

maximum simple frequency is 400π rad/s. In view of the convergence condition (12), the first pair of complex modes must be included to calculate the transient responses. The MDM is also considered to calculate the responses by using the first pair of complex modes. Figure 2 shows the responses at the first DOF (here N_{FT}=512). In this case, the responses calculated by the MDM is misleading and meaningless. The proposed method can show a good agreement with the exact responses by considering a few numbers of the correction terms.
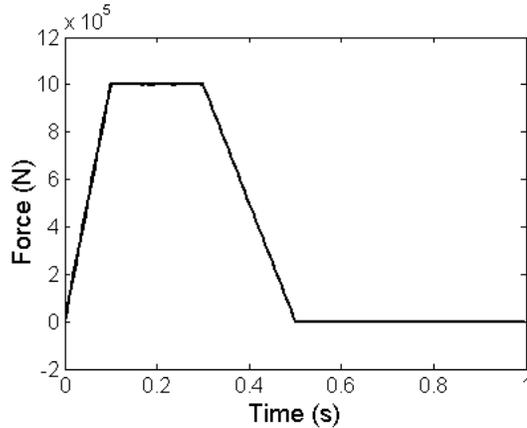
Figure 1. Applied force.   **Figure 2. Transient response.**
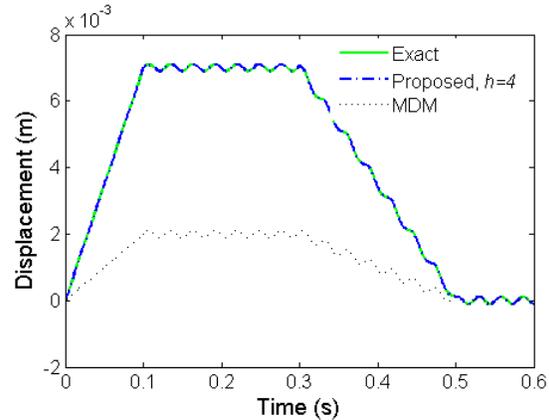
Conclusions

This paper consider the transient response analysis of non-classically damped systems. When the mode superposition method is used to calculate transient response, the modal truncation error is generally introduced since it is difficult, or even unnecessary, to obtain all the modes of a large-scaled model. An accurate modal superposition method is presented to calculate the transient response of non-classically damped systems based on the Neumann series and the FFT technique. The method maintains original-space without having to involve the state-space formula. The method can converge to exact results if and only if all the complex modes whose resonant frequencies are less than the maximal sampling frequency of the FFT must be available. The applicability of the method is investigated using a simple numerical example with non-classical damping. It is shown that, the responses calculated by the MDM is misleading and meaningless and the proposed method can show a good agreement with the exact responses.

Acknowledgments

References

Adhikari S. (2011) An iterative approach for nonproportionally damped systems, *Mechanics Research Communications* **38**, 226-230.

Bai Z and Su Y. (2005) Dimension reduction of large-scale second-order dynamical systems via a second-order Arnoldi method, *SIAM Journal on Scientific Computing* **26**, 1692-1709.

Barkanov E, Hufenbach W and Kroll L. (2003) Transient response analysis of systems with different damping models, *Computer Methods in Applied Mechanics and Engineering* **192**, 33-46.

Besselink B, Tabak U, Lutowska A, van de Wouw N, Nijmeijer H, Rixen D, Hochstenbach M and Schilders W. (2013) A comparison of model reduction techniques from structural dynamics, numerical mathematics and systems and control, *Journal of Sound and Vibration* **332**, 4403-4422.

Brigham EO. (1988) *The Fast Fourier Transform and Its Applications,* Englewood Cliffs, NJ, Prentice-Hall.

D'Aveni A and Muscolino G. (2001) Improved dynamic correction method in seismic analysis of both classically and non-classically damped structures, *Earthquake Engineering and Structural Dynamics* **30**, 501-517.

Duhamel P and Vetterli M. (1990) Fast Fourier transforms: a tutorial review and a state of the art, *Signal Processing* **19**, 259-299.

Fischer P. (2000) Eigensolution of nonclassically damped structures by complex subspace iteration, *Computer Methods in Applied Mechanics and Engineering* **189**, 149-166.

Freund RW. (2003) Model reduction methods based on Krylov subspaces, *Acta Numerica* **12**, 267-319.

Hetmaniuk U, Tezaur R and Farhat C. (2012) 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**, 1636-1662.

Hetmaniuk U, Tezaur R and Farhat C. (2013) An adaptive scheme for a class of interpolatory model reduction methods for frequency response problems, *International Journal for Numerical Methods in Engineering* **93**, 1109-1124.

Holz UB, Golub GH and Law KH. (2004) A subspace approximation method for the quadratic eigenvalue problem, *SIAM Journal on Matrix Analysis and Applications* **26**, 498-521.

Huang C, Liu Z-S and Chen S-H. (1997) An accurate modal method for computing response to periodic excitation, *Computers & Structures* **63**, 625-631.

Kawano DT, Morzfeld M and Ma F. (2013) The decoupling of second-order linear systems with a singular mass matrix, *Journal of Sound and Vibration* **332**, 6829-6846.

Komzsik L. (2001) *MSC/Nastran 2001 Numerical Methods User's Guide,* Santa Ana, MSC.Software Corporation.

Lee I-W, Kim M-C and Robinson A. (1998) Efficient solution method of eigenproblems for damped structural systems using modified Newton-Raphson technique, *Journal of Engineering Mechanics* **124**, 576-580.

Li L, Hu YJ and Wang XL. (2013) Improved approximate methods for calculating frequency response function matrix and response of MDOF systems with viscoelastic hereditary terms, *Journal of Sound and Vibration* **332**, 3945-3956.

Li L, Hu YJ and Wang XL. (2014a) Accurate method for harmonic responses of non-classically damped systems in the middle frequency range, *Journal of Vibration and Control* **in press**.

Li L, Hu YJ and Wang XL. (2014b) Eliminating the modal truncation problem encountered in frequency responses of viscoelastic systems, *Journal of Sound and Vibration* **333**, 1182-1192.

Li L, Hu YJ, Wang XL and Lü L. (2014c) A hybrid expansion method for frequency response functions of non-proportionally damped systems, *Mechanical Systems and Signal Processing* **42**, 31-41.

Ma F, Morzfeld M and Imam A. (2010) The decoupling of damped linear systems in free or forced vibration, *Journal of Sound and Vibration* **329**, 3182-3202.

Morzfeld M, Ma F and Parlett BN. (2011) The transformation of second-order linear systems into independent equations, *SIAM Journal on Applied Mathematics* **71**, 1026-1043.

Palmeri A and Lombardo M. (2011) A new modal correction method for linear structures subjected to deterministic and random loadings, *Computers & Structures* **89**, 844-854.

Qu Z-Q. (2007) Adaptive mode superposition and acceleration technique with application to frequency response function and its sensitivity, *Mechanical Systems and Signal Processing* **21**, 40-57.

Rajakumar C. (1993) Lanczos algorithm for the quadratic eigenvalue problem in engineering applications, *Computer Methods in Applied Mechanics and Engineering* **105**, 1-22.

Rumpler R, Göransson P and Deü JF. (2014) 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**, 759-784.

Reliability-Based Study of Well Casing Strength Formulation

L.P. Gouveia[1], *E.T. Lima Junior[1], J.P.L. Santos[1], J.L.R. Anjos[2] and W.W.M. Lira[1]

[1]Technology Center, Federal University of Alagoas, Maceió, Brazil.
[2]Research and Development Center, Petrobras, Rio de Janeiro, Brazil.

*Corresponding author: limajunior@lccv.ufal.br

Abstract

The increasing development of materials technology and the consequent rise in complexity of structures, demand a proper knowledge of the safety levels involved in the design. The reliability techniques applied to structural analysis allow dealing with the uncertainties inherent in the design of structures, so that the failure probability can be predicted since the design stage. In this context, the design variables are described as random variables, within the choice of an adequate distribution model to represent data is required. In an oil and gas well scenario, the casing design is a crucial stage of the whole project, representing the major structural elements responsible to maintain the well integrity through its lifetime, allowing adequate production activities. The occurrence of failures in casing systems can lead to irreversible safety problems in well operation. For the evaluation of the strength of the tubes used in casing design, the code API 5C3:1994 is widely applied by several companies overall. This deterministic formulation cannot deal with the uncertainties associated with the tube manufacturing process, as variations in geometrical and mechanical properties. This paper addresses the analysis of casing strength in a reliability-based approach, regarding the failure modes usually verified in well casing design. The reliability analysis is performed by the Monte Carlo simulation and the First/Second Order transformation methods (FORM/SORM). The safety levels associated to the referred formulation are estimated and discussed.

Keywords: OCTG, Well casing design, Structural reliability, Burst, Collapse

Introduction

Casing systems in wells play an important role as the major structural system that keeps the well integrity since the drilling and along its lifetime. The main elements in the casing are the tubulars and connections between them, which have to be designed in order to support the external loadings that they are subjected. In offshore well construction, which includes drilling, casing and completion stages, the casing system may represent around 15% to 20% of the total cost. The search for oil and gas in increasingly higher depths exposes the casing to extreme conditions, including high levels of pressure and temperature, besides chemical attack as, for instance, corrosion due to hydrogen sulfide gas. In this scenario, the adequate balance between cost and safety levels has to be reached in the design, and the structural reliability theory can assist the designer in this decision-making process.

The main loadings experienced by casing in vertical wells are represented by internal pressure, external pressure and axial force. The failure modes associated to internal and external pressure, so-called burst and collapse, respectively, are the governing failure modes in the problem. Tensile or compressive forces occur along the casing, but rarely induce tube failure alone. In the case of directional wells, not assessed in this work, additional torsion and bending effects can stand out. The strength equations of tubulars are posed in the code API 5C3:1994, which is widely applied by oil and gas companies. This formulation refers to a serviceability limit state (SLS) related to the elastic regime. Therefore, the tubes are designed to bear loads up to reaching the steel yielding

limit, disregarding its bearing capacity after this point. This is one of the reasons that makes this code seems conservative, underestimating the actual strength of the tube. The API 5C3:1994 normative code suggests the Barlow's equation for burst design strength, whose derivation is done by assuming thin wall hypothesis, which can be inadequate in some tubes commonly adopted in casing design. In terms of external pressure, the code provides four minimum collapse formulas for design strength, developed on the 1960's, being each formula suitable for a distinct D/t interval.

A new version of the code (API 5C3:2008, identical to ISO/TR 10400:2007) introduces the ultimate limit state (ULS) philosophy in casing design and suggests that reliability-based procedures can be performed in order to enhance the strength evaluation. An ultimate limit state equation is usually related with experimental rupture test results, since it tries to estimate the load at which the casing actually fails. Therefore, a proper equation is usually chosen by fitting experimental test data. Although this proposed improvements, the older version from 1994 is still the most adopted in design routines. The new paradigm of ULS associated to probabilistic design is slowly being introduced in the companies by consulting and research activities.

The increasing development of materials and structural modeling, and the consequent growing in complexity of structures, demand a proper knowledge of safety levels involved in the design. The structural reliability theory provides methods to evaluate these safety (or risk) levels, accounting for the uncertainties inherent to the design. In engineering applications, the uncertainties commonly verified relates to mechanical (material) parameters, as Young modulus or tensile strength, and dimensional parameters, as lengths and masses, for instance. It refers directly to non-uniformity on the manufacturing process of structural materials and elements. In the light of probability and statistics concepts, these uncertainties are modeled as random variables, and collected together in a framework of mathematical models that estimate the probability of failure associated to a specific failure mode defined by the user. Fundamentals and applications of the structural reliability theory can be found in Melchers (1999), Ang and Tang (2007), Ditlevsen and Madsen (2005), among others.

Specifically, in oil and gas industry, well casing design is related to steel tubular manufacturing, whose production quality and inspection procedures have been improved significantly in the last decades. The suggested casing strength formulas are dependent on the outside diameter (D) and wall thickness (t) – usually referred by the slenderness ratio D/t – and yield stress. In this case, the old version of API code recommends adopting conservative minimal or nominal values for these parameters and, additionally, in the design process, safety factors are applied to ensure implicitly a tolerable risk level (TRL). On the other hand, if the supplier guarantees accuracy in the manufacture and inspection processes, with lower dispersion levels in the tube performance values, and consequently the TRL can be kept by allowing the structural element to bear a higher load than the one predicted by the standard.

Some works in the literature address the recommended formulas from API 5C3:1994 and discuss its seeming conservative nature, besides proposing ULS formulations. Some works can be found in the literature with distinct ultimate limit state equation suggestions for burst [Klever and Stewart (1998); Klever (2010)] and collapse [Abbassian and Parfitt (1995); Klever and Tamano (2006); Tamano et al. (1985)]. As previously stated, new design codes also proposes to going beyond the elastic limits, enforcing this ultimate limit state philosophy.

The probabilistic analysis of casing design have been studied by some authors, since the 1990's. Adams et al. (1998) present the behavior of failure probabilities for API 5C3 strength in the ULS analysis provided by Tamano et al. (1985), and verify that it gives a wide range of variation for

failure probabilities over a D/t range, concluding that this behavior is not desirable and suggesting a new reliability-based method for collapse casing design. Ju et al. (1998) proposes a different formulation. The code API 5C3:2008 itself does a very similar development, but it adopts Klever and Tamano (2006) as the ultimate state limit equation suggesting a more robust probabilistic method for collapse casing design. Burres et al. (1998) propose an interesting discussion, working on the calibration of safety factors in design equations in order to reach a specific TRL.

This paper addresses two kinds of reliability analyses. The first one is done by verifying the probabilities of API 5C3:1994 design strength be exceeded when the maximum deterministic load that it was designed is reached, instead of the failure probability of the casing actually fails if the same load is achieved. It means that this paper is not going to analyze the probability of failure for an ultimate limit state, but it is going to do it for the design equation, which is really used in design procedures. This kind of analysis can be useful to verify the influence of the dispersion of design variables in each design equation, across the D/t range, and to check the probability of this design strength be exceeded, leading the casing tube to transcend the elastic limit. A second analysis is developed by addressing a hypothetical design scenario, including defined loading profiles, for which the probability failures are evaluated along the depth of the well. This application becomes useful when the results are compared to the safety factors adopted by each company for each failure mode in the deterministic design.

The First and Second Order Reliability Methods (FORM/SORM) and the Monte Carlo simulation are used in this paper to perform the probabilistic analysis. The evolution of safety levels implicitly associated to the referred equations across D/t and for distinct grades is investigated, and a performance comparison between these methods is carried out.

This work is divided in three main sections. The first one has a brief review of what is and how works a structural reliability analysis. The second one brings an overview of the recommended practices for casing burst and collapse design made by API 5C3:1994. Finally, in the last section, the concepts are combined and the simulations are presented, discussing the results.

Structural Reliability Analysis

Essentially, a structural reliability analysis needs a limit state function, some random variables and a reliability method. The limit state function must represent the problem which is going to be studied, in general, it gives positive values for safe events and negative values for failure events. The failure modes considered in this paper represent the safety margin of probabilistic API 5C3 casing design strength be exceeded by the deterministic corresponding strength. Equation (**Erro! Fonte de referência não encontrada.**) presents the failure mode adopted $G(X)$, where $R(X)$ is the resistance term and L is the load term.

$$G(X) = R(X) - L \tag{1}$$

where X is a vector containing the random variables. As the focus of this paper consists on the analysis of casing strengths, only the resistance term is going to be assumed as probabilistic.

Random variables must estimate the behavior of geometric/mechanical properties and other design variables related to the structural element, that influence the resistance term in limit state function. On the other hand, the load term in limit state function also could have random variables, which could be the self-weight, some external mechanical load or the one caused by temperature variation, for instance. The correlation between random variables also can be attached on reliability-based problems, although the literature states that adopting the variables as independent is a conservative procedure. In the following analysis, random variables are going to be independent, but the correlation between them can be adopted in future works.

The reliability-based method is going to link the limit state function with the random variables to compute a failure probability. For structural reliability the Monte Carlo Method, FORM and SORM are the most known ones [Melchers (1999)]. In the next sections, these methods are briefly detailed.

Monte Carlo Simulation

The method consists in generate n random scenarios to be tested in the limit state function, computing the number of failure events n_f (when $G(X) \leq 0$), and estimate the failure probability by $P_f = n_f/n$.

The random scenarios are defined by generating n aleatory values for each random variable assumed in the analysis. Therefore, the statistical characterization of each variable and a random number generator are required. An illustrative example of a Monte Carlo simulation is presented in Fig. **Erro! Fonte de referência não encontrada.**, in which a thousand events are generated. Each event is tested with the limit state function, where if the resistance (R) is higher than the load (L), there is a safe event, otherwise there is a failure event. In this hypothetical example R and L are Gaussian distributed random variables with means 115.0 and 90.0, and standard deviation equals to 4.0 and 10.0, respectively. It is usual to adopt the notation $R = N(115.0; 4.0)$ and $L = N(90.0; 10.0)$.
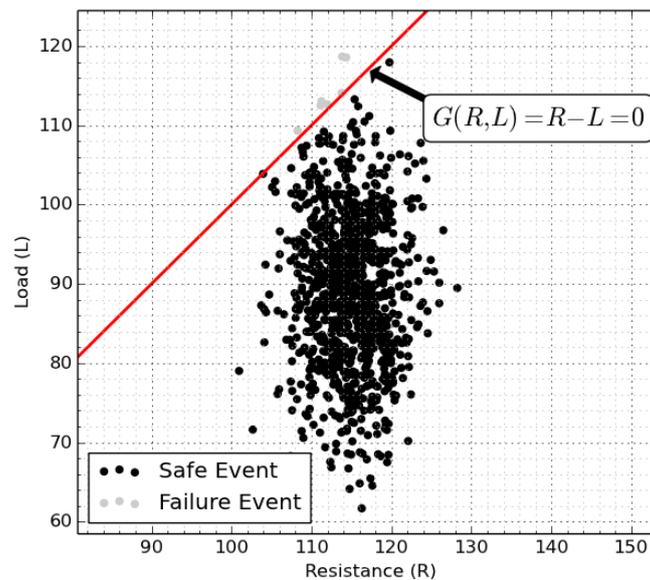
Figure 1. Monte Carlo illustrative example assuming R=N(115.0;4.0) and L=N(90.0;10.0)

By its nature, Monte Carlo provides very accurate results, since an adequate number of simulations is performed. However, this method may have issues with very low failure probabilities once it will need, at least, the inverse of the failure probability number of scenarios to possibly be capable to detect one failure event, i.e., if the problem has a probability of failure equals to 10^{-6}, a minimum of 10^{-6} scenarios has to be generated and simulated. It has to be regarded that the estimated value

P_f is also a random variable, and some expressions are proposed to define a minimum number of scenarios which leads to a desirable covariance of P_f. The computational cost in Monte Carlo simulations may increases substantially in some applications, being noted as a possible disadvantage. Some strategies of random data generation can be applied in order to improve performance in this method as, for instance, stratified sampling and other importance sampling procedures.

First/Second Order Reliability Methods

A reliability analysis problem can be mathematically expressed considering the limit state function ($G(X) = 0$) and the adopted random variables X, being P_f exactly evaluated by the integral:

$$P_f = P(G(X) \leq 0) = \int_{G(X) \leq 0} f_X(x) dx \tag{2}$$

in which $f_X(x)$ is the joint probability density function of the random variables X. However, depending on the number of random variables, this integral is not easy to solve and numerical approximations should be applied, where Monte Carlo simulation is an option. Transformation methods as the First Order Reliability Method (FORM), which one is analytically derived and iteratively solved, stand out as an interesting choice. The method consists in transforming all random variables (X) in its corresponding standardized normally distributed ones (U), this is done by first applying a normal tail approximation and then reducing them to standard normal probability distribution function. It is also necessary to rewrite the limit state function for this standard normal space ($G(U) = 0$). In this new space, the probability of failure concept can be associated with the shortest distance between the new adopted limit state function and the transformed random variables space origin. This distance is known as the reliability index β and its relation with P_f is provided by:

$$P_f = \Phi(-\beta) \tag{3}$$

The reliability analysis is posed as a nonlinear optimization problem, in which one wants to minimize the distance β subject to the constraint function $G(U) = 0$. The point U^* in which this condition is the most probable failure point, the so-called design point. Thus, the reliability index corresponds to the norm of the position vector of this point, i.e., $\beta = \|U^*\|$. The random variables transformation is made as suggested by Hasofer and Lind (1974). The limit state function is approximated by a first order Taylor series at the current search point from the iterative optimization problem. The optimization problem can be expressed as follows:

$$\beta = \min(\|U\|), \, constrained \, to \, G(U) = 0 \tag{4}$$

The algorithm HLRF (Hassofer, Lind, Rackwitz and Fiessler) is classically employed to the optimization problem solution. In general, in few iterations (less than 10) the convergence is reached. An advantage of this method is that it can be solved faster than Monte Carlo simulation, regarding a good level of accuracy, in many applications. Moreover, if the limit state function is linear on the random variables, these ones presenting Gaussian distribution, FORM results are exact. Figure 2 illustrates the procedure.
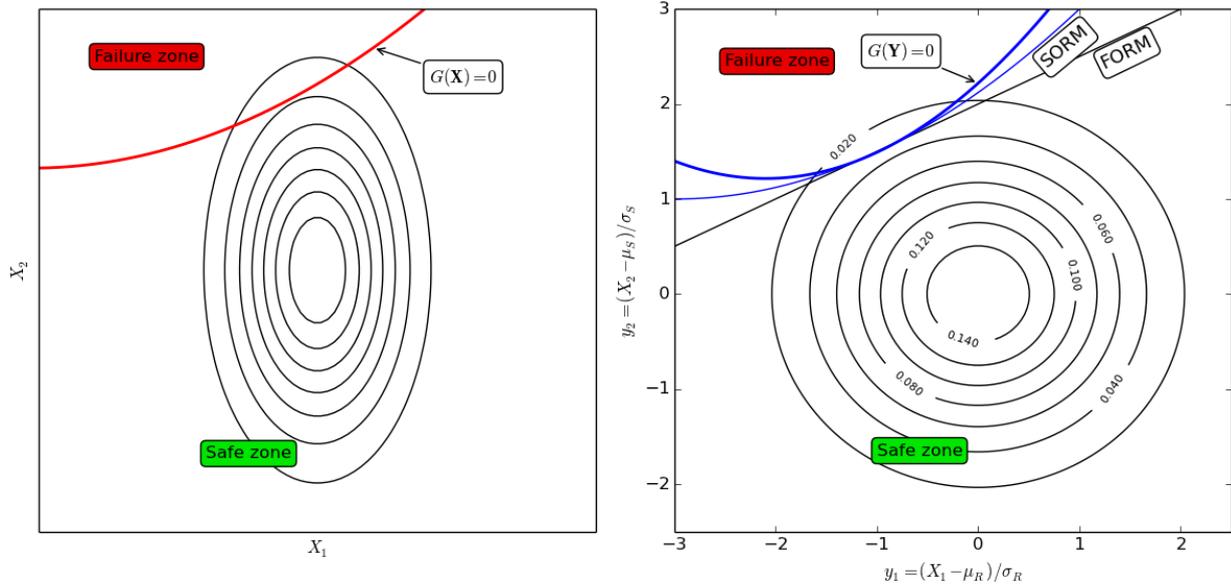
Figure 2. FORM illustrative Scheme: original and standard normal spaces

Another information extracted from FORM is the importance factor of each random variable, for the achieved failure probability. This information is associated with the position vector $\boldsymbol{U}^*$ and the partial derivatives $G(\boldsymbol{U})$ at this point. The importance factor give the influence of the random variable in the aleatory process.

It should be noted that, in the case of correlation between random variables, additional steps are necessary. In order to calculate the equivalent correlation coefficient for each pair of variables in the transformed normal space, the procedure proposed in Nataf (1962) can be applied. Moreover, this correlation has to be eliminated, so that the final transformation into standardized normal independent variables $\boldsymbol{U}$ can be performed. These procedures are detailed in Melchers (1999).

In some cases, in which the variables are tightly correlated, or present non-Gaussian distributions, or when the limit state function is strongly nonlinear, the use of a second order approximation of the limit state function can improve the accuracy of results obtained in the transformation method. This give rises to SORM (Second Order Reliability Method). This approximation demands more information over the limit state function, as its curvatures. The final approximation consists in a parabolic equation centered on the design point. In this work, the Breitung approximation is adopted [Breitung (1984)]. More details can be seen in Melchers (1999).

Casing Strength Formulation

The recommended practices for casing well design described by API 5C3:1994 are summarized in this section, focusing the axial, burst and collapse strengths.

Axial Strength

The axial strength that corresponds to a stress equal to the minimum yield strength, given as follows:

$$R_t = 0.7854(D^2 - d^2)Y_p \tag{5}$$

where:
R_t pipe body yield

Y_p material minimum yield strength
D specified outside diameter
d specified inside diameter

The axial force is the result of the balance between self-weight of the pipe and the pressure caused by the drilling fluid and other fluids from formation. As stated before in this text, it does not configure a governing failure mode by itself.

Burst Strength

The internal pressure that leads to a stress, on the inside wall, equals to the minimum yield strength. The failure mode associated is a brittle rupture of the tube. The equation is based on the Barlow's equation, suitable to thin wall tubes:

$$P_i = 0.875 \left(\frac{2\,Y_p\,t}{D} \right) \tag{6}$$

Where D is the outside diameter, t is wall thickness and Y_p the minimum yield strength of the steel. The reduction factor 0.875 refers to a tolerance of -12.5% in the wall thickness. This value is the allowable limit due non-uniformity in manufacture process, and is preconized by the code API 5CT:2010. This is one the reasons why the equation seems to be conservative.

Collapse Strength

When a pipe collapses due to external load, it changes the geometry to elliptical or other non-circular shape. It brings structural problems associated to loss of rigidity and local instability in the tubes, besides operational issues as blocking of passage of equipment into the tube. The external load is usually caused by pore-pressure, pressure from the drilling fluid, or fluid expansion due to temperature gradient. According to API 5C3:1994, four distinct casing slenderness D/t domains compose the collapse design strength. Yield strength collapse pressure formula (Eq. (7)) provides the load that generates minimum yield stress Y_p on the inside wall of the tube. This formula is achieved by means of Lamé's classical equation.

$$R_{cy} = 2Y_p \left[\frac{(D/t)-1}{(D/t)^2} \right] \tag{7}$$

Average plastic collapse pressure formula (Eq. (8)) was derived empirically from several collapse tests for casing tube grades K55, N80 and P110. This is the usual nomenclature for the steel which the casing tube has been made, where the first letter refers to its tensile strength and the following digits refers to its minimum yield stress. The data used by API 5C3:1994 authors was taken from a report made by a Workgroup composed by members from manufacturers and members from API itself. Collapse tests data were fitted separately for each grade, and then, constants A and B were empirically determined to generalize an average plastic collapse pressure formula. To obtain the minimum plastic collapse pressure formula a constant pressure for a particular grade, a constant C is subtracted from the average expression. This constant C is a tolerance limit and represents the conception that there is a 95% probability or confidence level that the collapse pressure will exceed the minimum stated with no more than 0.5% failures.

$$R_{cp} = Y_p \left[\frac{A}{(D/t)} - B \right] - C \tag{8}$$

Transition collapse pressure formula (Eq. (9)) overcomes an anomaly that happens between minimum plastic collapse formula and minimum elastic collapse formula: they do not intersect across the D/t range. Thus, this formula has been developed intersecting the D/t value where the

average plastic collapse pressure formula gives a collapse pressure of zero and is tangent to the minimum elastic collapse pressure.

$$R_{ct} = Y_p \left[\frac{F}{(D/t)} - G \right] \qquad (9)$$

Finally, the minimum elastic collapse pressure formula (Eq. (10)) was derived from theoretical elastic collapse pressure formula, resulting in the equation:

$$R_{ce} = \frac{46,95 \times 10^6}{(D/t) \times ((D/t)-1)^2} \qquad (10)$$

The D/t limits are the ones which define the collapse domains. They are determined by the intersection of the collapse pressure formulas described above and are shown below

$$(D/t)_{Yp} = \frac{\sqrt{(A-2)^2 + 8\left[B + \left(\frac{C}{Y_p}\right)\right]} + (A-2)}{2\left[B + \left(\frac{C}{Y_p}\right)\right]} \qquad (11)$$

$$(D/t)_{PT} = \frac{Y_p(A-F)}{C + Y_p(B-G)} \qquad (12)$$

$$(D/t)_{TE} = \frac{2+B/A}{3B/A} \qquad (13)$$

These D/t limits are dependent only on the yield stress (in psi) and must be calculated for each steel grade. Once they are determined, it is necessary to verify the casing design collapse domain by its thickness D/t. If casing D/t is lower than $(D/t)_{Yp}$, yield collapse pressure formula must be applied. If casing D/t is higher than $(D/t)_{Yp}$, but lower than $(D/t)_{PT,}$ minimum plastic collapse pressure formula must applied. If casing D/t is higher than $(D/t)_{PT}$, but lower than $(D/t)_{TE}$, transition collapse pressure formula must applied. If casing D/t is higher than $(D/t)_{TE}$, minimum elastic collapse pressure formula must be applied. The coefficients A, B, C, F and G are shown below:

$A = 2.8762 + 0.10679 \times 10^{-5} Y_p + 0.21301 \times 10^{-10} Y_p{}^2 - 0.53132 \times 10^{-16} Y_p{}^3$
$B = 0.026233 + 0.50609 \times 10^{-6} Y_p$
$C = -465.93 + 0.030867 Y_p - 0.10483 \times 10^{-7} Y_p{}^2 + 0.36989 \times 10^{-13} Y_p{}^3$

$$F = \frac{46.95 \times 10^6 \left[\frac{3(B/A)}{2+(B/A)}\right]^3}{Y_p \left[\frac{3(B/A)}{2+(B/A)} - (B/A)\right] \left[1 - \frac{3(B/A)}{2+(B/A)}\right]^2}$$

$$G = \frac{FB}{A}$$

The collapse resistance of casing in the presence of an axial stress is calculated by modifying the yield stress to an axial stress equivalent grade according to:

$$Y_{pa} = \left[\sqrt{1 - 0.75(S_a/Y_p)^2} - 0.5(S_a/Y_p)\right] Y_p \qquad (14)$$

where:
S_a axial stress (pounds per square inch)
Y_p minimum yield strength of the pipe

Y_{pa} yield strength of axial stress equivalent grade, pounds per square inch.

Analysis and Results

In order to associate the concepts presented in previous sections, some reliability analyses are performed as follows. The transformation methods FORM and SORM, besides crude Monte Carlo simulation are applied. The set of random variables contains the yield strength Y_p, the outer diameter D and the wall thickness t, and its statistical parameters are taken from the code ISO 10400:2007. The referred statistical database compiles several manufacturing production data, between 1977-2004, being representative of different manufacturing technologies and quality levels. The statistical parameters are evaluated by using the coefficients shown in Table 1, for the three steel grades adopted, K55, N80 and P110. The grades represents that the steel used has a minimum yield strength of 55000 psi, 80000 psi and 110000 psi, respectively. In this table, *mean* is equal to the actual mean value divided by the nominal value, and *COV* is the standard deviation divided by the actual mean value. According to the reference, the variables are normally distributed.

The tolerable failure probability values are not an unanimity over the scientifical/technical community in structural engineering in general. It depends on the class of the structure, the failure cost, among others. The implication of human lives and environmental risks are also determinant aspects on the definition of a required safety level. Recommendations on some normative codes just begin to appear, e.g., the ones based on JCSS (Joint Committee on Structural Safety) suggestions. In well design industry, it consists in a subject of relatively incipient discussion. In the present text, probabilities of failure higher than 10^{-3} are considered unallowable, based on technical literature for applications in engineering.

Table 1. Statistical coefficients used to characterize the random variables

Y_p		D		t	
mean	*COV*	*mean*	*COV*	*mean*	*COV*
K55					
1.23	0.0719				
N80		1.0059	0.00181	1.0069	0.0259
1.21	0.0511				
P110					
1.09	0.0377				

Collapse and Burst Design Formulation Analyses

In that follows, both for collapse and burst, the reliability evaluation uses a limit state function that assumes the strength as probabilistic and the load as deterministic. The probabilistic strength formula is obtained by adopting the variables as random in the API 5C3 design equation. The deterministic load is set equals to the value of API 5C3 design strength, calculated on the nominal values. It means that the failure probability achieved represents the probability of the design strength be exceeded if there is a load equal to the minimum strength currently used in design. Thus, the limit state function can be written as:

$$G(Y_p, D, t) = R_{conf}(Y_p, D, t) - L_{det} \tag{15}$$

It has to be noted that the correction factor 0.875 is not used in the term $R_{conf}(Y_p, D, t)$ in burst analysis. The variability of wall thickness is consistently treated here, by assuming it as a random variable. Figure 3 presents the results for collapse reliability analysis.
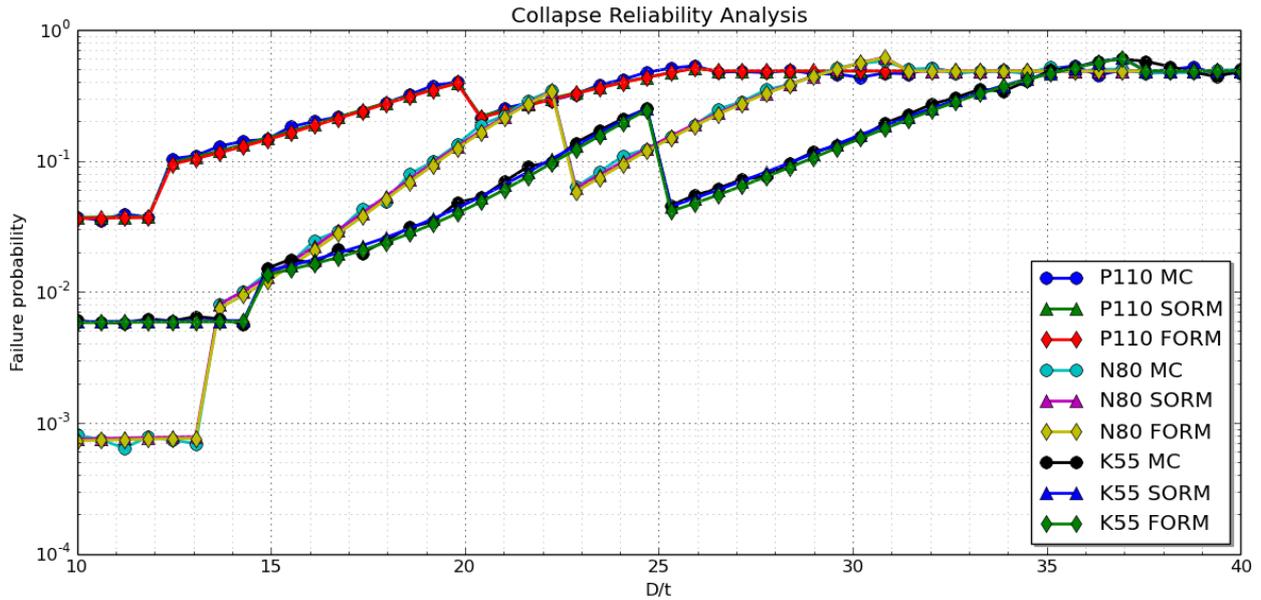
Figure 3. Collapse failure probabilities over D/t range

The four collapse domains described in the previous section present different failure probabilities over D/t range, besides the theoretical formulas give nearly constant failure probabilities (yield and elastic collapse pressure formulas) and the empirical formulas give variable failure probabilities (plastic and transition collapse pressure formulas). For the three grades it is noticed that the probability of failure grows as the casing tube D/t increases. This kind of behavior is not desirable in a design formulation, since it is expected constant safety level for all casing tubes designed by the same normative code. This is one of the main arguments posed by the committee which worked on the new version of the code (API 5C3:2008), proposing only one ULS formulation, and encouraging to use probabilistic analysis in design.

The analysis of burst formula is presented in Fig. 4.
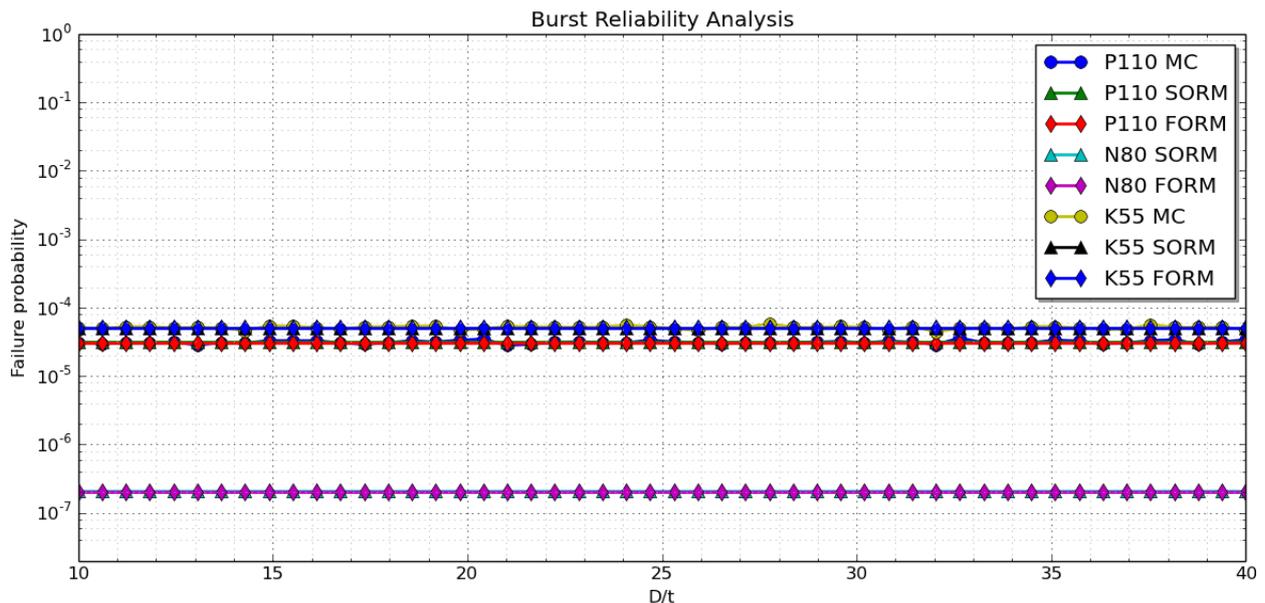
Figure 4. Burst failure probabilities over D/t range

In this case, the failure probabilities have a constant behavior over the D/t range. However, the failure probabilities are not constant when the casing grade varies. It is noticed that grade N80 has a much lower failure probability then the other two grades. This behavior may occur because *mean* and *COV* of Y_p vary with grade. However, the design formulation should have been calibrated to achieve a target reliability level considering the statistical data from production.

It should be noted that FORM results agree with SORM and Monte Carlo in both collapse and burst reliability analysis. The nature of these equations and its smooth nonlinearity contributes to this fact. It is possible that, in strongly nonlinear limit state functions, SORM provides quite different results. The maximum relative error observed between P_f values obtained by FORM and SORM is around XX% for collapse and XX% for burst. Monte Carlo is not compared with a numerical measurement error due its intrinsic random results, although the graphical visualization demonstrates a good agreement between Monte Carlo and the other reliability methods.

The importance factors obtained by FORM for collapse analysis are shown in the following Fig. 5.
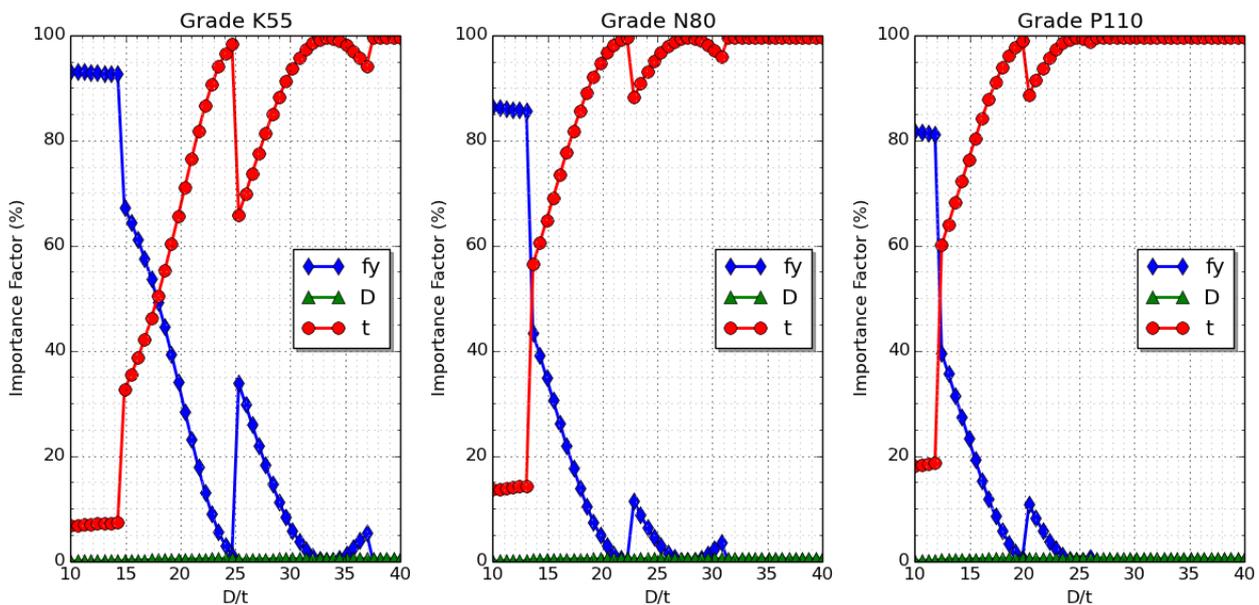
Figure 5. Importance factors over D/t range for collapse achieved probabilities of failure

For all grades, the most influent random variable in the process is the wall thickness, except for thick casing tubes in which the yield stress governs the probabilistic behavior. On the other hand, the diameter has a negligible influence in all these results. It means that considering it as deterministic will not affect significantly the failure probability values. These importance factors results are mainly impacted by the formulation used and by the adopted dispersion for each random variable, as it can be noticed in Table 1 that the diameter has the lowest *COV*.

FORM burst analysis provides the importance factors shown in Fig. 6. The yield stress is the most influent random variable in the achieved failure probability, followed by the wall thickness in all grades. Observing Table 1 and Fig. 6 it can be seen that the higher *COV* gives the higher importance factor for K55 grade. The others grades respect the following order. Once again, external diameter is the less important random variable, meaning that its dispersion is very small.
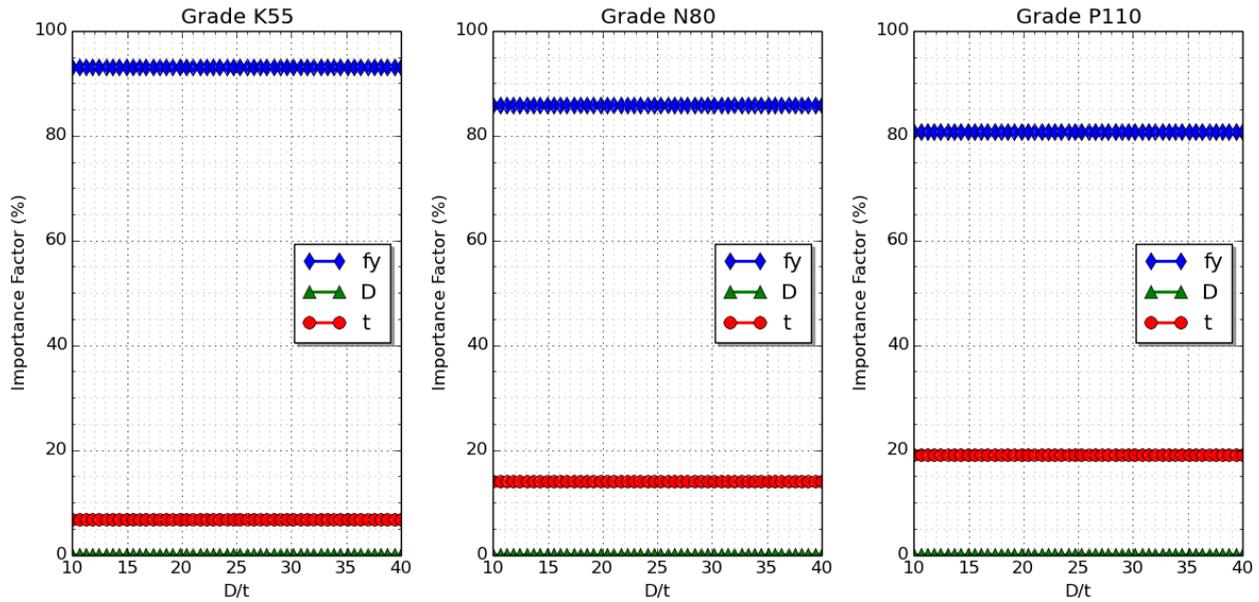
Figure 6. Importance factors over D/t range for burst achieved probabilities of failure

Analysis of an Extreme Design Scenario: Kick

In a casing design routine, the tubes are designed for different loading conditions throughout the well depth. Depending on the depth and the geomechanical conditions, extreme scenarios may occur along the drilling, casing, completion and production stages. These kind of scenarios has to be simulated in the well design. A *kick* situation is defined when a gas invades the drilling column, increasing drastically the expected internal pressure levels on the casing system, leading to possible accidents as a blow-out. For design purposes, it is considered that the last 2/3 of well depth are occupied by gas.

It is assumed a drilling of a 5700 m depth well, under 2000 m of water. The calculations of each pressure term are neglected, for sake of conciseness. The loading profile is shown in Fig. 7, in which is defined the differential pressure over the depth, resulting on a burst (internal pressure) failure mode overall. For this analysis is assumed a 10 3/4 in 85.3 lb/ft tube, which is widely applied in surface and intermediate casing structures. It has outer diameter of 10.75 in and wall thickness equals to 0.797 in.
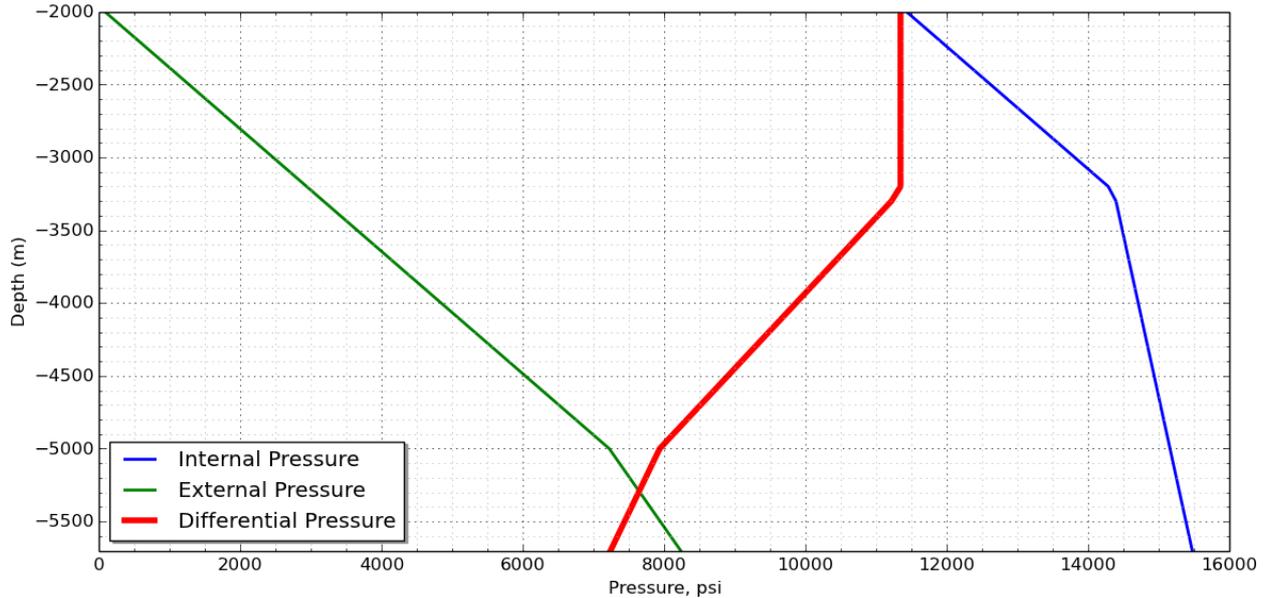
Figure 7. Loading Profile on the kick scenario

Regarding the reliability analysis, the limit state function adopted has the format:

$$G(Y_p, D, t) = R_{conf}(Y_p, D, t) - L_{det} \tag{16}$$

in which the load term is evaluated along the well depth according to the presented loading profile. The reliability evaluations are done every 100 m. The results are shown in Fig. 8, in which the failure probability values are log scaled.

As expected, the tube made of grade K55 reaches higher failure probability values. The severe values of pressure from 2000 m up to 4000 m leads to unallowable levels of P_f for this tube, showing its inadequacy for this scenario.

Considering that the differential pressure is constant up to around 3223 m, from which is considered the fluid inflow, the failure probabilities remain unchanged. From this point on, until 5700 m, it is observed a quasi-linear decrease of P_f values, referring to the reduction of applied pressure, due to the low specific weight of the invading fluid.
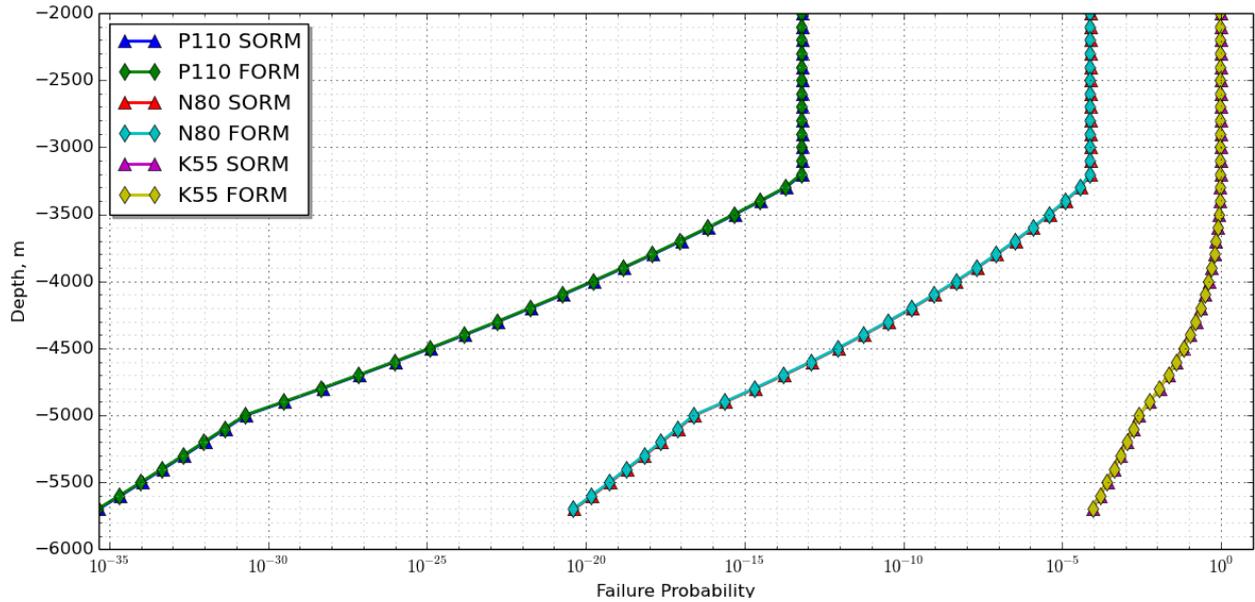
Figure 8. Evolution of failure probability along the depth

This kind of analysis allows to compare the P_f values in any point along the well depth with the safety factors usually employed in casing design. This discussion is not developed here, considering that these factors are defined by each oil company.

Conclusions

The classical casing collapse and burst strength equations are revisited in the light of a probabilistic approach. Moreover, the reliability analysis is applied to the verification of an extreme event kick scenario.

Regarding the analysis on the collapse design formulation, it is noticed that there are some high failure probability values associated to the collapse design strength, when a deterministic load equals to the minimum casing strength is considered. Moreover, the non-uniformity of the safety levels across the slenderness D/t is not a proper behavior, for structural design purposes. The importance factors values indicates that wall thickness is the most influent random variable in the achieved failure probability.

For burst analysis, moderate failure probability values are verified. The importance factors values indicates that the material yield limit is the most influent random variable in the achieved failure probability, followed by the wall thickness. The supposed conservative nature of the burst equation, posed by several authors in the literature, is apparently verified here.

The application of a probabilistic evaluation in the casing design practice can be done by procedures such as the scenario analysis presented. It brings robustness to the analysis, and assess the designer in decision-taking processes aiming both investment savings in simple wells and feasibility in complex wells.

In this context, the need of detailed analysis both on casing design formulation and about the non-deterministic nature of strength parameters stands out. The standardization codes and oil/gas companies are interested in these issues since the last two decades, and some scientific and technical publications has been developed. Some effort has also to be done in order to consider combined load cases in a probabilistic approach, focusing on the stochastic behavior of

environmental load scenario. Reliability-based analysis also proves to be useful for industry and designers as a device to identifying aspects in which the manufacturing process has to be improved in its accuracy and quality inspection.

It should be remarked that the results presented in this paper are only indicatives of the probabilistic behavior associated to the design formulations studied. The P_f values themselves have to be interpreted with caution, as they reflect the behavior of a specific statistical database, provided by the code ISO 10400:2007.

This research group is engaged in probabilistic analysis of combined failure modes for well casing by both SLS and ULS approaches. A graphical user interface have been developed in order to disseminate the reliability analysis practice among casing designers.

Acknowledgements

References

Abbassian, F. and Parfitt, S. (1995) Collapse and post collapse behaviour of tubulars: a simple approach. In *Proceedings of SPE Production Operations Symposium*.

Adams, A., Warren, A. and Masson, P. (1998) On the development of reliability-based design rules for casing collapse. In *Proceedings of SPE Applied Technology Workshop on Risk Based Design of Well Casing and Tubing*.

American Petroleum Institute – API (1994) API Bull 5C3: Bulletin on formulas and calculations for casing, tubing, drill pipe and line pipe properties. Bulletin.

American Petroleum Institute – API (2008) API TR 5C3: Technical report on equations and calculations for casing, tubing, and line pipe used as casing or tubing; and performance properties tables for casing and tubing. Technical report.

American Petroleum Institute – API (2010) API 5CT: Specification for casing and tubing.

Ang, A. H-S. and Tang, W. H. (2007) *Probability concepts in engineering: emphasis on applications to civil and environmental engineering*, John Wiley & Sons.

Breitung, K. (1984) Asymptotic approximations for multinormal integrals, *J. of Eng. Mech. Div.* **110**, 357–366.

Burres, C. V., Tallin, A. G. and Cernocky, E. P. (1998) Determination of casing and tubing burst and collapse design factors to achieve target levels of risk, including influence of mill source. In *Proceedings of SPE Applied Technology Workshop on Risk Based Design of Well Casing and Tubing*.

Ditlevsen, O. and Madsen, H. O. (2005) *Structural Reliability Methods*, Technical University of Denmark.

Hasofer, A. and Lind, N. (1974) An exact and invariant first order reliability format. *J. Eng. Mech. Div.* **100**, 111–121.

International Organization for Standardization – ISO (2007) ISO TR 10400: Petroleum, petrochemical and natural gas industries - Equations and calculations for the properties of casing, tubing, drill pipe and line pipe used as casing or tubing. Technical report.

Gouveia, L. P. (2014) *Avaliação da Confiabilidade de Tubos de Revestimento de Poços*. Federal University of Alagoas, Master's Thesis. *(in portuguese)*

Klever, F. (2010) Formulas for rupture, necking, and wrinkling of oil country tubular goods under combined loads, *SPE Journal* **15**, 834–855.

Klever, F. and Stewart, G. (1998) Analytical burst strength prediction of OCTG with and without defects. In *Proceedings of SPE Applied Technology Workshop on Risk Based Design of Well Casing and Tubing*.

Klever, F. and Tamano, T. (2006) A new OCTG strength equation for collapse under combined loads. *SPE Drilling and Completion* **21**, 164–179.

Melchers, R. E. (1999) *Structural reliability analysis and prediction*, Civil Engineering Series, John Wiley and Sons.

Nataf, M.A. (1962) Détermination des distributions de probabilités dont les marges sont données. *Comptes Rendus de l'Académie des Sciences* **225**, 42–43. *(in French)*

Scientific Computing and Visualization Laboratory – LCCV (2013) Development of failure criteria for well casing design based on risk analysis. Research Report.

Tamano, T., Mimaki, T. and Yanaimoto, S. (1985) A new empirical formula for collapse resistance of commercial casing. *Nippon Steel Tech. Report* **26**, 19–26.

Design and evaluation smart mandrels based on shape memory polymer

H.Y. Du[1], *L.W. Liu[1], F.L. Chen[1], Y.J. Liu[1], and J.S. Leng[2]

[1]Department of Astronautical Science and Mechanics, Harbin Institute of Technology (HIT),
P.O. Box 301, No. 92 West Dazhi Street, Harbin, China
[2]Centre of Composite Materials and Structures, Science Park of Harbin Institutive of Technology
(HIT), P.O. Box 3011, No. 2 Yikuang Street, Harbin, China

*Corresponding author: liuliwu_006@163.com

Abstract

A kind of smart mandrels (shape memory polymer (SMP) mandrels) have been demonstrated to make up the drawbacks of traditionally mandrels (multi-pieces metal mandrel, water-soluble mandrel and elastomeric rubber mandrel), such as time-consuming and difficulty to remove. In the paper, the styrene-based SMP has been used to fabricate the bottle-shaped and air-duct shaped smart mandrel. Firstly, the glass transition temperature of SMP has been obtained by dynamic mechanical analysis (DMA) test, the peak value of the loss angle has been chosen as the glass transition temperature (63°C). Secondly, the bottle-shaped and air duct-shaped smart mandrels have been manufactured by the process of curing, forming, heating, inflating, cooling and removing. At the same time, the corresponding recovery processes of smart mandrels have been measured to show the good shape recovery ability of SMP mandrels. Finally, the shape recovery process of bottle-shaped mandrel is simulated by finite element method. These results show that SMP can provide some experimental and theory guide for the application field of mandrel fabrication.

Keywords: Shape memory polymer, Smart mandrel, Shape recovery process, Finite element method

Introduction

Shape memory polymer is a kind of new smart materials, which can keep one temporary shape and recover the original shape under some special external stimulus [Leng et al. (2011); Baghani et al. (2012); Liu et al. (2006)], such as temperature [Tobushi et al. (1997); Tan et al. (2013)], electricity current [Liu et al. (2009); Lv et al. (2010)], light [Lendlein et al. (2005)], magnetic filed [Conti et al. (2007)] and solution [Wang et al. (2012); Lv et al. (2008)], and so on. Nowadays, SMPs have attracted a great deal of interest since the development of 1984, particular in the last few years [Leng et al. (2011)]. A typical thermomechanical cycle can be shown in Figure 1 [Lan et al. (2009)]:(1) Fabricating the original shape of SMP sample and heating it above the glass transition temperature; (2) Loading the sample to a temporary shape; (3) Keeping the deformation and cooling the temperature to the room temperature (lower than glass transition temperature); (4) Keeping the temperature and removing the external load; (5) Reheating the sample above the glass transition and recovering the original shape.
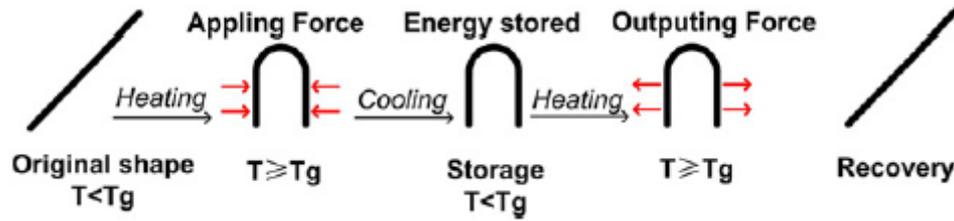
Figure 1. The schematic of a typical thermomechnical cycle of SMP sample
[Lan et al. (2009)]

Compared with traditionally metal materials and shape memory alloy (SMA), SMP possess the unique advantages of low density, high elastic deformation, low energy consumption for shape programming, excellent manufacturability, and low cost [Baghani et al. (2012); Xie (2010); Meng and Li (2013)]. Today, SMP have been synthesized, fabricated, designed, investigated, developed and utilized in a wide range of application [Baghani et al. (2012)]. Such as aerospace [Sokolowski and Tan (2007)], medical research [Lendlein and Langer (2002)], textiles industry [Hu et al. (2012)] and mandrel fabrication technology [Everhart et al. (2005; 2006)]. Cornerstone Research Group (CRG) Company have fabricated some smart mandrels by using shape memory polymer and verified the feasibility of this kind of technology [Everhart et al. (2005; 2006)]. In this paper, the application of SMP for mandrel fabrication in aerospace field is investigated, mainly including bottled-shaped mandrel and air duct-shaped mandrel.

Traditionally, the fabrication methods for complex shape composite parts are mainly multi-piece metal mandrel, water-solution mandrel and inflation elastomeric mandrel [Everhart et al. (2005; 2006), Kim et al. (2010)]. However, there are some inherit disadvantages for every kinds of mandrels. Multi-piece metal mandrel must be disassembled one by one after curing the composite parts and reassembled before next use; water-solution mandrel must collect the waste and dispose it for next use [Everhart et al. (2005; 2006)]; inflation elastomeric mandrel is difficult to provide enough stiffness for uncured composites and the mandrel is easy to damage. On the other hand, SMP mandrel can provide large deformation under high temperature and keep the shape until low temperature, when the temperature is lower than glass transition temperature, the materials can also meet the need of filament winding under small deflection [Everhart et al. (2005; 2006)]. Considering these factors, SMP can be expected to be used as a kind of mandrel for the fabrication of complex shape composite parts in the future.

Based on the above-mentioned state, the mainly content can be organized as follows. In section 2, the polymer materials to fabricate the SMP mandrels are selected and the mechanical properties (such as storage modulus and loss angle) are measured by DMA; In section 3, the SMP bottle-shaped mandrel and air duct-shaped mandrel are fabricated by inflation method; In section 4, the shape recovery ability of SMP mandrels has been demonstrated under some special temperature, then, shape memory process of bottle-shaped SMP mandrel can be simulated by finite element method; Finally, in section 5, the summary and conclusions are given to verify the feasibility of SMP mandrel for the future aerospace applications.

Materials

As shown in our prior paper, there are many kinds of SMP materials for application, mainly including styrene, epoxy, cyanate, and so on [Liu et al. (2014)]. In our work, the material for fabricated SMP mandrels is styrene-based SMP material due to the relatively large deformation ability and shape recovery ability. The general mechanical properties for styrene-based SMP can be measured by DMA device, the three-point testing is selected and the temperature range is designed from 298K(25°C) to 363K(90°C) with the heating and cooling rate 5K/min. The values of loss angle can

be shown in Figure 2. The peak of loss angle (63°C) can be selected as the glass transition temperature in our work.
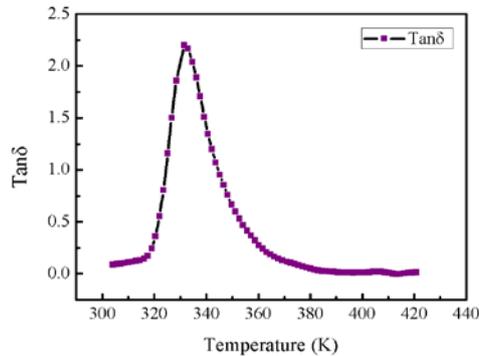
Figure 2. The DMA curve of styrene-based SMP

Fabrication of SMP mandrel

In this section, the bottle-shaped SMP and air duct-shaped mandrel are demonstrated by the fabrication technique of inflation method.

Fabrication of bottle-shaped SMP mandrel

The basic fabrication diagram of bottle-shaped SMP mandrel can be shown in Figure 3, mainly including curing, forming, heating, inflation, cooling and removing steps.
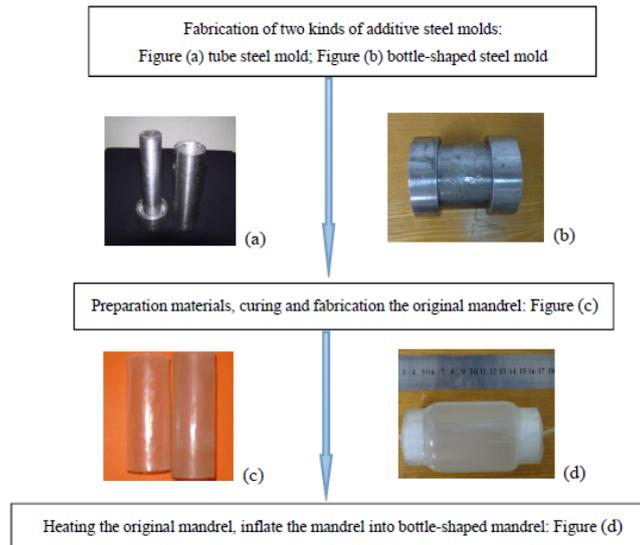
Figure 3. The design diagram of bottle-shaped SMP mandrel

As shown in Figure 3, the two kinds of additive molds for fabricating bottle-shaped mandrel are steel materials with high stiffness and low thermal expansion coefficient to make sure the accuracy of SMP mandrels. The original SMP mandrel is a tube with internal diameter 36mm, outer diameter 40mm and thickness 2mm, the length is 120mm; the final bottle-shaped mandrel with the maximum diameter 50mm at the middle part of the mandrel and 40mm at the two end of mandrel. The shape deformation can be defined by the change in the middle part of mandrel, so the maximum ratio is 25%. In addition, the bottle-shaped mandrel with 50%, 75% and 100% can be fabricated with the same method.

Fabrication of air duct-shaped SMP mandrel

After the bottle-shaped SMP mandrel is fabricated, the air duct-shaped mandrel is also designed and fabricated. The geometry of air duct-shaped mandrel is that a rectangle section at one end with the length 48mm and width 36mm, the other end is circular section with a diameter 55mm, the middle part is gradually transformed from the rectangle section to circular section [Lin et al. (1989)]. The original shape of air duct-shaped mandrel is a tube due to the simple fabrication process, the final air duct shape can be obtained by inflating method, the final shape and original shape of air duct-shaped SMP mandrel can be shown in Figure 4.

Figure 4. The air duct-shaped SMP mandrel with final deformable state and original state

As shown in Figure 4, the original shape of air duct-shaped SMP mandrel is a tube, the final shape is a serpentine shape, the transition range is continuous and smooth with a large curvature. The diameter of tube produces obviously change after carrying out the internal pressure. The left end is a circular section and the right section is rectangle section. The experimental results show that the SMP materials can be easy to duplicate the inner surface shape of outer steel mold with a small pressure; the SMP mandrel owns good shape fixity ability for temporary shape.

Experimental result and discuss

Shape recovery process of bottle-shaped mandrel

As a kind of reusable smart mandrel, SMP mandrel can retain the temporary shape and recover original shape under some external stimulus, the good shape fixity ability can make sure the accuracy of mandrel production and the good shape recovery ability is a key factor for successful extraction from curing composite part. The shape recovery process of bottle-shaped mandrel under 90°C (above transition temperature 63°C) can be shown in Figure 5.
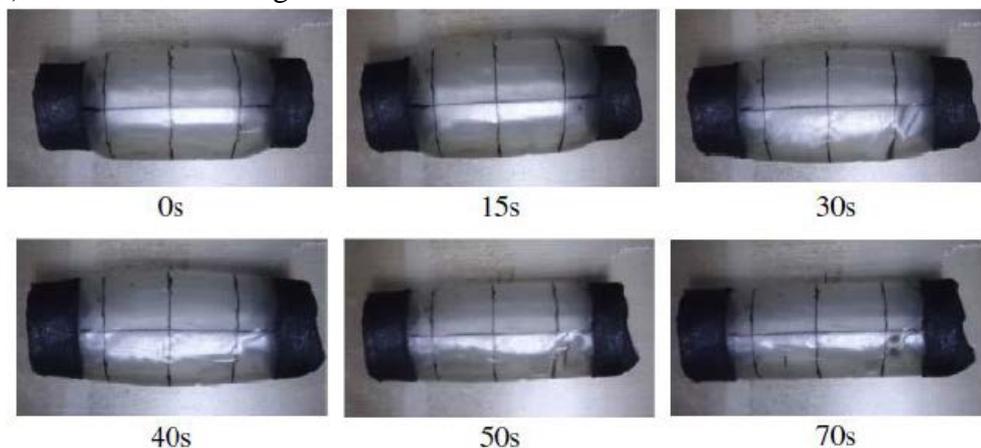
Figure 5. The shape recovery process of bottle-shaped smart mandrel [Zhang et al. (2014)]

As shown in Figure 5, there are some perpendicular lines on the surface of mandrel to locate the deformation. The shape recovery process mainly consists of three steps when placing the mandrel in the oven with temperature 90°C: step 1, the mandrel starts recovery after the mandrel putting into the oven, the rate is relative slow; step 2, The recovery rate markedly increases after 15s and gradually gets the peak; step3, the rate is declined to recover the original shape, when the time is 70s, the bottle-shaped mandrel nearly recovers the original shape. The total recovery rate curve is similar to Gauss function with a single peak. The recovery trend can be effective to alleviate the shock effect when SMP materials are used to fabricate the deployable space structures, such as hinge, solar array, and so on.

As above-mentioned state, the shape fixity ability and shape recovery ability are two very important factors to the application of SMP. In our work, the shape fixity ability and recovery ability are obtained by measuring the maximum diameters change in the deforming process and recovering shape. The maximum diameters change are 25% and 50% in the paper, the definition can be shown as follows:

Shape fixity ratio R_f

$$R_f = \frac{D_L}{D_H} \times 100\% \tag{1}$$

Where D_H, D_L represent the diameter after high temperature deformation and the diameter after cooling to room temperature, respectively.

Shape fixity ratio R_r:

$$R_r = \frac{D_O}{D_R} \times 100\% \tag{2}$$

Where D_R, D_O represent the diameter after high temperature recovery and the diameter before deformation (original shape), respectively.

The shape fixity ratio and shape recovery ratio can be measured and computed as Table 1.

Table 1. Shape fixity ratio and shape recovery ratio of bottle-shaped mandrel

Measuring Location	Mandrel with 25% deformation			Mandrel with 50% deformation		
	Upper section	Middle section	Lower section	Upper section	Middle section	Lower section
D_H	40.00	50.00	40.00	40.00	60.00	40.00
D_L	39.85	49.72	39.81	39.92	59.67	39.66
R_f	99.63%	99.44%	99.53%	99.80%	99.45%	99.15%
D_O	38.49	38.59	38.57	38.90	38.81	38.85
D_R	38.62	38.71	38.87	38.95	38.91	38.97
R_r	99.66%	99.69%	99.22%	99.87%	99.74%	99.69%

Table 1 shows that the shape fixity ratio and shape recovery ratio are more than 99%, which verify the good shape fixity ability and shape recovery ability of SMP mandrel and provide deformation indexes for the application of SMP mandrel in the future.

Shape recovery process of air duct-shaped mandrel

Air duct-shaped mandrel is more complex component-curved shaped structure, in order to better show the good shape recovery ability and reusable ability, the shape recovery experiment is carried out under some special temperature. In our work, the shape recovery temperature is selected as 55°C and the shape recovery process can be shown in Figure 6.
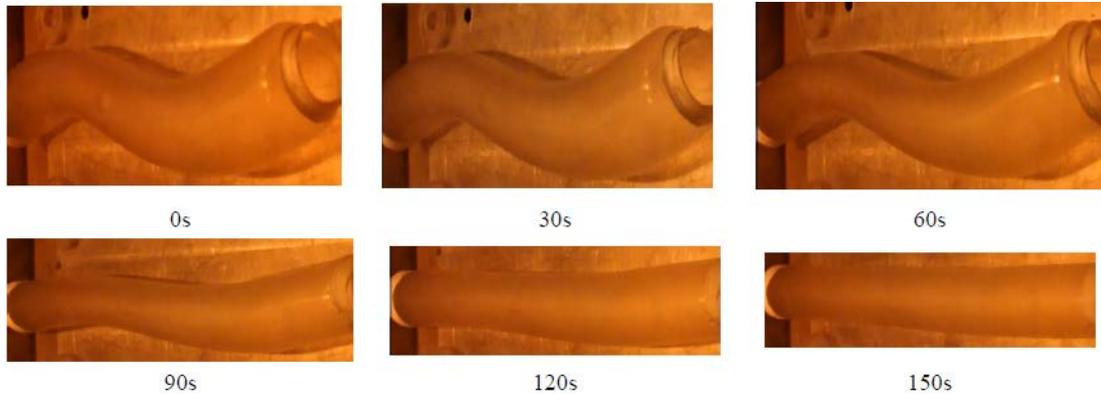
Figure 6. The shaper recovery process of air duct-shaped mandrel

As shown in Figure 6, due to the environment temperature (55°C) is relatively lower than glass transition temperature (63°C), the air duct-shaped mandrel can not obviously change during the first 30s; in the second 30s, the mandrel starts to provide some markedly recovery on the two end. Particular in the circular section, which undergoes the maximum deformation during the deforming process; in the third 30s, the rectangle section gradually transforms to circular section and the curvature of transition zone declines significantly; in the fourth 30s, the curvature zone continuously transform to circular section but the rate is slow; in the fifth 30s, the air duct-shaped mandrel nearly recover the original tube shape with a diameter 40mm. It is noted that the recovery process can take place in the around of transition temperature, when the temperature is less than glass transition temperature, the recovery time is increased obviously.

Shape memory process simulation of bottle-shaped smart mandrel

It is well known that the SMP is in the rigid elastic state when the temperature is low, the deformation is stored and stiffness is commonly enough for filament winding and curing the fiber and matrix to obtain the final composite part. After that, the temperature is reheated above glass transition temperature and the mandrel comes into the soft rubber state with low stiffness, the storage deformation is released and the mandrel recovers the original tube shape to remove the mandrel from composite part. In our work, the shape memory process of bottle-shaped mandrel is simulated to verify the feasibility of SMP mandrels; the finite element software is applied by UMAT function based on three-dimensional viscoelastic model [Zhou et al. (2014)]. The deformation and recovery process of bottle-shaped mandrel can be shown in Figure 7.

Firstly, the bottle-shaped smart mandrel is inflated by internal pressure effect on the high temperature environment, as shown in Figure7(a), then, the mandrel is kept the temporary shape and the temperature is cooled below glass transition temperature, as shown in Figure 7(b); after that, removing the outer force some elastic deformation is springback, as shown in Figure 7(c); Finally, reheating the SMP mandrel above the transition temperature, the mandrel recovers the original tube state, as shown in Figure 7(d). It is noted that the springback effect can be controlled by changing material parameters in the simulation process, In addition, the surface of polymer tube

is smooth and without damage, which can save the manufacturing cycle and reduce the cost significantly.
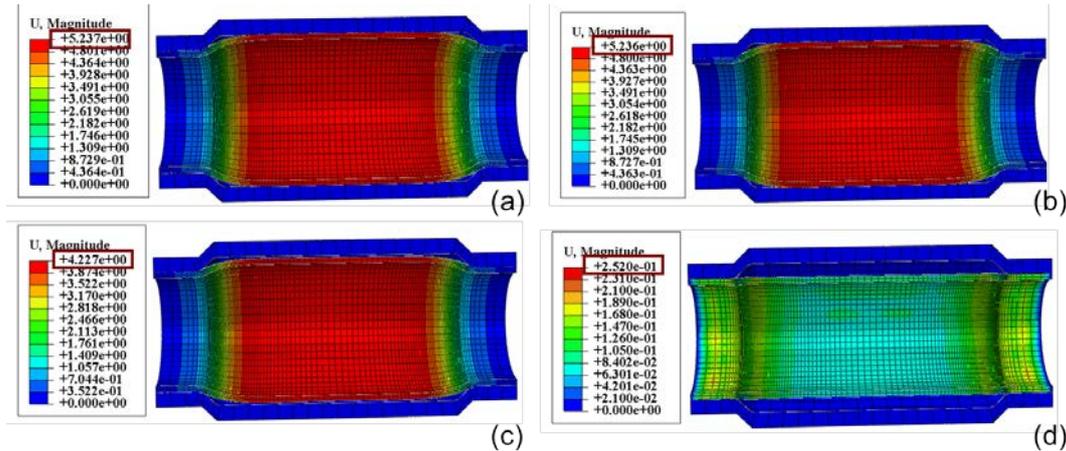
Figure 7. Shape memory process simulation of bottle-shaped mandrel with 25% deformation

Conclusions

The aim of this paper is design and evaluation of bottle-shaped smart mandrel and air duct-shaped smart mandrel based on shape memory polymer. The method is different from the traditional techniques and owns many unique advantages, such as low weight, short production cycle and low cost. The shape recovery testing has been carried out and some results can be shown as follows:

(1) Bottle-shaped SMP mandrel owns good shape fixity ability and shape recovery ability, which is large than 99% in the deformation process and recovery process. The advantage is obviously for SMP mandrel as a kind of mandrel manufacturing techniques for future applications;

(2) Bottle-shaped SMP mandrel and air duct-shaped mandrel can recover the original shape (tube) under some special temperature environment. It is noted that the air duct-shaped mandrel even recovers the original shape under the temperature lower than glass transition temperature.

(3) Bottle-shaped SMP mandrel deformation and recovery process can be simulated by finite element method, the three-dimensional viscoelastic model is applied, the results show the good feasibility of SMP mandrel for the shape memory effect of complex shape composite parts.

References

Baghani, M., Naghdabadi, R., Arghvani, J. and Sohrabpour, S. (2012) A thermodynamically-consistent 3D constitutive model for shape memory polymer. *International Journal od Plasticity*, **35**, 13-30.

Conti, S., Lenz, M. and Rumpf, M. (2007) Modeling and simulation of magnetic-shape-memory-polymer composites. *Journal of the Mechanics and Physics of Solid*, **55**, 1462-1486.

Everhart, M. C. and Stahl, J. (2005) Reusable shape memory polymer mandrels. *Proceedings of SPIE, Smart Structures and Materials: Industrial and Commercial Applications of Smart Structures Technologies*, **5762**, 27-34.

Everhart, M. C., Nickerson, D. M. and Hreha, R. D. (2006) High-temperature reusable shape memory polymer mandrels. *Proceedings of SPIE, Smart Structures and Materials: Industrial and Commercial Applications of Smart Structures Technologies*, **6171**, 61710K

Hu, J. L., Meng, H., Li. G. Q. and Ibekwe, S. I. (2012) A review of stimuli-responsive polymers for smart textile applications. *Smart Materials and Structures*, **21**, 053001.

Kim, G. H., Choi, J. H., and Kweon, J. H. (2010) Manufacture and performance evaluation of the composite hat-stiffened panel. *Composite Structures*, **92**, 2276-2284.

Lan, X., Liu, Y. J., Lv, H. B., Wang, X. H., Leng, J. S. and Du, S. Y. (2009) Fiber reinforced shape-memory polymer composite and its application in a deployable hinge. *Smart Materials and Structures*, **18**, 024002.

Lendlein, A., Jiang, H. Y., Junger, O. and Langer, R. (2005) Light-induced shape memory polymers. *Nature*, **434**, 879-882.

Lendlein, A. and Langer, R. (2002) Biodegradable, elastic shape-memory polymers for potential biomedical applications. *Science*, **296**, 1673-1676.

Leng, J.S., Lan, X., Liu, Y. J. and Du, S. Y. (2011) Shape memory polymers and their composites: stimulus methods and applications. *Progress in Materials Science*, **56**, 1077-1135.

Lin, Q, and Guo, R. H. (1989) Flow characteristics in an s-shaped rectangular-round diffuser at high incidence. Acta Aerodynamica Sinica, **7,** 220-226

Liu, Y. J., Lv, H.B., Lan, X., Leng, J. S. and Du, S. Y. (2009) Review of electro-active shape memory polymer composite. *Composites Science and Technology*, **69**, 2064-2068.

Liu, Y. J., Du, H. Y., Liu, L. W. and Leng, J. S. (2014) Shape memory polymers and their composites in aerospace applications: a review. *Smart Materials and Structures*, **23**, 023001.

Liu, Y. P., Gall, K., Dunn, M. L., Greenberg, A. R. and Diani, J. (2006) Thermomechanics of shape memory polymers: uniaxial experiments and constitutive modeling. *International Journal of Plasticity*, **22**, 279-313.

Lv, H. B., Liu, Y. J., Gou, J., Leng, J. S. and Du, S. Y. (2010) Electrical properties and shape memory behavior of self-assembled carbon nanofiber nanopaper incorporated with shape memory Polymer. *Smart Materials and Structures*, **19**, 075021.

Lv, H. B., Liu, Y. J., Zhang, D. X., Leng, J. S. and Du. S. Y. (2008) Solution-responsive shape-memory polymer driven by forming hydrogen bonding. *Advanced Materials Research*, **47-50**, 258-261.

Meng H, and Li GQ. (2013) A review of stimuli-responsive shape memory polymer composites. *Polymer*, **54**, 2199-2221.

Sokolowski, W. M. and Tan, S. C. (2007) Advanced self-deployable structures for space applications. *Journal of Spacecraft and Rockets*, **44**, 750-754.

Tan, Q., Liu, L. W., Liu, Y. J. and Leng, J. S. (2013) Post buckling analysis of the shape memory polymer composite laminate bonded with film. *Composites Part B: Engineering*, **53**, 218-225.

Tobushi, H., Hashimoto, T., Hayashi, S. and Yamada, E. (1997) Thermomechanical constitutive modeling in shape memory polymer of polyurethane series. *Journal of Intelligent Material System and Structures*, **8**, 711-718.

Wang, C. C., Huang, W. M., Ding, Z., Zhao, Y. and Purnawali, H. (2012) Cooling/Water -responsive shape memory hybrids. *Composites Science and Technology*, **72**, 1178-1182.

Xie T. (2010) Tunable polymer multi-shape memory effect. *Nature*, **464**, 267-270.

Zhang, L., Du, H. Y., Liu, L. W., Liu, Y. J. and Leng JS. (2014) Analysis and design of smart mandrels using shape memory polymers. *Composites Part B: Engineering*, **59**, 230-237.

Zhou. B., Liu. Y.J., and Leng. J.S. (2009) Finite element analysis on thermo-mechanical behavior of styrene-based shape memory polymer polymers. *Acta Polymerica Sinca*, **6,** 525-529.

Influence of Random Nucleation Condition on Transformation Kinetics in Phase Field Simulations

* Takuya Uehara[1]

[1]Department of Mechanical Systems Engineering, Yamagata University, Japan
4-3-16, Jonan, Yonezawa 992-8510, Japan

*Corresponding author: uehara@yz.yamagata-u.ac.jp

Abstract

Influence of nucleation condition in phase field simulation is systematically investigated. Two-dimensional multi-phase-field model for poly-crystalline material was used, and the transformation kinetics was compared with conventional Johnson-Mehl-Avrami-Kolmogorov (JMAK) model. At first, the nucleation was set as the initial condition, and the nuclei arrangement was varied as on regular lattice points or randomly distributed. In such a model, the kinetics did not correspond well to JMAK plot. Time-dependent nucleation was then considered, and it revealed that the kinetic curve agrees well to JMAK plot. Finally, limitation was imposed on the setting of nucleation sites. As a result, it revealed that restriction in the nucleation site interfere the free growth and that the kinetics deviate from the ideal one. It is concluded that proper time-dependent condition with nucleation site set make better correspondence in the transformation kinetics with the JMAK plots.

Keywords: Phase transformation, Phase field model, Microstructure, JMAK model, Nucleation, Uncertainty, Computer simulation

Introduction

Phase field (PF) model has become a useful tool for simulating microstructure formation process of engineering materials, and various complex patterns, such as dendrite, cells, lamella, and poly-crystals, have been regenerated successfully [e.g. Provatas and Elder (2010)]. The model has a basis on the thermodynamics, and free-energy minimization is ensured for growing process of the precipitated phase. Nucleation of the new phase is, however, out of the framework of the PF model, and nuclei are disposed as a computational condition. In solidification process, the melt is usually homogeneous and specific site-dependency in the melt is not found except the wall and surfaces. Randomness is then unavoidably introduced such that the nucleation site is scattered in the model using random numbers, but it brings uncertainty in the obtained results. For re-crystallization during hot-work process, recently, site-dependent nucleation is modeled based on the finite-element analysis using crystal-plastic theorem and sophisticated methods have been in great progress [e.g. Takaki et al.(2009], but the random nucleation model is still a major tool for PF simulations in general. On the contrary, macroscopic kinetics of phase transformation has long been studied. Johnson-Mehl, Avrami and Kolmogorov independently derived kinetic equation, currently known as JMAK equation, in which time evolution of the volume fraction is described using exponential term with empirical parameters. This equation is applied for various processes such as solidification, in-solid phase transformation and re-crystallization processes, and has been built in finite-element codes for engineering use. Phase field simulation exhibits microstructure in detail, while the total or averaged evolution of transformation area is not necessarily accorded with the kinetics [Jou and Lusk (1997), Li et al. (2007), Simmons et al. (2004), Alekseechkin (2011), Uehara (2014)] which is an outstanding problem for bridging the scale. Therefore, in this study, the comparison of the kinetics between PF simulation and JMAK algorithm is systematically demonstrated. A simple two-dimensional model is used to pick out the dominative factors.

Phase Transformation Kinetics

The conventional transformation kinetics known as JMAK form is summarized in this section. Assuming solidification process, solid region is initiated with nucleation, and it grows concentrically at a constant rate G. Then a spherical region of the radius $R = G t$ is formed at time t, and the volume is $V = 4\pi R^3/3$. When the frequency of the nucleation is N per unit time, the number of nuclei generated in the duration between t and $t + d\tau$ is $N d\tau$, and the resulting volume of solid region is $V = 4\pi G^3(t - \tau)^3/3$. Important notice is that every sphere is assumed to continue growing despite they collide to each other in reality. The volume calculated based on this assumption is termed extended volume V_e, which is represented by

$$V_e = \int_0^t \frac{4}{3}\pi NG^3(t-\tau)^3 d\tau = \pi NG^3 t^4 , \qquad (1)$$

where the second equality holds when both G and N are constant. Now, denoting the volume fraction of solid at time t by $x(t)$, the liquid fraction is $1-x(t)$, and the solidification actually occurs in this region. Then the increment of the solid fraction is $dx = (1-x(t)) dx_e$, where x_e is the extended volume fraction. Substituting Eq. (1) into this relation,

$$x = 1 - \exp(-At^4), \quad \text{where} \quad A = \frac{\pi}{3} NG^3 \qquad (2)$$

is obtained. Generally, this formula is expanded as $x = 1 - \exp(-At^n)$, where exponent n is a parameter or named Avrami number, which is one of the fitting parameters. The theoretical value in the two-dimensional model is $n = 3$, which is used as a reference in this paper.

Fundamental Equation for Phase Field Simulation

A poly-crystalline microstructure formation is considered in this study, and multi-phase-field model originally proposed by Steinbach et al. [Steinbach et al. (1996)] is used. Only the fundamental form is described here:

$$\dot{\phi_i} = -\frac{2}{n}\sum_{j=1}^n m_{ij}(f_{ij}\sqrt{\phi_i\phi_j} + \Sigma_k(w_{ik} - w_{jk})\phi_k + \Sigma_k(a_{ik}^2 - a_{jk}^2)\nabla^2\phi_k) , \qquad (3)$$

where ϕ_i is the multi-phase field which is assigned for every grain of precipitated phase for $i = 1$ to N while $i = 0$ is assigned for the original phase, m_{ij}, f_{ij}, w_{ij} and a_{ij} are the parameters and n is the number of existing phase. The values of these parameters, of course, affect the transformation kinetics and should be discussed in detail. However, in this study, to focus on the fundamental characteristics of the pattern formation, constant values for every combination i and j are assumed, and standard values are used. Instead, the total number N of precipitated phase considered is the focused parameter in this study. In the poly-crystalline model, the number of grains N_g is consistent with the number of multi-phase-field variable, i.e. $N_g = N$, to identify each grain as a different phase, but it is time-consuming since the number of combination increases. In this paper, N is taken as 16 even for models with N_g larger than 16, since significant difference was not found in the preliminary calculation.

Model and Conditions

Phase field equation (3) is numerically solved using finite difference method. Two-dimensional square domain is divided by 400x400 lattice points, and periodic boundary condition is imposed on every directions. The whole domain is initially homogeneous original phase. Various nucleation conditions are applied, and three cases reported in this paper are listed in Table 1. Commonly the nucleus position $r_i = (x_i, y_i)$ is selected using random number and the phase field value of ϕ_i is changed to 1, if it is still 0. If the position is already in precipitated phase, the nucleation procedure

is skipped. Firstly, in Case 1, nuclei are set as the initial condition; all nuclei are initially disposed on random position, while regular arrangement is also tested for comparison. Time-dependent nucleation is assumed in Case 2; n_c nuclei are generated every t_c time steps. Finally in Case 3, the nucleation site is limited on specific position; the case for square lattice is demonstrated in this paper. Volume fraction, or actually area fraction, is calculated by $x(t) = N_p / N_T$, where N_p is the number of grids where the phase field value ϕ_0 representing the original phase is 0 and N_T is the total number of grids (N_T=160000). Time-evolution curve of the volume fraction is compared with JMAK plot according to equation (4). The exponent is set as $n = 3$ as the two-dimensional ideal model, and the parameter A is determined so that the time for $x = 0.5$ coincides; i.e. $A = -\ln(1-x)/t^3 = -\ln(0.5)/t_h^3$, where t_h is the time when x reaches 0.5 in the PF simulation.

Table 1. Simulation condition

Case No.	Arrange	Timing
1	Regular / Random	Initial set
2	Random	1 nucleus per 10 steps / 1 nucleus per 2 steps / 2 nuclei every step
3	R2 lattice / R3 lattice	Initial set / 1 nucleus per 10 steps / 1 nucleus per 2 steps

Results and Discussion

Case 1 --- Nucleation as Initial Condition

Results for Case 1 are shown in Figs 1 and 2. Figure 1 shows the phase-field distributions at the early stage which represents the initial arrangement of nuclei, and the resultant polycrystalline structure just before the phase transformation completes in the whole model. The cases for total number of nuclei $n = 16$ and 100 are presented. The color indicates the identifying number of the multi-phase-field variable, where ϕ_0, i.e. the original phase, is shown in blue. As a matter of course, the growing domains collide to each other, and grain boundaries are formed; regularity of the obtained structure is dependent on the initial nuclei arrangement.

Figures 2 (a) and (b) represent variation of the volume fraction for regular and random arrangement, respectively. In addition to $n = 16$ and 100, the cases for $n= 36$ and 64 are also plotted. Since the results for random arrangement are dependent on the random number generated on the computer, two trials for every condition are plotted. Fig. 2 (c) is comparison between the PF result and JMAK model which is fitted according to the above-mentioned procedure, where the average values of two trials for random case are exhibited. Overall, faster growth is observed for regular arrangement, because the distance between the nuclei is uniformly large, and hence free growth duration lasts long. Compared to JMAK plots, PF result

(a) n=16: regular   (b) n=16: random   (c) n=100: random
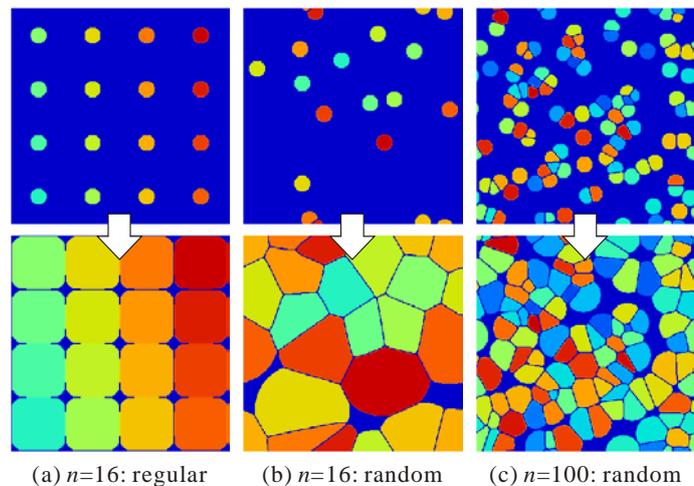
Figure 1. Variation of the phase-field distribution for Case 1.

shows faster growth in the case of regular arrangement for both *n*=16 and 100. For the random arrangement, on the contrary, PF result shows remarkable delay from JMAK plot, and the delay is significant in the latter stage. This tendency is more apparent in Fig. 2(d); all plots are re-drawn against normalized time *t**, in which time is divided by the time when the volume fraction reached 0.5. Note that JMAK plot is identical independent of the fitting parameter *A*.
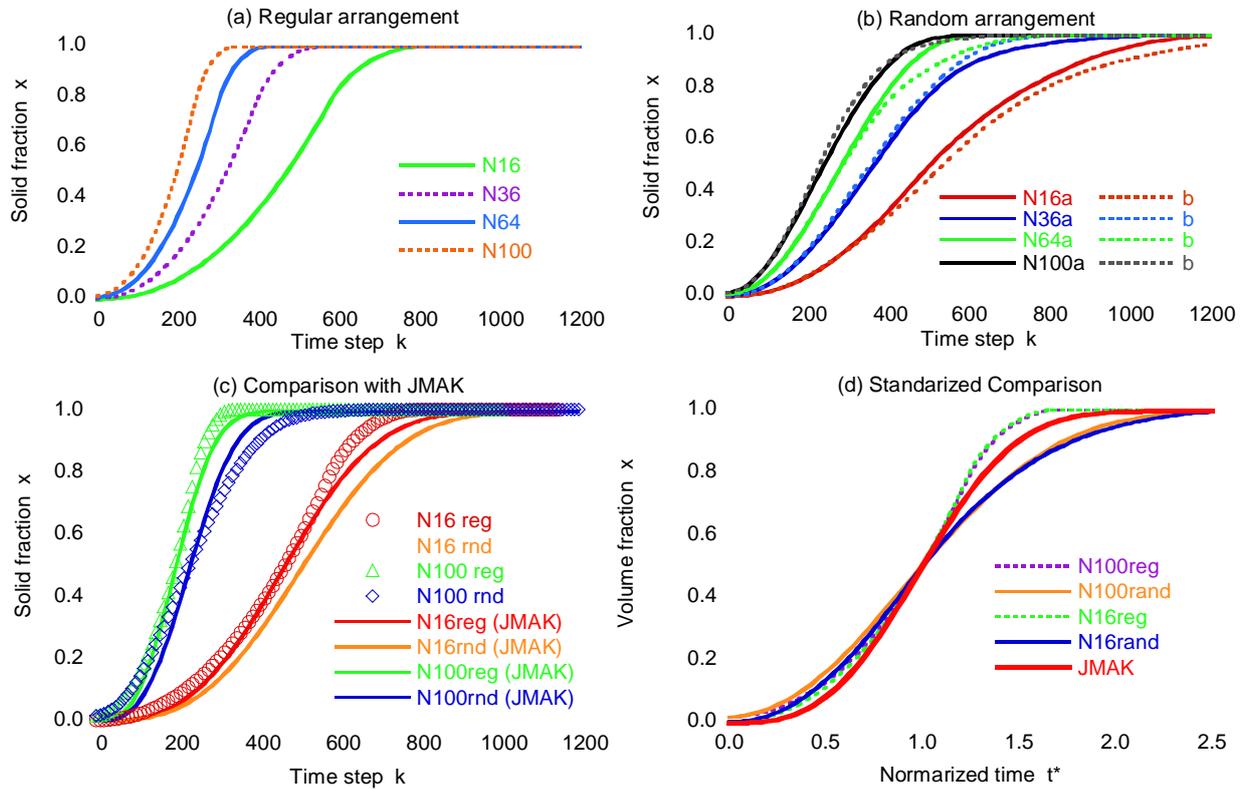
Figure 2. Variation of the volume fraction for Case 1.

Case 2 --- Time-dependent Nucleation

Results for time-dependent nucleation condition in Case 2 are shown in Figs 3 and 4. Here, three conditions are considered: a) one nucleus is generated every 10 steps, (b) one nucleus per 2 steps, and (c) 2 nuclei are generated every time step. Upper figures in Fig. 3 are the phase-field distributions at the 200th time step, and lower figures are those at the time step just before the transformation completes. Naturally, fine grains are formed when nucleation is more frequent, and coarse structure is formed when the nucleation rate is slower. Despite of this obvious difference in the microstructure, the kinetic curves correspond well to each other and also fit the JMAK plot. As shown in the magnified view in Fig. 4, a little discrepancy is found, but they are thought to be in the negligible range. Therefore, it is concluded that the time-dependency of the nucleation plays dominant role for transformation kinetics.

Case 3 --- Restriction in Nucleation Site

Finally, the effect of site-dependent nucleation is investigated. As shown in the result for Case 1, initial nuclei arrangement influences the transformation kinetics. In this section, the time-dependency is also considered. Nucleation site is limited on the square lattice, while the actual
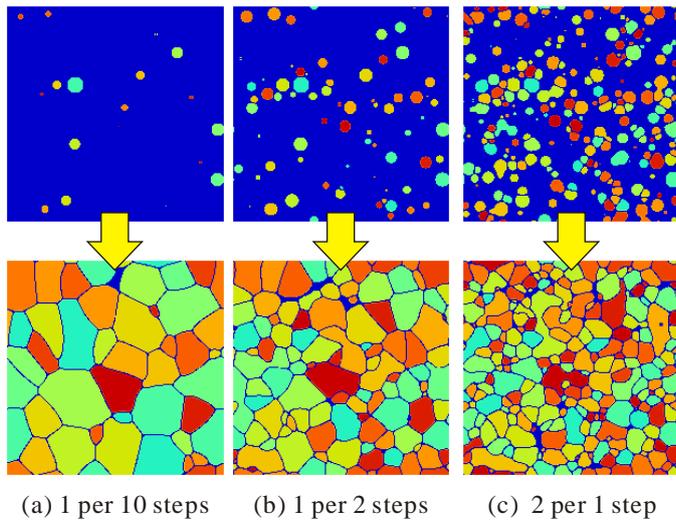
Figure 3. Distribution of phase field for Case2.   **Figure 4. Time evolution for Case 2.**
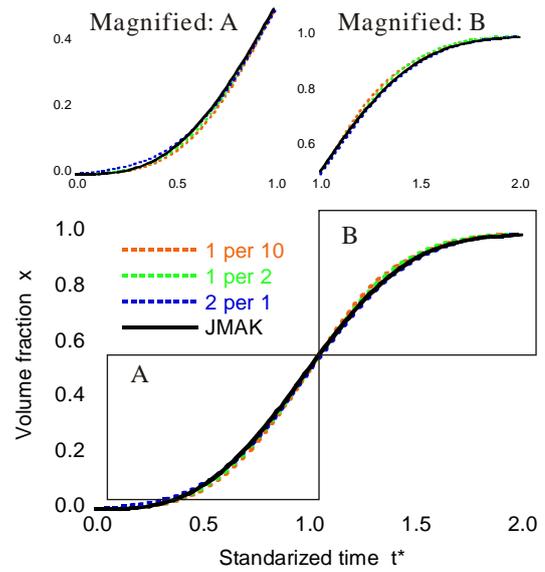
position on the lattice is set at random. Two types of lattice (2x2 (denoted as R2) and 3x3 (R3)) are used. The nucleation rates are assumed as same as those in the previous section, while the case for nucleation at the initial condition is also used for comparison.

Simulation results are shown in Figs 5 and 6. In Fig. 5, variation of the phase field represents the grain growth from the nucleation sites on the R2 or R3 lattice. Before two grains on the lattice collide on the lattice, they grow freely, but after the collision, the growth direction is limited inside the lattice. This causes one-dimensional grain growth, and hence deviation from JMAK plot becomes significant, as shown in the kinetic curve in Fig. 6. Especially when many nuclei are set as the initial condition, (see "R2 init"), there is no duration of free growth, and the resultant kinetics appears to be rather linear. As the nucleation rate becomes slower, the kinetics becomes closer to the ideal curve, which is apparent in Fig. 6(b). Also, the case for R3 lattice shows better correspondence to JMAK. This is because relative duration of free growth to the entire transformation is larger for R3. In other words, as the nucleation sites are strictly limited relative to the whole

(a) R2 lattice: initial set

(b) R2 lattice; 1 nucleus per 2 steps

(c) R2 lattice; 1 nucleus per 10 steps

(d) R3 lattice, 1 nucleus per 2 steps
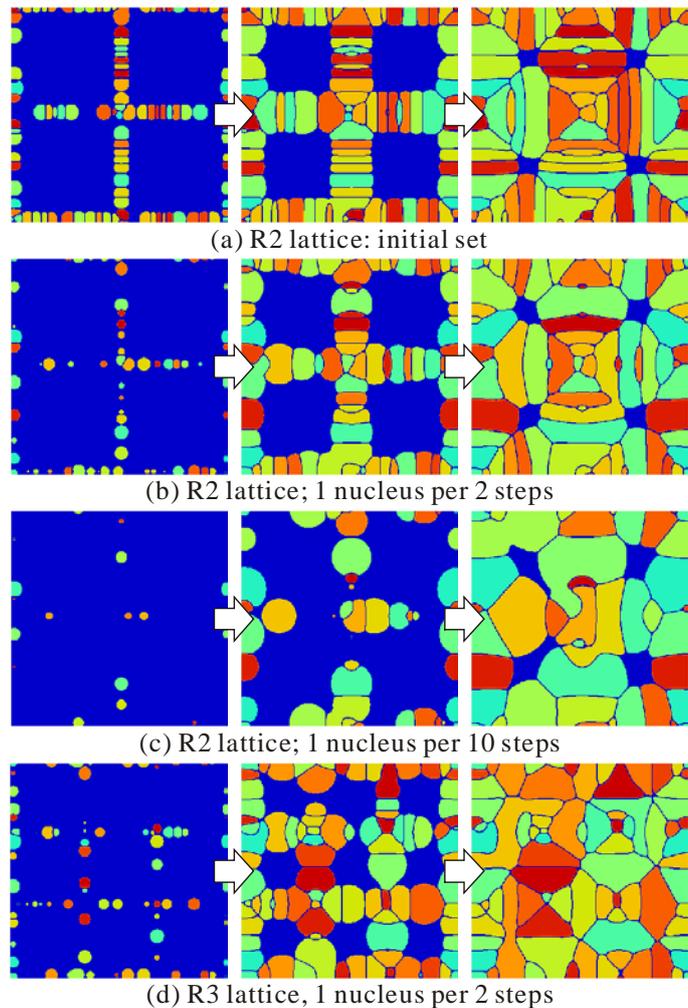
Figure 5. Distribution of phase field for Case 3.

domain, the kinetics deviates from the ideal kinetics.

Conclusions

Nucleation condition is determinative for microstructure formation, and the influence in the phase field simulation was systematically investigated. When the nuclei are set as the initial condition, the kinetics, i.e. the evolution of volume fraction of the transformed region, does not well corresponds to JMAK plot. When time-dependent nucleation condition was introduced, the kinetic curve revealed to agree well to JMAK curve. Restriction on the nucleation site was also determinative for deviation from ideal state. It is then concluded, in this paper, that proper time-dependent nucleation condition makes better correspondence in the transformation kinetics with JMAK theory.

The model considered here is, of course, too simple to be compared with realistic model or experimental results. Three-dimensional model may have additional effects in the kinetics. Physical understanding and its modeling on nucleation

Figure 6. Evolution of volume fraction for Case 3.

process, as well as environmental conditions such as temperature and material composition, are also to be discussed. Nevertheless, fundamental feature of the transformation kinetics presented in this paper is considered to be valuable for further modeling and utilization of the phase field model.

References

Alekseechkin N.V. (2011) Extension of the Kolmogorov-Johnson-Mehl-Avrami theory to growth laws of diffusion type, *Journal of Non-Crystalline Solids* **357**, 3159-3167.

Jou, H.-J. and Lusk, M. T. (1997) Comparison of Johnson-Mehl-Avrami-Kologoromov kinetics with a phase-field model for microstructural evolution driven by substructure energy, *Physical Reviw B* **55**, 8114-8121.

Li, J. J., Wang, J. C., Xu, Q. and Yang, G. G. (2007) Comparison of Johnson-Mehl-Avrami-Kologoromov (JMAK) kinetics with a phase field simulation for polycrystalline solidification, *Acta Materialia* **55**, 825-832.

Provatas, N. and Elder, K. (2010), *Phase-Field Methods in Materials Science and Engineering*, Wiley-VCH, Germany.

Simmons, J. P., Wen, Y. and Wang Y. Z. (2004) Microstructural development involving nucleation and growth phenomena simulated with the phase field method, *Materials Science and Engineering A* **365**, 136-143.

Steinbach, I., Pezzolla, F., Nestler, B., Seeßelberg, M., Prieler, R., Schmitz, G. J., and Rezende, J. L. L. (1996) A phase field concept for multiphase systems, *Physica D* **94**, 135-147.

Takaki T., Hisakuni, Y. Hirouchi, T. Yamanaka, A. and Tomita, Y (2009) Multi-phase-field simulations for dynamic recrystallization, *Computational Materials Science* **45**, 881-888.

Uehara, T. (2014) Influence of randomness in nucleation condition on solidification structure generated by numerical simulation, *Proceedings of the JSCES Conference on Computational Engineering and Science* **19**, to appear.

Reduced-Order Modeling for Transonic Wing Flutter Analysis Including Effects of Control Surface and Nonzero Angle of Attack

T. Kim and K.L. Lai
National University of Singapore, Singapore

Extended Abstract for The 5th International Conference on Computational Methods (ICCM2014)

Introduction

In the past much effort has been made to utilize advanced computational fluid dynamic (CFD) programs for aeroelastic simulations and analyses of military and civil aircraft. Although the use of CFD has become broad for static aerodynamic calculations nowadays, it is limited in the field of unsteady aeroelasticity due to enormous size of computer memory and unreasonably long CPU time associated with the large number of mode shapes in the structural model. While a military airplane model may need 20-50 modes, commercial aircraft models typically require as many as 200 modes to describe the motion of the structure with sufficient accuracy. Thus, both aeroelastic and CFD researchers have explored and developed various ways to reduce the size of the unsteady aerodynamic system and minimize the memory and CPU time. Unfortunately, although these reduced-order models (ROM) retain much of the characteristics of the original full-order models and reproduce the full responses quickly and faithfully, very few of them can be constructed in a time period short enough to justify such an effort, especially when faced with the multiple mode inputs.

In this paper, a new aeroelastic reduced-order modeling based on coupled CFD-CSD aeroelastic responses with a reduced set of state variables is presented. Recently, a novel system identification and model reduction technique, also known as "Aerodynamics is Aeroelasticity Minus Structure" (AAEMS) was developed for linear time-invariant, coupled fluid-structure systems [1]. The method has been successfully applied based on numerical simulations of a scaled Boeing Commercial Aircraft model and AGARD Wing model [1]-[2] (See Figs. 1 and 2), and experimental data obtained of a rigid wing in subsonic wind tunnel [3] (Fig. 3).

The objective of this paper will be to continue this effort and find flutter instabilities of the AGARD Wing in transonic flows including effects of control surface and nonzero angle of attack. Towards this end, a control surface will be attached to the trailing edge of the wing and deflected at an angle. Then, the entire wing plus control surface will be put at a nonzero angle of attack. It is noted that the effects the angle of attack including the deflected control surface on the flutter instability in transonic flow zone have rarely been explored in the literature.
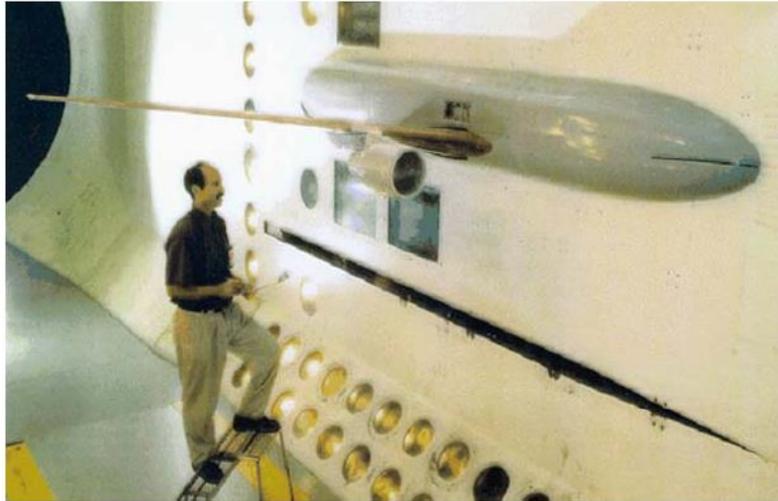
Figure 1. The Twin-Engine Transport Flutter Model (TETFM) in the Transonic Dynamic Tunnel
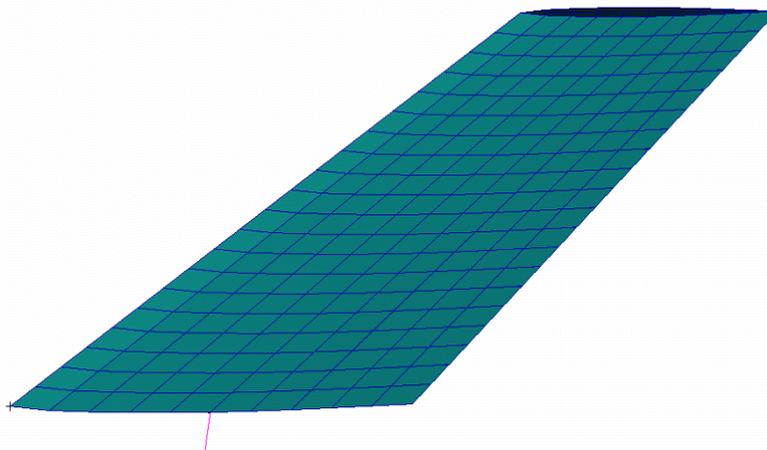
Figure 2. AGARD 445.6 Wing
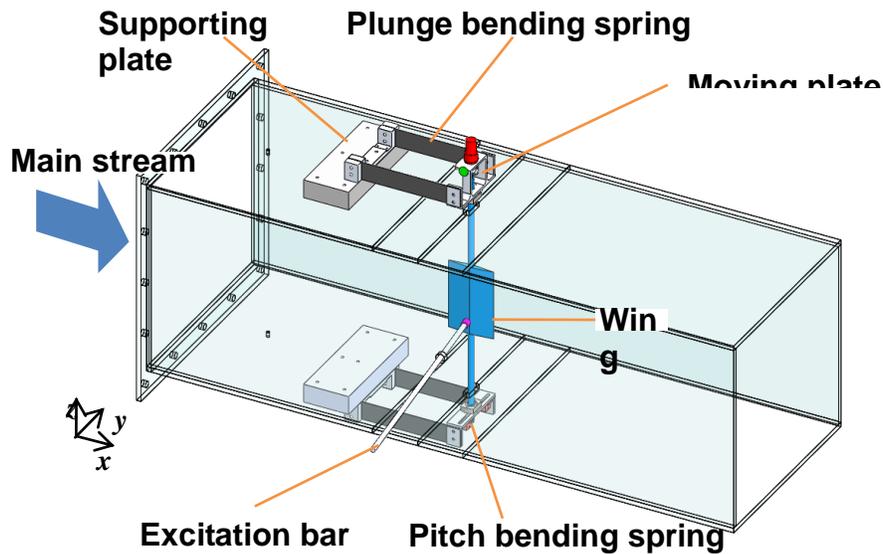
Figure 3. 3D Experimental Wing

AAEMS

Unlike all of the previous ROM methods, the AAEMS works directly on time history data of the coupled aeroelastic system and therefore provides a realistic, easy and efficient tool to construct the aeroelastic ROM. Most importantly, the traditional mode-by-mode excitation of the unsteady aerodynamics is avoided saving a significant amount of model construction time and hence making the method very attractive for the practical applications. Assuming that structural properties are known a priori, and using linear transformations between the structural and aeroelastic states, it extracts and models the underlying unsteady aerodynamic system in discrete-time, state-space format with a finite number of state variables. The displacements and velocities of the structural coordinates are recorded in real time during the numerical simulations. In addition, unsteady pressures are recorded at various points on the lifting wing surface. All the responses are obtained for a fixed Mach, at a low sub-critical dynamic pressure value.

Method of Excitation and Static Condensation

To search for an efficient way to excite the aeroelastic model, various combinations of the structural coordinates as well as control surfaces can be used as potential inputs. It is important to make sure that all the important system modes

are perturbed by the excitation. If control surface inputs are not available, one can use initial conditions instead because a system response due to an initial condition is mathematically equivalent to a response due to an impulse input. Also, to find an optimum number and locations of the aerodynamic samples, different combinations of the pressures at different locations will be explored and the result will be reported in the paper. Considering that the aerodynamic flow will be highly nonlinear in transonic zone, it is critical to apply an input with a very small magnitude to extract only a statically nonlinear but dynamically linear (SNLDL) aerodynamic system. To this end, it is useful to check coherence functions of the various measurements and select only the responses that are linear or sufficiently linear for the system identification. Another issue encountered in the data processing is how to use the time histories that have mixed static/dynamic effects. Normally, unsteady CFD simulation is run after a steady-state equilibrium condition is established first because it might be different than the initial state at t=0. However, running the steady solution separately increases the CPU time significantly and therefore in this paper we will directly advance to the unsteady simulation without the steady calculation, mixing the steady and unsteady parts together. An analytic scheme will be developed such that the resulting time history contains only the effects of the unsteady part and hence is appropriate for the system identification. This so called 'static condensation' is necessary not only for extracting the unsteady solution part but also to avoid 'drifting' mode in the resulting reduced-order model. After the aerodynamic model is identified, an aeroelastic model can be constructed in discrete-time, state-space format by coupling the structural model and the aerodynamic system. The resulting reduced-order model is suitable for constant Mach, varying density (CMVD) analysis including flutter prediction and dynamic response calculation.

CFD Solver and AGARD Wing Model

For application and verification of the method, the aforementioned AGARD Wing model with five structural modes will be studied. An in-house Euler solver developed by Temasek Lab will be used for the CFD part. Since the wing will be at nonzero angle of attack and produce nonzero static response, the static condensation method mentioned above will be critical in eliciting dynamic responses necessary for the system identification. It is also expected that the coupled responses will be highly nonlinear (the structural model is assumed to be linear) due to strong shock waves and numerical artifacts, so the optimum signal processing will be very important in constructing the SNLDL ROM successfully. Aeroelastic results including flutter speeds obtained from the reduced-order model

will be compared to results of the full-order models, and whenever available will be compared with experimental results for various Mach conditions.

It is expected that the proposed new process will generate aerodynamic and aeroelastic ROMs that are useful for the analysis of flutter under the influence of nonzero angle of attack and deflected control surface with minimum amount of effort and time.

References

1. Kim, T., "System Identification for Coupled Fluid–Structures: Aerodynamics is Aeroelasticity Minus Structure" AIAA Journal Vol. 49, No. 3, March 2011.

2. Lai, K.L, and Kim, T., "Reduced-Order Aeroelastic Modeling Using Coupled CFD-CSD Simulations and System Identification Technique," 54th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference, Boston, Massachusetts, April, 2013.

3. Song, J.E., Kim, T., and Song, S.J., "Experimental Determination of Unsteady Aerodynamic Coefficients and Flutter Behavior of a Rigid Wing," Journal of Fluids and Structures, Vol. 29, February 2012, pp. 50-61.

Efficient Prediction of Bending Deformation with Eigenstrain for

Laser Peen Forming

*Yongxiang Hu[1], Zhi Li[2], and Zhenqiang Yao[1]

[1] State Key Laboratory of Mechanical System and Vibration, School of Mechanical Engineering, Shanghai Jiao Tong University, Shanghai, 200240, China
[2] School of Mechanical Engineering, Shanghai Jiao Tong University, Shanghai, 200240, China

*Corresponding author: huyx@sjtu.edu.cn

Abstract

Laser peen forming, is a purely mechanical forming method achieved through the use of laser energy to form large-scale metal plate with small curvatures. The eigenstrain modeling method is used to cut the computational cost to obtain predictions in an efficiently way. The eigenstrain in one representative cell of overlapping laser shocks is obtained by an explicit model to simulate short shock induced plastic deformation. Then, the bending deformation of metal plate is analyzed with an elastic model using the simulated eigenstrain as initial strain. The eigenstrain is incorporated into the elastic model by thermal expansion with a predefined unit temperature field and different anisotropic thermal expansion coefficients. The model can give a consistent prediction for the deformed shape on aluminum alloy 2024-T351.

Keywords: Laser Peen Forming, Eigenstrain, Bending, Aluminum alloy, Finite Element Analysis

Introduction

Laser peen forming (LPF), a derivative of laser shock processing technology, is a locally effective forming process to form complex curvatures without dies. It is now emerging as a viable means for the shaping of metallic components. As a purely mechanical forming method, LPF has advantages of non-contact, tool-free and high efficiency and precision. Its non-thermal process also makes it possible to form without material degradation or even improve them by inducing compressive stress over the target surface, which is desirable because it is important in industry for shaped metal parts to resist cracks from corrosion and fatigue [Ocaña et al., (2007)].

The forming process of LPF has attracted many concerns of researchers. Hackle and Harris demonstrated that it could contour the thick part over its large area and showed that an enhanced convex curvature was achieved [Hackel and Harris, (2002)]. Hu et al investigated the mechanism of laser peen forming that can induce two bending directions under different plate thickness and laser conditions [Hu et al., (2010)]. However, a large number of trial-and-error experiments are typically required on components such as integral wing panel forming before practical applications can be achieved. Beyond direct explicit modeling of material dynamic response, the eigenstrain-based modeling method, which has been intensively investigated in the prediction of welding induced residual stress and distortion, has recently received attention for use in predicting residual stress for LP processes. Korsunsky adopted this method first to predict residual elastic strain with a simplified analytical model, where the eigenstrain distribution is determined by solving an inverse problem with the residual elastic strain using X-ray diffraction [Korsunsky, (2006)]. Achintha & Nowell have just adopted this method to predict laser-peening-induced residual stress to demonstrate feasibility [Achintha and Nowell, (2011)]. Further attention must be paid on the eigenstrain modeling of LPF with overlapping patterns to predict the deformed shape efficiently for a large-scale component.

The aim of this work is to propose a model approach using eigenstrain methodology to predict the deformed shape of a square aluminum plate by LPF. The repeating pattern of plastic strain field is identified and averaged in two axis directions. Then the approximate uniform eigenstrain field is applied to a thermo-elastic model to predict the deformed shape to compare with experiments.

LPF Experiments

The typical application of LPF is carried out under a confined regime configuration. The specimen is undergo a high strain rate deformation and be dynamically yielded due to the rapid laser induced shock pressure. A large number of laser shocks are applied successively to the specimen surface according to the specified path. It will generate incremental deformations in the specimen, which can be accumulated to obtain bending with convex or concave shape depending on process parameters.

Experiments were conducted with a Q-switched Nd:YAG pulsed laser source in the fundamental transverse electro-magnetic mode. The laser was operated at the repetition frequency of 10 Hz and the pulse duration of about 10 ns in FWHM (Full Width at Half Maximum). A wavelength of 532 nm was selected for experiments. The laser output pulse energy measured by a power meter was about 0.93J/pulse. The expanded laser beam passed through a positive long-focus lens to the target surface with the desired beam diameter of 2.0 mm. One kind of black tape, thick enough to maintain its integrity after irradiation of laser pulses, was used as the sacrificial overlay. Water was used as the transparent overlay to confine the generated plasma.

4 mm thick plate sample of aluminum alloy 2024-T351 is prepared with the size of 67 mm ×67 mm. The sample was clamped by the fixture as shown in Fig. 1a. They were manipulated by an industrial robot to move with the predefined scanning path of the fixed laser beam. The overlap of laser spots was set to be 50%, and the shocked-covered region was 40×40 mm in the center area. Fig. 1b shows top view of the sample with scanning laser shocks on the surface. It can be found that the shocked region has some shallow indentations. And the contour surface of sample was quantified by surface profile measurements with a Keyence KS-1100 optical surface profilometer as shown in Fig.1c. It can be found that the pate is deformed with convex curvatures in two directions after square distributed laser shocks.
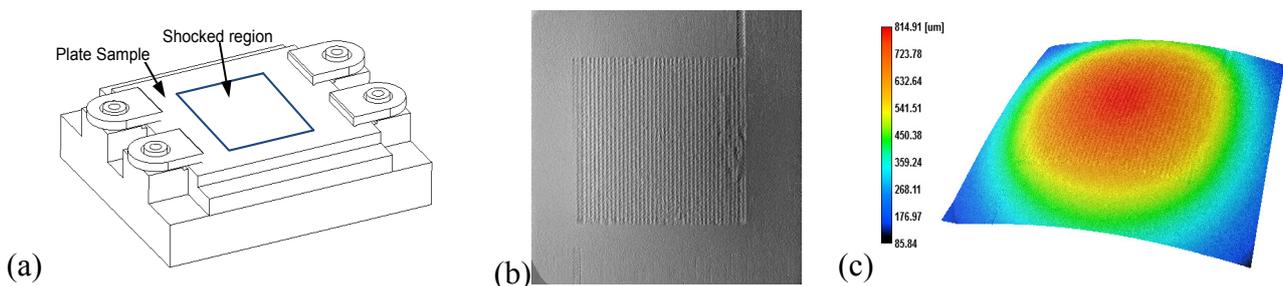
Fig.1 The plate sample formed by laser peen forming with fixture: (a) the device to clamp sample ; (b) top view of the shocked surface; (c) measured surface contour

Eigenstrain-based Modeling Approach

The term eigenstrain, noted by ε*, was introduced by Toshio Mura (1987) to indicate any strain arising in material due to inelastic processes such as plastic deformation, crystallographic transformation, or thermal expansion mismatch between components of an assembly [Mura, (1987)]. Eigenstrain accounts for all permanent strains that arise in material exhibiting inelastic behavior.

The process-induced residual stress and deflections can be predicted through including the eigenstrain as the initial elastic strain distribution [Hu and Grandhi, (2012)].

The generation of deflections due to laser shocks can be regarded as a pure mechanical process. The eigenstrain in laser peen forming is only represented by the plastic strain. Once the representative eigenstrain distribution has been determined under process conditions, the bending deflection and residual stress can be solved with a finite element (FE) model with full scale of component. The corresponding bending moments M_x^* and M_y^* mainly depend on the in-plane eigenstrain components of ε_x^* in the longitude direction and ε_y^* in the transverse direction, respectively [Murakawa et al., (2009)]:

$$M_x^* = E \int \varepsilon_x^* (z - h/2) dx dz \tag{1}$$

$$M_y^* = E \int \varepsilon_y^* (z - h/2) dy dz \tag{2}$$

where h is the plate thickness.

The eigenstrain value is mainly related to the process parameters, material properties, and the specimen thickness in the LPF process, but it is geometry insensitive. Moreover, the eigenstrain generated by each shock is confined to a local region. It allows us to adopt a very small representative cell model to obtain the eigenstrain field, and then apply them to a large-scale model for efficient predictions under the same process conditions and material. For the prediction of deformed shape induced by large-scale array of overlapping shocks, the plastic strains in one repeating pattern of overlapped laser shocks as shown in Fig.2 is very useful to significantly reduce the computation cost. Considering an example of 50% overlap with each laser shock separated with $0.5R$ in Fig.2a, the representative cell size can be selected to be square with the dimension of $2R$ according to the distribution of laser shocks. The dimension with spot size $2R$ is proposed as the simplest repeating pattern for most conditions because it is consistent with the covered area of each shock. The plastic strain field in one representative cell can be reproduced and applied to the practical part model as eigenstrain, one after another, to predict the residual stress and deformation fields based on the characteristic of the repeating pattern. The eigenstrain in one representative cell can be determined efficiently through an explicit infinite-square plate model by simulating multiple sequential laser shocks as shown in Fig.2b. Then the determined eigenstrain field can be applied to the elastic model one by one as shown in Fig.2c for the prediction of deformation fields.
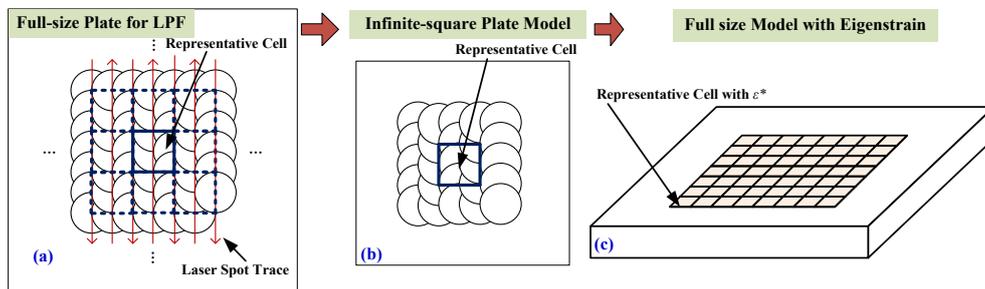
Fig.2 The repeating pattern of a large scale overlapping laser shocks with the percentage overlap of 50%: (a) full-size plate with arrays of overlapping laser shocks; (b) infinite-square pate model to determine the eigenstrain in one representative model; (c) full-size plate model to predict the deformed shape and residual stresses.

Therefore, the eigenstrain-based modeling of LPF process includes two FE models. One is the explicit FE model of a infinite-square plate to determine the eigenstrain in one representative cell. As shown in Fig. 3, this dynamic model is composed of finite part and infinite part. The size of

finite part is determined by the critical distance analysis to be 14R×14R, and shock-covered region is 10R×10R on the top surface.
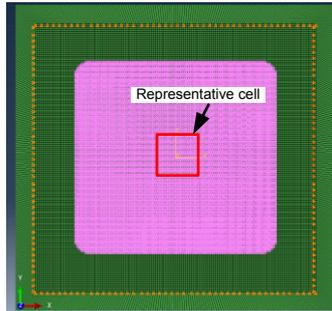
Fig.3 Infinite-square plate model to determine eigenstrain in one representative cell

The time history of shock pressure loading p(t) is calculated by one-dimensional analytical model, proposed by Berthe et al [Berthe et al., (1997)], and the spatial distribution of the pressure considers the difference on the transverse and longitudinal direction with the Eq. 3

$$p(r,t) = p(t)e^{-\frac{(x^2/\beta_1 + y^2/\beta_2)}{R^2}}$$

(3)

Some uncertain parameters existing in the Berthe's model and spatial equation were calibrated first to give a consistent prediction of indentation profiles of single shock. In the infinite-square plate model, the target material is subjected to a shock pressure of few GPa with a short interaction time. The material model for simulation must consider the material behavior dependence of high strain rate to simulate the high-velocity process. The simplified Johnson-Cook model without thermal effect is adopted as the constitutive equation in the explicit model to consider the high-strain-rate effect on the flow behavior of metals:

$$\sigma = \left(A + B\varepsilon^n\right)\left(1 + C\ln\dot{\varepsilon}^*\right)$$

(4)

where $\dot{\varepsilon}^* = \dot{\varepsilon}/\dot{\varepsilon}_0$ is the dimensionless strain rate, and A, B, C and n are considered to be material constants [Johnson, (1985)]. Table 1 provides the material properties of aluminum alloy 2024-T351 required for dynamic simulation.

Table 1 Material properties of aluminum alloy 2024-T351

Properties		2024-T351	Units
Density, ρ		2770	kg m^{-3}
Poisson's ratio, v		0.33	
Elastic modulus, E		73.1	GPa
JC model	A	265	MPa
	B	426	MPa
	n	0.34	
	C	0.015	

After the eigenstrain is determined, the eigenstrain field in the representative cell are averaged in x and y directions, respectively. Then they are imported into the elastic model in the center region with laser shocks for the computation of deformation field. The thermal analysis is an optional method to incorporate the determined eigenstrain field into the model. In the thermal analysis, if the model is defined with the material property of anisotropic thermal expansion coefficients α, the corresponding strain can be calculated as $\varepsilon = \alpha\Delta T$ under the temperature field ΔT. Therefore, the eigenstrain field can be imported by assigning the anisotropic thermal expansion ratios equal to the

value of eigenstrain with six components at each position (Eq.5) and applying a unit temperature variation (Eq.6) to the full model in the Abaqus/implicit module:

$$\alpha(x, y, z) = \varepsilon^*(x, y, z) \qquad (5)$$
$$T(x, y, z) = 1 \qquad (6)$$

Results and discussion

Figure 4 provides the model output of dynamic model and static elastic modes for eigenstrain and deformed shape, respectively. The contours of simulated equivalent plastic strain field on the top surface are shown in Fig. 4a. It can be observed that the plastic strain distribution is periodic with repeating patterns for the model. Due to the 2-axis distribution of shock pressure, the plastic strain in x-direction is a little more than that in y-direction. With the eigenstrain in one representative cell determined by the infinite plate model, the in plane plastic strain ε_x^* and ε_y^* along the depth is averaged within the representative cell. After importing to the full size model of plate sample, the deformed shape as shown in 4b is provided with the predicted contour of deformation with the z-displacement. It can be found that a convex shape can be obtained with square shock in the center area.

(a)

(b)
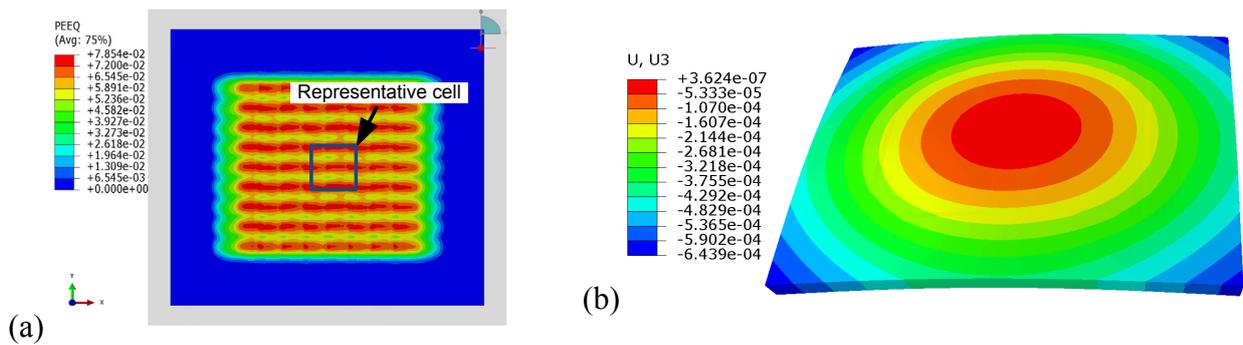
Fig.4 Model output of dynamic model and static elastic modes: (a) eigenstrain in the dynamic model; (b) z-displacement of deformed shaped in the static elastic model

To validate the eigenstrain strain modeling method, the bending profiles obtained by experiments are compared with the predicted values in Fig.5. Fig.5a shows the z-displacement of the centerline along x direction on the top surface for the experiment and simulation. The experimental results shows that the downward bending is about 534 μm, while the model predicted value is about 433μm, producing a consistent agreement but a little underestimated. Fig.5b shows the z-displacement of the centerline along y direction. The experimental results shows that the downward bending is about 388 μm, while the model predicted value is about 380μm, also producing a good agreement with the bending profile.

(a)

(b)
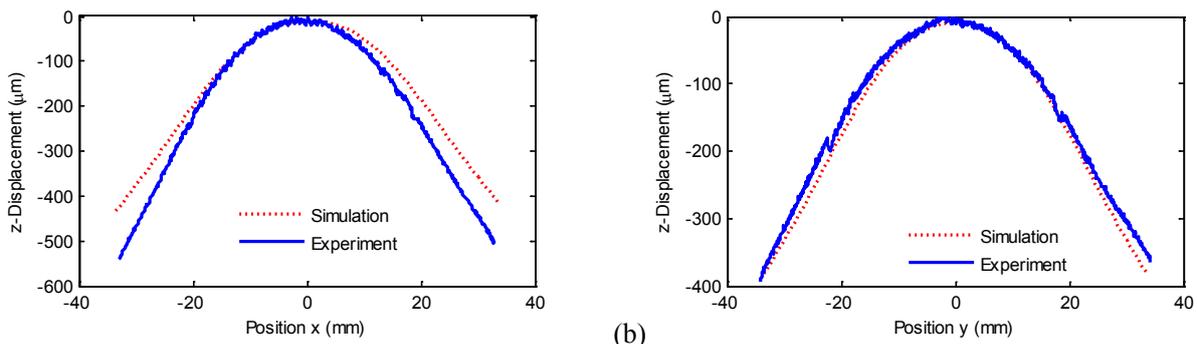
Fig.5 Infinite-square plate model to determine eigenstrain in one representative cell

Conclusions

The eigenstrain modeling method is developed to predict the deformation shape of square plate after laser peen forming. The predictions are validated by the experimental results of overlapping LPF bending of 2024-T351aluminum alloy plate. The eigenstrain modeling method is verified to be an effective approach to simulate the LPF process to predict the bending deformation. The prediction for the large-scale LPF forming process can be completed by identifying the eigenstrain in one representative cell and applying it one after another in the shocked region as an approximation of the actual eigenstrain field. Compared with the direct explicit modeling methods, a shorter computation time is taken for the eigenstrain modeling method to complete the simulation for the LSP application on the large-scale components.

Acknowledgements

References

Achintha, M., Nowell, D. (2011). Eigenstrain modelling of residual stresses generated by laser shock peening. *Journal of Materials Processing Technology* 211, 1091-1101.

Berthe, L., Fabbro, R., Peyre, P., Tollier, L., Bartnicki, E. (1997). Shock waves from a water-confined laser-generated plasma. *Journal of Applied Physics* 82, 2826-2832.

Hackel, L., Harris, F., 2002. Contour Forming of Metals by Laser Peening, U.S. Patent.

Hu, Y., Grandhi, R.V. (2012). Efficient numerical prediction of residual stress and deformation for large-scale laser shock processing using the eigenstrain methodology. *Surface and Coatings Technology* 206, 3374-3385.

Hu, Y.X., Xu, X.X., Yao, Z.Q., Hu, J. (2010). Laser peen forming induced two way bending of thin sheet metals and its mechanisms. *Journal of Applied Physics* 108, 073117.

Johnson, G.R., 1985. Strength and Fracture Characteristics of a Titanium Alloy (.06Al, .04V) Subjected to Various Strains, Strain Rates, Temperatures and Pressures. Naval Surface Weapons Center, Dahlgren, VA .

Korsunsky, A.M. (2006). Residual elastic strain due to laser shock peening: modelling by eigenstrain distribution *Journal of Strain Analysis for Engineering Design* 41, 195-204.

Mura, T., 1987. Micromechanics of defects in solids, 2nd ed. Martinus Nijhoff, Boston.

Murakawa, H., Deng, D., Rashed, S., Sato, S. (2009). 2009-Prediction of Deformation for Large Welded Structures Based on Inherent Strain. *Transactions of JWRI* 38, 63-69.

Ocaña, J.L., Moralesa, M., Molpeceresa, C., Garcíaa, O., Porroa, J.A., J.J. García-Ballesterosa (2007). Short pulse laser microforming of thin metal sheets for MEMS manufacturing. *Appl. Surf. Sci.* 254, 997-1001.

Nonlinear Control of Systems with Non-smooth Nonlinearities

*S. Jiffri[1], J.E. Mottershead[1]

[1] Centre for Engineering Dynamics, School of Engineering, The University of Liverpool,
Brownlow Hill, Liverpool L69 3GH, United Kingdom.

*Corresponding author: s.jiffri@liv.ac.uk

Abstract

Application of feedback control in nonlinear systems is an area of active research. Control algorithms utilising Lyapunov methods, Describing Functions, backstepping etc. are some of the approaches being explored. Feedback Linearisation, which effectively renders the nonlinear system exactly linear through the application of nonlinear feedback, is another approach that has been investigated. Many publications presenting analytical, numerical and also experimental findings have emerged. Much of this work addresses systems with smooth nonlinearities, often described by a polynomial function. The underlying theory of feedback linearisation is well-defined for such systems and is readily available through classical texts and also other publications. For non-smooth systems, however, the applicability of the method is not quite as obvious. The present work aims to demonstrate that at least for some types of non-smooth nonlinearity, the theory of feedback linearisation holds soundly. Successful application of the method in closed-loop control is demonstrated through a numerical example.

Keywords: nonlinear, control, feedback linearisation, non-smooth

Nomenclature and abbreviations

$\omega_\alpha, \omega_\beta, \omega_\xi$ $\quad$ – uncoupled natural frequency in α, β, ξ DOFs

$\zeta_\alpha, \zeta_\beta, \zeta_\xi$ $\quad$ – viscous damping coefficients in α, β, ξ DOFs

a $\quad$ – distance from aerofoil mid-chord to rotational axis, normalised by b

b $\quad$ – aerofoil semi-chord

c $\quad$ – distance from aerofoil mid-chord to aileron hinge line, normalised by b

K_α, K_β, K_ξ $\quad$ – structural stiffness in α, β, ξ DOFs

$\mathbf{D}, \mathbf{E}_1, \mathbf{E}_2, \mathbf{F}$ $\quad$ – matrices relating to the augmented states of the aeroelastic system

$\mathbf{M}_t, \mathbf{C}_t, \mathbf{K}_t$ $\quad$ – overall mass, damping and stiffness matrices of aeroelastic system

r_α, r_β $\quad$ – radius of gyration in α, β normalised by b

U, U^* $\quad$ – air velocity, reduced air velocity ($U^* = U / b\omega_\alpha$)

x_α $\quad$ – COM distance of wing+aileron from rotational axis, normalised by b

x_β $\quad$ – COM distance of aileron from hinge line, normalised by b

COM $\quad$ – centre of mass
DOF(s) $\quad$ – degree(s) of freedom
LFS $\quad$ – linear flutter speed
UoL $\quad$ – University of Liverpool
WTAR $\quad$ – wind tunnel aerofoil rig

1   Introduction

Suppression of vibration is among the major considerations not only in the design and manufacture of new systems, but also in improving existing and well-established ones. A variety of active and passive control methods have been explored. Active control poses the advantage of being able to alter the control inputs based on observed response, thus allowing greater control of the plant. The modelling of nonlinearities in the system being controlled is becoming increasingly important, fuelled by the ever-growing desire to increase effectiveness of existing control methods or develop new ones altogether. In this work, the numerical illustration considered is that of flutter suppression in a 3-DOF pitch-plunge-flap aeroelastic system.

There have been many publications in the literature dealing with the control of systems with smooth nonlinearities, including aeroelastic systems. The application of feedback linearisation on nonlinear aeroelastic systems with smooth structural nonlinearities, mainly of the hardening type, was investigated in (Platanitis and Strganac 2004, Strganac, et al. 2000, Ko, et al. 1999, Jiffri, et al. in press, Jiffri, et al. 2013, Jiffri, et al. 2013, Jiffri, et al. 2013); both theoretical and experimental aspects have been addressed. Papers related to non-smooth systems are also available, albeit in less abundance. A method for adaptive control with feedback linearisation of systems containing a freeplay input was presented in (Recker, et al. 1991), which was extended subsequently to include also a freeplay output (Tao and Kokotovic 1997). The cases of partial feedback linearisation with and without relative degree were addressed subsequently in (Ma and Tao 2000). Other papers related to control of non-smooth nonlinear systems include (Zheng, et al. 2013, Tao, et al. 2013).

The present work applies partial input-output feedback linearisation on a 3-DOF aeroservoelastic numerical model with a piece-wise linear stiffness in the pitch DOF, with the aim of stabilising the linearised response through pole-placement. The model employed is that developed by Edwards et al. (Edwards, et al. 1979), which includes actuator dynamics and approximates unsteady behaviour using two additional augmented aerodynamic states. Other work in which this model has been used include (Conner, et al. 1997, Li, et al. 2010). In the present work, the parameters of the model are tuned to match the dynamics of the wind tunnel aerofoil rig (WTAR) at the University of Liverpool.

This paper commences with a description of the nonlinear aeroelastic system. Equations of motion are given, and are followed by frequency and time-domain simulation results based on the WTAR parameters. Expressions for input-output linearisation of the plunge DOF are derived, including those for the zero-dynamics. Numerical simulation results from the closed-loop system are then presented, demonstrating successful control of the system with a piecewise linear non-smooth nonlinearity both when full knowledge of the nonlinearity is assumed, and when there is uncertainty associated with the nonlinearity.

2   Model description

In this section, a detailed description of the aeroelastic model employed in this work is given. Thereupon, numerical simulation results performed using aeroelastic parameters pertaining to the WTAR at the University of Liverpool will be presented.

2.1 Equations of motion

The aeroelastic model of Edwards et al. (Edwards, et al. 1979) featuring approximation of the unsteady aerodynamic loads through the use of augmented states is employed in the present work. This model consists of a total of 8 states in the first-order state-space representation. Six of these are structural states, namely plunge (normalised with respect to the semi-chord b), pitch, aileron flap (ξ, α, β respectively) and their time-derivatives ($\dot{\xi}, \dot{\alpha}, \dot{\beta}$ respectively). The remaining two are the augmented aerodynamic states mentioned above (x_{a_1}, x_{a_2}). Equations of motion for the model are given as

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}u, \quad \text{where} \quad u = \beta_{com}, \quad \mathbf{x} = \{x_1 \quad x_2 \quad \cdots \quad x_8\}^T = \{\mathbf{q} \quad \mathbf{v} \quad \mathbf{q}_a\}^T,$$

$$\mathbf{q} = \{\xi \quad \alpha \quad \beta\}^T, \quad \mathbf{v} = \dot{\mathbf{q}} = \{\dot{\xi} \quad \dot{\alpha} \quad \dot{\beta}\}^T, \quad \mathbf{q}_a = \{x_{a_1} \quad x_{a_2}\}^T, \quad (1)$$

$$\mathbf{A} = \begin{bmatrix} \mathbf{0}_{(3\times3)} & \mathbf{I}_{(3\times3)} & \mathbf{0}_{(3\times2)} \\ -\mathbf{M}_t^{-1}\mathbf{K}_t & -\mathbf{M}_t^{-1}\mathbf{C}_t & \mathbf{M}_t^{-1}\mathbf{D} \\ \mathbf{E}_1 & \mathbf{E}_2 & \mathbf{F} \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} \mathbf{0}_{(3\times1)} \\ \mathbf{M}_t^{-1}\mathbf{G} \\ \mathbf{0}_{(2\times1)} \end{bmatrix}, \quad \mathbf{G} = \begin{Bmatrix} 0 \\ 0 \\ r_\beta^2 \omega_\beta^2 \end{Bmatrix}$$

and the definition of all quantities appearing within the above equation may be found in (Edwards, et al. 1979, Li, et al. 2010). The input in the above equation is the desired flap angle of the aileron.

This particular model is chosen as it models the dynamics of the actuator, the means through which the input will be applied. As will be seen later, the existence of a non-smooth nonlinearity in the system will necessitate a non-smooth input during closed-loop control. Since such an input cannot be achieved in practice, modelling of the actuator dynamics is necessary to produce a numerical model that is representative of reality. This becomes an even more important issue when implementing feedback linearisation in practice, as using a control law that is based on a model without actuator dynamics will give rise to a discrepancy between the required non-smooth input and the actual smooth input provided by the actuator. It is expected that such a discrepancy will degrade controller performance.

The nonlinear case of the above model may be readily expressed in the affine form

$$\dot{\mathbf{x}} = \underline{\mathbf{f}}(\mathbf{x}) + \underline{\mathbf{g}}(\mathbf{x})u, \quad (2)$$

where, in the present case $\underline{\mathbf{g}}(\mathbf{x}) \equiv \mathbf{B}$ and the use of different symbols is aimed at maintaining conventionally accepted notation in the linear and affine nonlinear cases.

2.2 Aeroelastic Parameters of the WTAR

From previous experiments and related numerical simulations performed on the WTAR at UoL (Papatheou, et al. 24-26 June, 2013), aeroelastic parameters that describe well the aerofoil behaviour were extracted. These are given in Table 1, along

with estimates for the parameters describing flap dynamics (not found during the experiments).

Table 1 – Parameters of the UoL wind tunnel aerofoil rig, used in the present numerical model

Parameter	Value	Parameter	Value	Parameter	Value
ω_α (rad/s)	35.354	c	0.5428	ω_β (rad/s)	100
r_α	0.4	b (m)	0.175	r_β	0.079057
x_α	0.09	μ	69.0	x_β	0.0125
ω_ξ (rad/s)	22.948	ζ_ξ	0.002	ζ_β	0.002
a	-0.33333	ζ_α	0.015		

This format of parameters is widely used in the literature, and is defined in (Theodorsen 1935) in addition to the papers referenced earlier.

2.3   Frequency domain results for the linear system

For the linear case of the aeroelastic system, one may plot the variation of the eigenvalues with respect to reduced air speed. For a speed range of $U^* = 0.1 - 3.0$, the resulting plot is shown in Fig. 1.
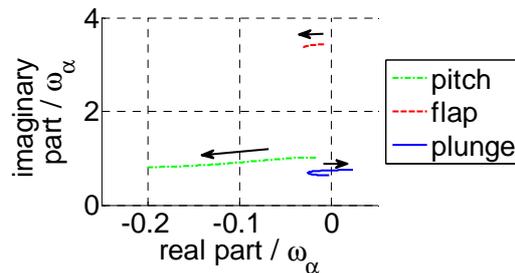
Fig. 1 – Normalised eigenvalues of structural modes varying with airspeed

The linear flutter speed (LFS) is located at the point where the normalised real part of an eigenvalue becomes positive. It is evident from Fig. 1 that this occurs with the plunge mode. The reduced LFS in the present system is found to be $U^* = 2.793$ (this translates to an absolute airspeed of 17.28 m/s).

2.4   Nonlinear time-domain response with piece-wise linear stiffness in pitch

A symmetric piece-wise linear nonlinearity is now introduced into the pitch DOF. The parameters describing the nonlinearity are given in Table 2.

Table 2 – Nonlinearity parameters for piece-wise linear pitch stiffness

Parameter	Description	Value
g_α	initial (lower) stiffness region on either side of $\alpha = 0°$	$1°$
λ	$K_{\alpha_{(|\alpha|\leq g_\alpha)}} = (1-\lambda) K_\alpha$, where $K_{\alpha_{(|\alpha|\leq g_\alpha)}}$ is the initial (lower) stiffness	0.6
K_α	stiffness in the outer regions ($|\alpha| > g_\alpha$), chosen to be equal to linear pitch stiffness K_α	$r_\alpha^2 \omega_\alpha^2$

The resulting pitch moment profile is depicted by the solid line in Fig. 2.
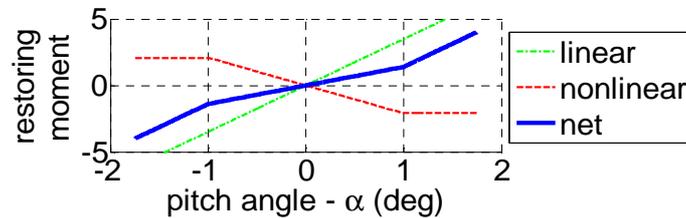
Fig. 2 – Pitch moment profile in the presence of piece-wise linear stiffness

For the purpose of applying feedback linearisation, it will be necessary to define a target linear system, i.e. the desired system once the nonlinearity has been eliminated. This is especially relevant if the feedback linearisation cancels out only the nonlinear terms and not the entire open loop dynamics. Naturally, the target linear system may be chosen as a system whose pitch stiffness is equal to the slope of the outer regions in the nonlinear case. The pitch moment profile in this case is shown by the dash-dot line in Fig. 2. It is now possible to define also the nonlinear moment, i.e. the moment which, when added to the linear moment produces the net, nonlinear moment profile. This non-smooth nonlinear moment profile is shown by the dashed line in Fig. 2.

The nonlinear system is now simulated at a reduced velocity $U^* = 2.0$ with plunge and pitch initial values of $\xi = 0.01$, $\alpha = 3°$ respectively, and with all other states set to zero. The resulting structural responses are shown in Fig. 3.
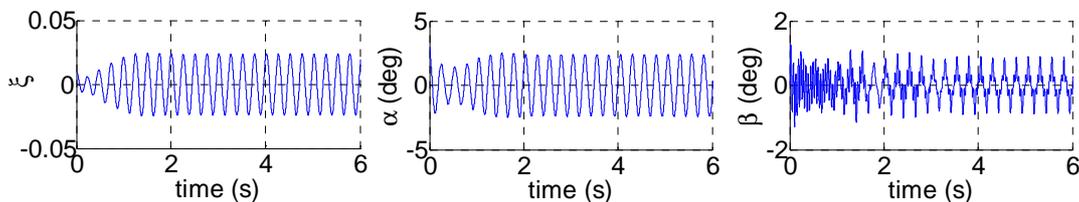
Fig. 3 – Structural states of nonlinear system at U^*=2.0

It is evident that the response settles into an LCO, which occurs at an airspeed which is less than the LFS $U^* = 2.793$, which is expected as the initial stiffness in the nonlinear case is lower than that of the linear system.

3   Feedback linearisation

This section presents the application of feedback linearisation on the nonlinear aeroelastic system described by eq. (2). Feedback linearisation (Isidori 1995, Khalil 2002) is a process whereby a nonlinear system is rendered linear through the application of nonlinear feedback and a co-ordinate transformation. The system in (2) is first expressed as

$$\mathbf{f}(\mathbf{x}) = \left\{ \begin{array}{c} \mathbf{v} \\ \mathbf{\Psi q} + \mathbf{\Phi v} + \mathbf{\Lambda q}_a + \mathbf{\Omega f}_{nl} \\ \mathbf{E}_1 \mathbf{q} + \mathbf{E}_2 \mathbf{v} + \mathbf{F q}_a \end{array} \right\}, \qquad \mathbf{g}(\mathbf{x}) = \begin{bmatrix} \mathbf{0} \\ \mathbf{\Xi} \\ \mathbf{0} \end{bmatrix}, \qquad (3)$$

where

$$\mathbf{\Psi} := -\mathbf{M}_t^{-1}\mathbf{K}_t, \quad \mathbf{\Phi} := -\mathbf{M}_t^{-1}\mathbf{C}_t, \quad \mathbf{\Omega} := -\mathbf{M}_t^{-1},$$

$$\mathbf{\Xi} := \mathbf{M}_t^{-1}\mathbf{G}, \quad \mathbf{\Lambda} := \mathbf{M}_t^{-1}\mathbf{D}, \qquad (4)$$

The nonlinear force vector arising from the piece-wise linearity, as illustrated in Fig. 2 above, is expressed as

$$\mathbf{f}_{nl} = -\lambda \Delta \mathbf{K}_s \mathbf{q} \quad \text{for} \quad |\alpha| \le g_\alpha$$

$$\mathbf{f}_{nl} = -\lambda \Delta \mathbf{K}_s \mathbf{g}_\alpha \quad \text{for} \quad |\alpha| > g_\alpha, \ \alpha > 0 \qquad (5)$$

$$\mathbf{f}_{nl} = \lambda \Delta \mathbf{K}_s \mathbf{g}_\alpha \quad \text{for} \quad |\alpha| > g_\alpha, \ \alpha < 0$$

where

$$\Delta \mathbf{K}_s = K_\alpha \mathbf{e}_2 \mathbf{e}_2^T, \qquad \mathbf{g}_\alpha = \{0 \quad g_\alpha \quad 0\}^T, \qquad K_{\alpha_{(|\alpha| \le g_\alpha)}} = (1 - \lambda) K_\alpha \qquad (6)$$

and where $\mathbf{e}_2$ is the second column of a 3×3 identity matrix. The feedback linearisation method requires that the outputs are continuously differentiable, and therefore smooth. The non-smooth nature of the nonlinearity would result in non-smooth – but continuous – forces/accelerations. However, the resulting changes in the system states (both displacement and velocity) will be smooth, as they are obtained as time-integrals of the accelerations (which are continuous, albeit non-smooth). Thus, all the states of the system are continuously differentiable, satisfying the condition for feedback linearisability.

3.1   Plunge output linearisation

The classical input-output linearisation approach (Isidori 1995, Khalil 2002) is now followed to apply feedback linearisation by controlling the plunge displacement. The co-ordinates of the linear system are obtained as

$$z_1 = y = x_1, \qquad z_2 = \dot{y} = \dot{x}_1 = x_4 \qquad (7)$$

using equation (1). Here, the output y is chosen as the plunge displacement $\xi = x_1$. The partially linearised system may then be obtained as

$$\left\{\begin{array}{c} \dot{z}_1 \\ \dot{z}_2 \end{array}\right\} = \left[\begin{array}{cc} 0 & 1 \\ 0 & 0 \end{array}\right]\left\{\begin{array}{c} z_1 \\ z_2 \end{array}\right\} + \left[\begin{array}{c} 0 \\ 1 \end{array}\right]v, \quad v = \mathbf{f}(\mathbf{x})_{(4)} + \xi_1 u, \quad \Xi = \{\xi_1 \quad \xi_2 \quad \xi_3\}^T, \quad (8)$$

with v being an artificial input associated with the linearised system. Since there remains an un-linearised set of 6 states, it is necessary to examine the zero-dynamics to ensure their stability when designing a controller. Expressions for the remaining linear co-ordinates are first required to complete the transformation. These are chosen as

$$z_3 = x_2, \qquad z_4 = x_3 + \xi_2 x_4 - \xi_1 x_5, \qquad z_5 = \xi_3 x_4 - \xi_1 x_6,$$
$$z_6 = \xi_3 x_5 - \xi_2 x_6, \qquad z_7 = x_7, \qquad z_8 = x_8, \tag{9}$$

completing the 8×8 transformation from nonlinear to linear co-ordinates as

$$\mathbf{z} = \mathbf{T_{zx}x}. \tag{10}$$

The resulting zero-dynamics are found as

$$\dot{\mathbf{z}}_{(3:8)_{ZD}} = \left[\begin{array}{c} \mathbf{P}_1 \\ \mathbf{T_{zx}}_{(4:6,4:6)}[\boldsymbol{\Psi} \quad \boldsymbol{\Phi} \quad \boldsymbol{\Lambda}]\left(\mathbf{T_{zx}}^{-1}\right)_{(1:8,3:8)} + \mathbf{P}_2 \\ \left[\mathbf{E}_1 \quad \mathbf{E}_2 \quad \mathbf{F}_p\right]\left(\mathbf{T_{zx}}^{-1}\right)_{(1:8,3:8)} \end{array}\right]\mathbf{z}_{(3:8)_{ZD}} + \left[\begin{array}{c} \mathbf{0} \\ \mathbf{T_{zx}}_{(4:6,4:6)}\boldsymbol{\Omega}\mathbf{f}_{nl}(z_3) \\ \mathbf{0} \end{array}\right] \quad (11)$$

where

$$\mathbf{P}_1 = \left[\begin{array}{cccccc} 0 & 0 & -\dfrac{\xi_2}{\xi_1\xi_3} & \dfrac{1}{\xi_3} & 0 & 0 \end{array}\right], \quad \mathbf{P}_2 = \left[\begin{array}{ccccccc} 0 & 0 & -1/\xi_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]. \tag{12}$$

The zero-dynamics are checked to verify stability of the internal dynamics of the partially linearised system. A stability investigation of the zero-dynamics yields that there exist 3 equilibrium points – one zero-equilibrium and two non-zero equilibria. The eigenvalues pertaining to the trivial equilibrium point are found to have negative real parts, viz.,

-4.31592 + 98.20627i	-0.84754 + 16.22586i	-37.91684 + 0.00000i
-4.31592 - 98.20627i	-0.84754 - 16.22586i	-3.79316 + 0.00000i

demonstrating stability of this equilibrium point.

3.2   Linearised response with pole-placement implemented

A desired natural frequency ω_{n_ξ} and damping ratio ζ_{n_ξ} may be set for the controlled DOF ξ by choosing the artificial input as

$$v = -\omega_{n_\xi}^2 z_1 - 2\zeta_{n_\xi}\omega_{n_\xi} z_2. \tag{13}$$

For this simulation, target values are chosen as $\omega_{n_\xi} = 1\,\text{Hz}$, $\zeta_{n_\xi} = 0.1$. The resulting closed-loop response, for the same initial conditions as the open-loop case, is shown in Fig. 4.
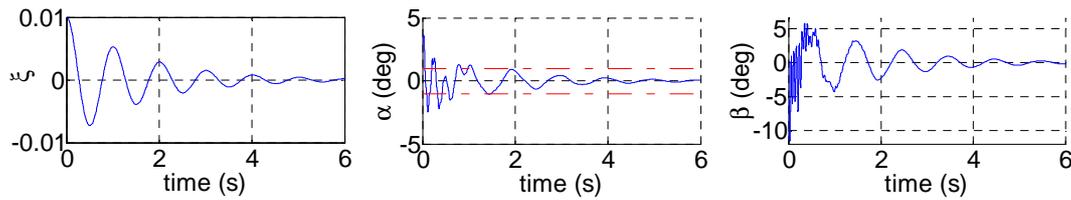
Fig. 4 – Closed-loop response of system at *U**=2.0

It is evident from the first subplot that the target natural frequency of 1 Hz is achieved in the plunge motion, as expected. The pitch motion, confined to the internal dynamics settles down to the stable zero equilibrium, as seen in the middle plot. The flap motion, given by the final subplot, is plotted alongside the commanded input in Fig. 5, where the difference between the two is highlighted.
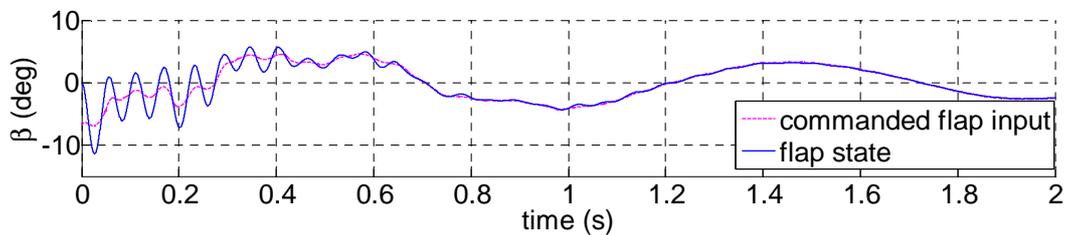
Fig. 5 – Comparing commanded and actual flap angles *U**=2.0

Closer inspection of the input reveals non-smooth changes corresponding to the switching points between the two stiffness regimes in the pitch DOF. This is expected, as the input is designed to cancel the system dynamics which include the non-smooth nonlinear forcing terms. Since the dynamics of the actuator are accounted for in the model and consequently in the computation of the non-smooth input, there will be no degradation on closed-loop response during feedback linearisation, and exact pole placement will be achieved in the absence of nonlinearity parameter errors.

4   Adaptive feedback linearisation

In real situations, complete cancellation of the nonlinearity will not be achievable. This could be due to a variety of reasons such as inaccurate measurement of the nonlinearity, incorrect assumption of the form of the nonlinearity etc. Adaptive Feedback Linearisation is a method that may be used to guarantee asymptotic closed-loop

stability in the presence of a discrepancy between the actual nonlinearity parameters and those assumed in the design of the controller. The assumed nonlinearity parameters are updated at every time step according to an adaptive law, which has the effect of driving the closed-loop controlled responses to zero.

The previous numerical simulation is continued. The inclusion of uncertainty/error in the description of the piece-wise linear stiffness would ideally require a few nonlinearity parameters (to describe the inner and outer stiffness and the range of the inner stiffness), but in this work we assume symmetry, knowledge of the inner stiffness; only the stiffness parameter K_α is considered uncertain. Since the zero-dynamics have an asymptotically stable equilibrium, and the nonlinearity is linearly parameterisable, the conditions for Adaptive Feedback Linearisation are satisfied. A 40% error in K_α is now assumed. Thus, $K'_\alpha = 1.4\,K_\alpha$. Commencing with a scalar quadratic Lyapunov function

$$V = \mathbf{z}^T_{(1,2)}\mathbf{P}\mathbf{z}_{(1,2)} + \tilde{K}^2_\alpha, \qquad \tilde{K}_\alpha = K_\alpha - K'_\alpha, \quad \mathbf{P} > 0, \tag{14}$$

it can be shown that a parameter update law

$$\dot{K}'_\alpha = \mathbf{r}^T \left(\mathbf{\Omega}_{(1,:)}\right)^T \underline{\mathbf{B}}^T_l \mathbf{P}\mathbf{z}_{(1,2)} \tag{15}$$

can be derived, which asymptotically drives the closed-loop controlled response to zero by ensuring that V is a decreasing function. Inclusion of this update law translates to an increase in the dimension of the state vector. In (15),

$$\mathbf{r}(\alpha) = -\mathbf{e}_2\mathbf{e}^T_2\mathbf{q} \quad \text{for} \ |\alpha| \le g_\alpha$$
$$\mathbf{r}(\alpha) = -\mathbf{e}_2\mathbf{e}^T_2\mathbf{g}_\alpha \quad \text{for} \ |\alpha| > g_\alpha, \alpha > 0 \tag{16}$$
$$\mathbf{r}(\alpha) = \mathbf{e}_2\mathbf{e}^T_2\mathbf{g}_\alpha \quad \text{for} \ |\alpha| > g_\alpha, \alpha < 0$$

and $\mathbf{B}_l = \begin{bmatrix} 0 & 1 \end{bmatrix}^T$ is the input matrix of the partially linearised system (equation (8)). The entries in the arbitrary matrix $\mathbf{P}$ are chosen judiciously so as to ensure rapid convergence of K'_α. For the same initial conditions as before and the same pole-placement requirement from the exact linearisation case above, the close-loop responses for the structural DOFs are given in Fig. 6.
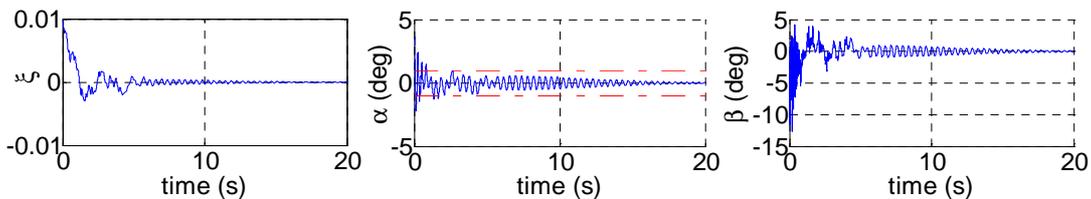
Fig. 6 – Closed-loop response of system with Adaptive Feedback Linearisation at U^*=2.0

It can be seen that the closed-loop response is characterized by higher frequency harmonics as compared with the exact linearisation case. Furthermore, the response

takes longer to settle, although it eventually decays to zero. The pitch response is again driven to the zero equilibrium. A noticeable difference between the controlled response in this case and in the case of exact feedback linearisation is that the pole-placement objective is not achieved here. This is expected, as the adaptive law does not take into account this objective, and merely guarantees the convergence of the response to the origin.

5   Conclusions

This work has presented the application of partial feedback linearisation on a dynamical system having a piece-wise linear structural stiffness nonlinearity. Although the nonlinear forces and the required inputs are non-smooth, the structural states themselves are smooth and continuously differentiable, thereby satisfying the requirements for feedback linearisability. The non-smooth nature of the inputs necessitates modelling of the actuator dynamics, so as to replicate the situation one would encounter in practice, namely that a real actuator is only capable of applying smooth inputs. Numerical simulation results from the 3-degree of freedom aeroelastic model demonstrate successful linearisation of the plunge response, whilst driving the uncontrolled pitch response to zero, as expected from the zero dynamics. The final section presents a simple case of nonlinearity parameter uncertainty and application of the associated adaptive algorithm during feedback linearisation; it is shown from numerical results that the system responses are successfully driven to zero.

Acknowledgements

References

Platanitis, G. and Strganac, T. W. (2004) Control of a nonlinear wing section using leading- and trailing-edge surfaces, *Journal of Guidance Control and Dynamics* **27**(1): 52-58.

Strganac, T., Ko, J. and Thompson, D. (2000) Identification and control of limit cycle oscillations in aeroelastic systems, *Journal of Guidance Control and Dynamics* **23**(6): 1127-1133.

Ko, J., Strganac, T. W. and Kurdila, A. J. (1999) Adaptive Feedback Linearization for the Control of a Typical Wing Section with Structural Nonlinearity, *Nonlinear Dynamics* **18**(3): 289-301.

Jiffri, S., Paoletti, P., Cooper, J. E. and Mottershead, J. E. (in press) A tutorial introduction to feedback linearisation for nonlinear vibration problems, *Shock and Vibration*.

Jiffri, S., Mottershead, J. E. and Cooper, J. E. (2013) Adaptive Feedback Linearisation and Control of a Flexible Aircraft Wing, *Topics in Modal Analysis, Volume 7, Proceedings of the 31st IMAC, A Conference on Structural Dynamics Orange County CA, USA*, Springer.

Jiffri, S., Mottershead, J. E. and Cooper, J. E. (2013) Active Control of a Nonlinear Flexible Aircraft Wing, International Conference on Structural Engineering Dynamics, Sesimbra, Portugal.

Jiffri, S., Mottershead, J. E. and Cooper, J. E. (2013) Nonlinear Control of a Flexible Aeroelastic System, International Forum on Aeroelasticity & Structural Dynamics, Bristol Marriott Royal Hotel, Bristol, UK.

Recker, D. A., Kokotovic, P. V., Rhode, D. and Winkelman, J. (1991) Adaptive Nonlinear Control of Systems containing a Deadzone, Proceedings of the 30th IEEE Conference on Decision and Control, Vols 1-3, New York, IEEE**:** 2111-2115.

Tao, G. and Kokotovic, P. V. (1997) Adaptive control of systems with unknown non-smooth non-linearities, *International Journal of Adaptive Control and Signal Processing* **11**(1): 81-100.

Ma, X. L. and Tao, G. (2000) Adaptive actuator compensation control with feedback linearization, *Ieee Transactions on Automatic Control* **45**(9): 1705-1710.

Zheng, Y. F., Wen, C. Y. and Li, Z. G. (2013) Robust adaptive asymptotic tracking control of uncertain nonlinear systems subject to nonsmooth actuator nonlinearities, *International Journal of Adaptive Control and Signal Processing* **27**(1-2): 108-121.

Tao, G., Burkholder, J. O., Guo, J. X. and Ieee (2013) Adaptive Output Feedback Actuator Nonlinearity Compensation for Multivariable Systems, 2013 American Control Conference, New York, Ieee**:** 3356-3361.

Edwards, J. W., Ashley, H. and Breakwell, J. V. (1979) Unsteady aerodynamic modeling for arbitrary motions, *AIAA Journal* **17**(4): 365-374.

Conner, M. D., Tang, D. M., Dowell, E. H. and Virgin, L. N. (1997) Nonlinear behavior of a typical airfoil section with control surface freeplay: A numerical and experimental study, *Journal of Fluids and Structures* **11**(1): 89-109.

Li, D., Guo, S. and Xiang, J. (2010) Aeroelastic dynamic response and control of an airfoil section with control surface nonlinearities, *Journal of Sound and Vibration* **329**(22): 4756-4771.

Papatheou, E., Tantaroudas, N. D., Da Ronch, A., Badcock, K., Mottershead, J. E. and Cooper, J. E. (24-26 June, 2013) Active Control for Flutter Suppression: An Experimental Investigation, *International Forum on Aeroelasticity and Structural Dynamics (IFASD), Bristol, UK*.

Theodorsen, T. (1935) General theory of aerodynamic instability and the mechanism of flutter, *NACA Report no. 496*: 413-433.

Isidori, A. (1995) Nonlinear Control Systems, Berlin Heidelberg New York, Springer.

Khalil, H. K. (2002) Nonlinear Systems, Prentice Hall.

An error bound for statistically invariant fields in stochastic PDEs

D. A. Paladim [*][1], P. Kerfriden[1], and S. P. A. Bordas[1,2]

[1]Cardiff School of Engineering
[2]Université du Luxembourg

24/04/2014

Abstract

A methodology to compute the error due to homogenization of an heterogeneous stochastic problem is presented. The method is developed for the heat equation, and in order to be applicable, it is required that the conductivity is statistically invariant over the domain. Examples illustrate the behaviour of the bounds.

1   Introduction

In this paper a methodology to compute the error due to homogenization of an heterogeneous stochastic problem is presented. By bounding the error, we mean to bound the difference between the expectation of a quantity dependent of the solution that is defined by the analyst and its estimation from the homogenized problem. The computation of the bound is purely deterministic, however, it involves the solution of another problem, called the dual problem.

All the theory regarding the bound is presented for a Poisson problem (heat equation), though this method can be extended to much wider that fulfill some characteristics listed in this work.

The main features of the bound are its simplicity and low computational cost. On the other hand, under some circumstances, the interval defined by the bound can be very wide. An example is shown and an explanation to this behaviour is given in the Appendix.

The paper is organized as follows. In section 2, the problem is defined and the required notation is introduced. In section 3, the bounds are derived. The following section, presents a couple of simplifications which make the computation of the error bound very efficient. Section 5 presents two test cases to validate the presented bounds.

---

[*]Contact e-mail: daalpa@gmail.com

2 Problem and notation

2.1 Stochastic problem

The problem is defined on a domain $\Omega \times \Theta$, where $\Omega \subseteq \mathbb{R}^2$ is the spatial domain and Θ is the stochastic domain. The boundary of Ω is denoted by Γ, which can be further divided in two subsets Γ_D and Γ_N. Deterministic Dirichlet boundary conditions (temperatures) are prescribed on Γ_D, while deterministic Neumann boundary conditions (fluxes) are prescribed on Γ_N. The conductivity k depends on the spatial coordinate and also on the realization; however, it is statistically invariant,

$$\int_\Theta k = \frac{1}{|\Omega|} \int_\Omega k. \tag{1}$$

With that notation, the partial differential equation reads,

$$
\begin{aligned}
-\nabla \cdot (k\nabla u) &= f(\boldsymbol{x}) & \forall \boldsymbol{x} \in \Omega \times \Theta \\
-k\nabla u \cdot \boldsymbol{n} &= g(\boldsymbol{x}) & \forall \boldsymbol{x} \in \Gamma_N \times \Theta \\
u &= h(\boldsymbol{x}) & \forall \boldsymbol{x} \in \Gamma_D \times \Theta
\end{aligned}
$$

where, f, g and h are known deterministic functions.

Through Green's lemma, the problem can be rewritten in its weak form:

Find $u \in H^1$

$$\int_{\Omega \times \Theta} k\nabla u \cdot \nabla v = \int_{\Omega \times \Theta} fv - \int_{\Gamma_N \times \theta} vg \qquad \forall \boldsymbol{v} \in H_0^1 \tag{2}$$

where

$$H^1 = \left\{ u \in L^2(\Omega \times \Theta) | u = h(\boldsymbol{x}) \text{ on } \Gamma_D \times \Omega \text{ and } \frac{\partial u}{\partial x_i} \in L^2(\Omega \times \Theta) \right\}$$

and

$$H_0^1 = \left\{ u \in L^2(\Omega \times \Theta) | u = 0 \text{ on } \Gamma_D \times \Omega \text{ and } \frac{\partial u}{\partial x_i} \in L^2(\Omega \times \Theta) \right\}$$

We conclude this section introducing more notation. The left hand side of (2) is a bilinear form, that will be denoted by

$$a(u, v) = \int_{\Omega \times \Theta} k\nabla u \cdot \nabla v$$

and its induced norm (since it defines an inner product)

$$\|v\|_a = \sqrt{a(u, u)}$$

; while the right hand side will be denoted by

$$L(v) = \int_{\Omega \times \Theta} fv - \int_{\Gamma_N \times \theta} vg$$

2.2 Quantity of interest

The quantity of interest is the result of applying a certain functional to the solution field. This functional must be expressed as the integral of a linear operator J over a subset of Ω, therefore, the average temperature on a square or the average flux on part of the boundary are examples of quantity of interest. Due to the stochastic nature of the model, this paper is concerned with the expectation of the quantity of interest. The quantity of interest will be denoted by

$$q_\theta(u) = \int_{\Omega'} J(u)$$

and its expectation will be denoted by

$$q(u) = E[q(u)] = \int_{\Theta} \int_{\Omega'} J(u)$$

Now, we proceed to introduce the adjoint problem,

Find $\phi \in H_0^1$

$$\int_{\Omega \times \Theta} k \nabla \phi \cdot \nabla v = q(v) \qquad \forall \boldsymbol{v} \in H_0^1 \qquad (3)$$

Its relevance will be made clear in the following sections.

2.3 Homogenized problems

In this section, we proceed to introduce two more problems. Both problems are deterministic since k is substituted by $\bar{k}$ obtained through homogenization. Since the following problems are easier to solve, their solution is going to be used to compute the quantity of interest and the error due to the substitution will be bounded. The problem reads

Find $\bar{u} \in \bar{H}^1$

$$\int_{\Omega} \bar{k} \nabla u \cdot \nabla v = \int_{\Omega} f v - \int_{\Gamma_N} v g \qquad \forall \boldsymbol{v} \in \bar{H}_0^1 \qquad (4)$$

where

$$\bar{H}^1 = \left\{ u \in L^2(\Omega) | u = h(\boldsymbol{x}) \text{ on } \Gamma_D \text{ and } \frac{\partial u}{\partial x_i} \in L^2(\Omega) \right\}$$

and

$$\bar{H}_0^1 = \left\{ u \in L^2(\Omega) | u = 0 \text{ on } \Gamma_D \text{ and } \frac{\partial u}{\partial x_i} \in L^2(\Omega) \right\}$$

We will denote the approximation of the solution of this problem (obtained by finite elements, for instance) by u^h. And the error field, $u - u^h$ will be denoted by e. The second problem reads:

Find $\bar{\phi} \in \bar{H}_0^1$

$$\int_\Omega \bar{k} \nabla \bar{\phi} \cdot \nabla v = q(v) \qquad \forall \boldsymbol{v} \in \bar{H}_0^1 \qquad (5)$$

The approximation of its solution will be denoted by ϕ^h, while its associated error field, $\phi - \phi^h$, will be denoted by e_ϕ.

3 Derivation of the error bounds

In this section, we proceed to bound the following quantity,

$$q(u) - q(u^h)$$

with an upper and lower bound. The computation of those bounds will not involve the solution of a stochastic problem.

We start by using using equation (3) and the fact that a is linear with respect to its second argument,

$$q(u) - q(u^h) = a(\phi, e)$$

To this expression, we add and subtract $a(\phi^h, e)$ to obtain

$$q(u) - q(u^h) = a(e_\phi, e) + a(\phi^h, e) \qquad (6)$$

Now, we make use of the following result related to the residue,

$$R(v) = L(v) - a(u^h, v) = L(v) + [a(u, v) - a(u, v)] - a(u^h, v) = a(e, v)$$

which allows us to rewrite (6)

$$q(u) - q(u^h) = R(\phi^h) + a(e_\phi, e)$$

It is relevant to emphasize that $R(\phi^h)$ is a deterministic quantity, since

$$R(\phi^h) = \underbrace{L(\phi^h)}_{\text{deterministic}} - \int_\Omega \int_\Theta k \nabla u^h \cdot \nabla \phi^h = L(\phi^h) - \underbrace{E[k] \int_\Omega \nabla u^h \cdot \nabla \phi^h}_{\text{deterministic}}$$

since k is statistically invariant.

Since the bilinear form $a(\cdot, \cdot)$ defines an inner product, we can use the Cauchy-Schwarz inequality to obtain,

$$R(\phi^h) - \|e_\phi\|_a \|e\|_a \le q(u) - q(u^h) \le R(\phi^h) + \|e_\phi\|_a \|e\|_a$$

Now, we proceed to bound the norms of the errors since they are not computable. In order to do that, we introduce an inner product, its induced norm,

$$<u, v>_{k^{-1}} = \int_{\Omega \times \Theta} k^{-1} u v \qquad \|u\|_{k^{-1}} = \sqrt{<u, u>_{k^{-1}}}$$

and a vectorial field $\hat{\boldsymbol{Q}}$ that fulfills,

$$\nabla \cdot \hat{\boldsymbol{Q}} = f(\boldsymbol{x}) \quad \forall \boldsymbol{x} \in \Omega$$
$$\hat{\boldsymbol{Q}} = g(\boldsymbol{x}) \quad \forall \boldsymbol{x} \in \Gamma_N$$

In other words, a flux field that fulfills the prescribed flux boundary conditions. Notice that all the terms in those two equations are deterministic. In [1] several techniques to compute a field with such characteristics are compared.

The aim is to prove that,

$$\|e\|_a \le \|\hat{\boldsymbol{Q}} + k\nabla u^h\|_{k^{-1}} = \eta \tag{7}$$

Firstly, we turn our attention to the following equalities,

$$-\int_{\Omega \times \Theta} \hat{\boldsymbol{Q}} \cdot \nabla v = L(v) \qquad \forall v \in H_0^1$$

$$\int_{\Omega \times \Theta} k\nabla u \cdot \nabla v = L(v) \qquad \forall v \in H_0^1$$

and to their difference,

$$0 = \int_{\Omega \times \Theta} (\hat{\boldsymbol{Q}} + k\nabla u) \cdot \nabla v \qquad \forall v \in H_0^1$$

Setting $v = -u + u^h$, we obtain that

$$< \hat{\boldsymbol{Q}} + k\nabla u, -k\nabla u + k\nabla u^h >_{k^{-1}} = 0$$

in other words, $\hat{\boldsymbol{Q}} + k\nabla u$ is orthogonal to $-k\nabla u + k\nabla u^h$ in the k^{-1}-inner product. This allows us to use the Pythagoras theorem to obtain the desired result

$$\|\hat{\boldsymbol{Q}} + k\nabla u^h\|_{k^{-1}}^2 = \|\hat{\boldsymbol{Q}} + k\nabla u - k\nabla u + k\nabla u^h\|_{k^{-1}}^2 =$$
$$= \|\hat{\boldsymbol{Q}} + k\nabla u\|_{k^{-1}}^2 + \| - k\nabla u + k\nabla u^h\|_{k^{-1}}^2 = \|\hat{\boldsymbol{Q}} + k\nabla u\|_{k^{-1}}^2 + \|e\|_a^2 \ge \|e\|_a^2$$

In a similar manner, the result can be extended to $\|e_\phi\|_{k-1}$. The bounds can be summarized in the following equation

$$\zeta_l = R(\phi^h) - \eta\eta_\phi \le q(u) - q(u^h) \le R(\phi^h) + \eta\eta_\phi = \zeta_u$$

Finally, we show that η is a also deterministic quantity. By expanding its square,

$$\|\hat{\boldsymbol{Q}} + k\nabla u^h\|_{k^{-1}}^2 = \int_{\Omega \times \Theta} k^{-1}\hat{\boldsymbol{Q}} \cdot \hat{\boldsymbol{Q}} + \int_{\Omega \times \Theta} k\nabla u^h \cdot \nabla u^h + 2\int_{\Omega \times \Theta} \hat{\boldsymbol{Q}} \cdot \nabla u^h$$
$$= E[k^{-1}]\int_\Omega \hat{\boldsymbol{Q}} \cdot \hat{\boldsymbol{Q}} + E[k]\int_\Omega \nabla u^h \cdot \nabla u^h + 2\int_\Omega \hat{\boldsymbol{Q}} \cdot \nabla u^h \tag{8}$$

it becomes explicit that it is only a sum of integrals over the spatial domain.

Before concluding this section, it is worth to mention this error bound can be extended to other problems, such as linear elasticity. The results required that the bilinear form to define an inner product, in order to be able to use the Cauchy-Schwarz inequality.

3.1   Deterministic heterogeneous problem

We would like to emphasize that the derivations exposed in the previous could be session could be repeated to compare the solution of an heterogeneous deterministic problem and the solution of it homogenized counterpart. However, in this case, an integral would have to be performed over the heterogeneous domain which may involve the generation of an integration mesh, while for stochastic problem, since the material properties are statistically invariant, the integral is performed over an homogeneous domain, which greatly reduces the computational costs.

4   Simplifications

In this section, we intend to show how under some assumptions, it becomes very efficient to compute the error bound proposed. Our first assumption considers that the homogenized problems were solved through the finite element method with a mesh that is fine enough to assume $\hat{\boldsymbol{Q}} = -\bar{k}\nabla u^h$. We will be also assuming that the Dirichlet boundary conditions are homogeneous. Given that, the finite element method would produce the following two system of equations, for the primal and the dual problem:

$$[\boldsymbol{A}][\boldsymbol{u}] = [\boldsymbol{l}]$$
$$[\boldsymbol{A}][\boldsymbol{\phi}] = [\boldsymbol{q}]$$

Now, the computation of $R(\phi^h)$ reduces to

$$R(\phi^h) = [\boldsymbol{l}]^T[\boldsymbol{\phi^h}] - \frac{E[k]}{\bar{k}}[\boldsymbol{u}]^T[\boldsymbol{A}][\boldsymbol{\phi^h}]$$

and the three integrals that define η^2 are reduced to,

$$E[k^{-1}]\bar{k}[\boldsymbol{u}]^{\boldsymbol{T}}[\boldsymbol{A}][\boldsymbol{u}] + \frac{E[k]}{\bar{k}}[\boldsymbol{u}]^{\boldsymbol{T}}[\boldsymbol{A}][\boldsymbol{u}] - 2[\boldsymbol{u}]^{\boldsymbol{T}}[\boldsymbol{A}][\boldsymbol{u}]$$

5   Numerical results

In this section, two numerical examples are presented to illustrate the behaviour of the bound.

5.1   Low conductivity contrast

The domain of consideration is a L-shape (figure 5.1) made of matrix of conductivity $k_m = 1$ and filled with 75 circular particles of radius 0.05 and conductivity $k_i = 0.6$, resulting in approximate volume fraction of 0.196. The centers of the particles follow an uniform random variable inside, thus, the particles are not allowed to intersect with each other or with the boundaries of the domain. The functions f, g and h are defined as follows,

$$f(\boldsymbol{x}) = 0 \qquad \forall \boldsymbol{x} \in \Omega$$

$$g(\boldsymbol{x}) = \begin{cases} -10(y+1) & \forall \boldsymbol{x} \in \{1\} \times [-1,1] \\ -10(x+1) & \forall \boldsymbol{x} \in [-1,1] \times \{1\} \\ 0 & \forall \boldsymbol{x} \in \{-1\} \times [0,1] \cup [0,1] \times \{-1\} \end{cases}$$

$$h(\boldsymbol{x}) = 0 \qquad \forall \boldsymbol{x} \in [-1,0] \times \{0\} \cup \{0\} \times [-1,0]$$

On the other hand, an approximation of the solution is obtained in a domain with an homogeneous conductivity. The homogeneous conductivity $\bar{k}$ is obtained through the rule of mixture (which in this case, is the same as the $E[k]$) which gives a value of We define our quantity of interest to be the average temperature on $\omega = \{1\} \times [-1,1] \cup [-1,1] \times \{1\}$,

$$q_\theta(u) = \frac{1}{|\omega|}\int_\omega u$$
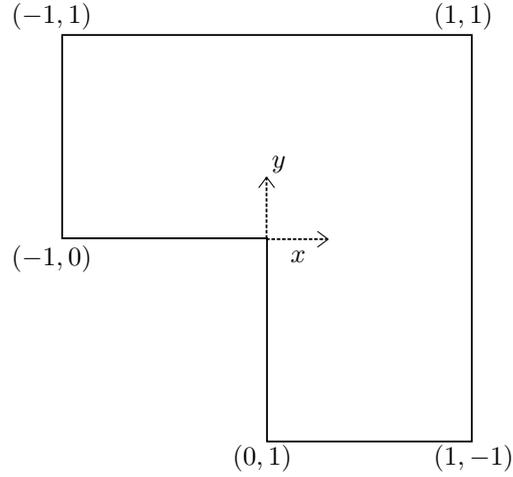
Figure 1: Domain

To validate the bounds a reference quantity of interest will be computed through Monte Carlo using 64 realizations.

Figure 5.1 and 5.1 show the temperature and the gradient field for the homogenized domain and for a random realization.

The following table summarizes the results

$q(u^h)$	$\zeta_l \leq$	$q(u) - q(u^h)$	$\leq \zeta_u$	$\zeta_l + q(u^h) \leq$	$q(u)$	$\leq \zeta_u + q(u^h)$
22.10	-1.20	0.46	1.20	20.90	22.56	23.30

It can be observed that the bounds for $q(u) - q(u^h)$ are symmetric, since ϕ^h belongs to the test space of the u^h.

5.2 High conductivity contrast

In this section, we are going to repeat the previous numerical example with a high contrast between the conductivity of the matrix and the inclusion. The conductivity of the matrix remains the same $k_m = 1$, however, the conductivity of the inclusion is set to $k_i = 0.1$. The following table summarizes the results:

$q(u^h)$	$\zeta_l \leq$	$q(u) - q(u^h)$	$\leq \zeta_u$	$\zeta_l + q(u^h) \leq$	$q(u)$	$\leq \zeta_u + q(u^h)$
25.88	-42.71	5.33	42.71	-16.83	31.22	68.59

The purpose of this numerical example was to show that the interval described by the bounds grows in length very fast when there is a high contrast between the material conductivities, making it unusable. In the appendix an explanation for this behaviour is given.

6 Conclusions

An bound on the error due to homogenization of an heterogeneous stochastic problem was presented in this paper. The bounds can help to asses the validity
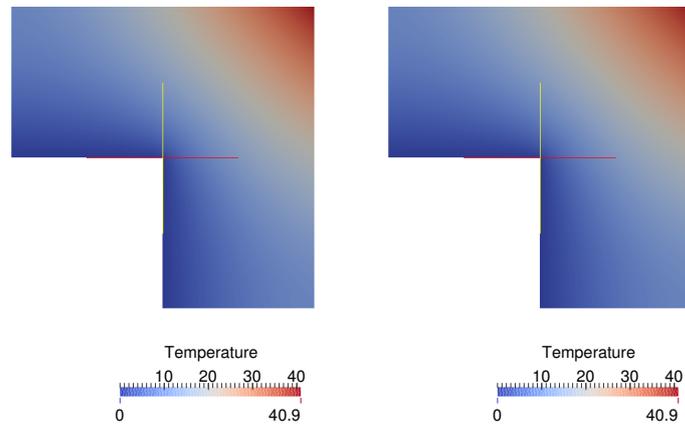
Figure 2: Left: Temperature field for a realization Right: Temperature field for the homogenized field
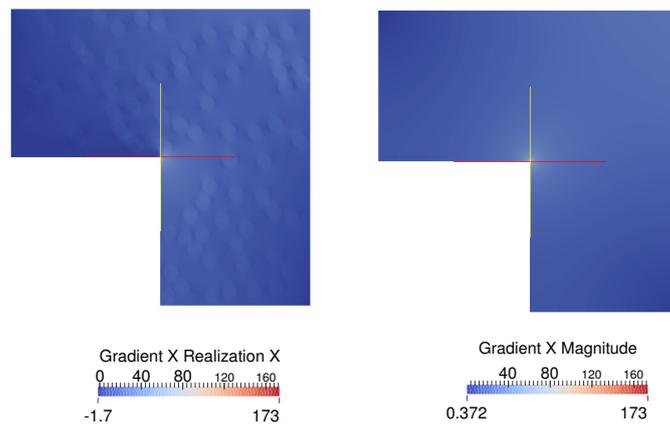
Figure 3: Left: Gradient in the X direction for a realization Right: Gradient in the X direction the homogenized field

of the approach. The computational cost, in some circumstances, is very low, specially, when compared to the deterministic approach, since it only involves an integral over an homogeneous domain.

On the other hand, the numerical examples have highlighted one of the limitations of this work. The error bounds do not behave well, when there exists a huge contrast between the conductivities of the several materials. This will be explored in future works.

That is not the only limitation of this work. The bounds presented are only valid for the expectation. For some analysis, the analyst might also be interested on how the quantities are spread around the expectation. Also, the bound does not take into account the effect of the shape of the particles. The authors have extended the present work to deal with the last two limitations, a bound on the variance, and a bound on the expectation that takes into account the particle shapes. Both ideas will be the object of upcoming publications.

Acknowledgements

References

[1] F Pled, L Chamoin, and P Ladevèze. On the techniques for constructing admissible stress fields in model verification : Performances on engineering examples. (April):409–441, 2011.

APPENDIX

In this section, we intend to explain the behaviour shown in section (5.2). In order to do so, we start by finding the homogenized conductivity that minimizes η and η_ϕ under the following considerations:

- The Dirichlet boundary conditions are homogeneous.

- $\bar{u}$ is the solution of the homogenized problem for a conductivity $\bar{k}$

- $\hat{\boldsymbol{Q}}$ is chosen to be the flux of $\bar{u}$, in short, $\hat{\boldsymbol{Q}} = -\bar{k}\nabla\bar{u}$

Under those considerations, if the conductivity is changed to $\bar{k}'$, the solution field becomes $\frac{\bar{k}}{\bar{k}'}u$. Applying the assumptions to eq. (8) gives

$$\left[\frac{E(k)}{\bar{k}'^2} - \frac{2}{\bar{k}'} + E(k^{-1})\right]\bar{k}^2\int_\Omega \nabla\bar{u}^2$$

an expression that can be minimized with respect of $\bar{k}'$. By taking its derivative with respect to $\bar{k}'$, we find that this expression reaches a minimum for $\bar{k}' = E(k)$, which means that the best bound is obtained, under the stated

assumptions, when rule of mixture is applied. Substituting this value in the previous expression

$$\left[E(k^{-1}) - \frac{1}{E(k)} \right] \bar{k}^2 \int_{\Omega} \nabla \bar{u}^2 \qquad (9)$$

If the expectation of the conductivity has the following form,

$$E(k) = \sum_{i=1}^{N} \alpha_i k_i \qquad \text{where} \qquad \sum_{i=1}^{N} \alpha_i = 1$$

then, the expectation of its reciprocal is

$$E(k^{-1}) = \sum_{i=1}^{N} \frac{\alpha_i}{k_i}$$

Bringing this into the expression inside brackets of equation (9), we obtain

$$E(k^{-1}) - \frac{1}{E(k)} = \sum_{i=1}^{N} \frac{\alpha_i}{k_i} - \frac{1}{\sum_{i=1}^{N} \alpha_i k_i}$$

which is the difference between the reciprocal of the weighted harmonic mean and the reciprocal of the weighted arithmetic mean of k_i ($i = 1...N$). By the generalized mean inequality, it is known that this term will be always equal or greater than 0, having the equality only when $k_1 = k_2 = ... = k_N$. The greater the contrast between the conductivities, the further the expression is from 0. Figure 4 represents the difference between reciprocal of the means for $N = 2$.
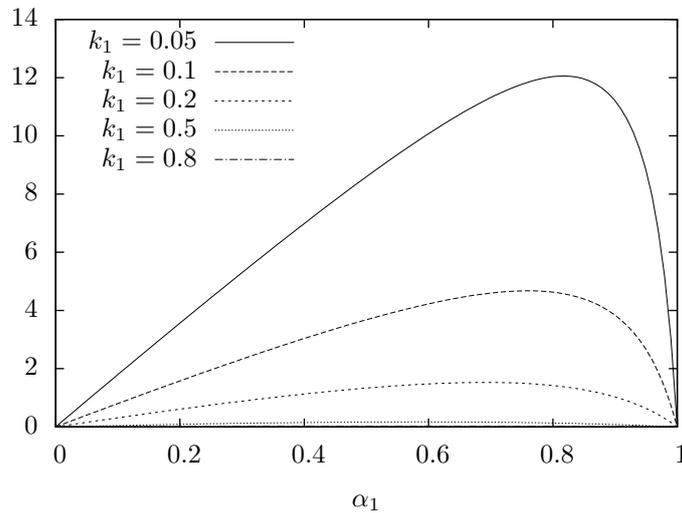
Figure 4: Difference between the reciprocals of the harmonic and the arithmetic mean for $k_2 = 1$ and $N = 2$.

Generalized Integral Transform Solution of Extended Graetz Problems with

Axial Diffusion

N. R. Braga Jr., L. S. de Barros, and L. A. Sphaier*

Laboratory of Theoretical and Applied Mechanics — LMTA/PGMEC,
Department of Mechanical Engineering, Universidade Federal Fluminense,
Rua Passo da Pátria 156, bloco E, sala 216, Niterói, RJ, 24210-240, Brazil

*Corresponding author: lasphaier@id.uff.br

Abstract

A methodology for obtaining fully analytical solutions for the extended Graetz-Brinkman problem including the effects of axial conduction in infinite and semi-infinite domains has been proposed. The infinite domain formulation consists of a preparation (unheated) region (x < 0) followed by a heated one (x > 0) such that back diffusion effects can have an influence on the temperature fields. The solution methodology is based on the Generalized Integral Transform Technique, in which eigenfunction expansions in terms of orthogonal bases are employed. A simple eigenfunction basis in terms of Helmholtz problems are used to maintain the calculation of integral coefficients and the solution of the involved eigenproblems themselves analytical and simple. With the exception of matrix eigenvalues calculation the solution process is mainly analytical.

Keywords: Generalized integral transform, Laminar flow, Forced convection, Duct flow.

Introduction

Graetz problems have been studied for a long time, since early works that include those of Graetz himself [Graetz, 1882]. Initially, simpler versions of the problem, having negligible axial diffusion, simple wall heating conditions (isothermal and isoflux), simple geometry cross-section (either parallel plates or circular channels), and no fluid flow heating effects were investigated, which can be generally denoted the Classical Graetz Problem. This problem is basically of parabolic nature, such that its solutions is given by from a known inlet condition, and extents infinitely in the direction of the flow. Over the course of the decades that followed, a number of extensions to the problem were presented. Since the Graetz problem itself, regardless of the proposed extension, is generally linear, many analytical solutions were proposed, but numerical solution schemes are also found in the literature for the more complicated extensions of the problem.

A common extension to the Graetz is the introduction of axial diffusion into the problem, which complicates matters notably due to the back-diffusion effects that is increased for smaller values of the Péclet number. For this situation, the solution domain must be altered as a preparation unheated (uncooled) region must be considered prior to the heated (or cooled) section. This is usually carried-out by considering an infinite domain that extends from −∞ to ∞ in the flow direction, with a step change in the wall boundary condition at the center of the domain.

Hsu [1968] studied a Graetz problem with axial diffusion in a circular tube, using a semi- infinite domain formulation with a specified inlet condition, while Michelsen and Villadsen [1974] analyzed the effects of axial diffusion in a infinite domain formed by a insulated preparation region followed by an isothermal wall. Both studies used a numerical scheme to complete the solution of

the problem.

Vick et al. [1980] considered a similar Graetz problem with an insulated preparation region followed by a finite isoflux region, and presents an approximate (lowest order analysis) analytical solution in terms of eigenfunctions that arise from the solution of case without axial diffusion. A similar approximate solution procedure was carried out by Bayazitoglu and Ozisik[1980], which solved the axial diffusion Graetz problem with convective boundary conditions at the duct walls and internal energy sources (i.e. flow heating effects). Further on, Vick and Ozisik [1981] and Vick et al. [1983] analyzed similar problems (isoflux walls and convective wall, respectively), but proposed a solution in terms of an alternative eigenfunction basis arising from a problem that does not belong to the traditional Sturm-Liouville class.

Ku and Hatziavramidis [1984] also solved the extend problem subjected to a step-change in wall temperature such that the infinite domain is divided in two portions having an isothermal wall conditions, but with different temperatures. A numerical solution by means of expansions using Chebyshev polynomials was carried out.

Laohakul et al. [1985] and Najjar and Laohakul [1986] considered the infinite domain formulation with two different wall boundary condition arrangements: a heated isothermal wall region surrounded by two cooled isothermal walls, and a heated isoflux wall region surrounded by two insulated regions, respectively. Both studies presented approximate analytical solutions for both large and small Péclet number values in terms using integral transforms with orthogonal eigenfunction expansions.

Ebadian and Zhang [1989, 1990] presented an alternate analytical solution to the extended Graetz problem, writing the temperature field in the infinite domain in terms of a Fourier integral. Although the authors presented a closed-form solution for the temperature field, a Runge-Kutta method was required for performing the required integration to obtain the temperature distribution and Nusselt values.

Johnston [1991] investigated the Graetz problem with axial diffusion in a semi-infinite domain with a prescribed inlet condition, and also used an integral transform solution technique with an orthogonal eigenfunction basis arising from the simpler version of the problem without axial diffusion. However, the author was able to calculate the complete solution of the system (beyond the lower order approximation) in an analytical fashion, by rewriting the transformed ODE system in a first-order differential form.

Min et al. [1997] presented a solution for a Graetz problem with axial diffusion and flow heat- ing effects in a semi-infinte domain with a given inlet condition, for the velocity profile obtained for a Bingham plastic. The solution is expressed in terms of an eigenseries expansion using the same non-Sturm-Liouville basis employed in previous investigations. Olek[1998] also presented a solution for similar extended Graetz problem with a non-Newtonian velocity profile; nevertheless, a solution method similar to that of [Johnston, 1991] was employed.

Çetin et al. [2008] considered a Graetz-type problem for a slip-flow regime with viscous dissipation heating and included the effects of axial diffusion. The authors considered a semi-infinite domain with a prescribed inlet condition, in which a coordinate transformation was employed for arriving at a finite domain. The resulting problem was numerically solved via finite differences. The same authors Çetin et al. [2009] proposed a solution to a similar problem, using a different wall heating condition (step-change in wall temperature), and considering an eigenfunction expansion approach

in terms of a non-Sturm-Liouville eigenfunction basis. Sharma and Chakraborty [2008] also adopted a similar solution technique for tackling an extended Graetz problem with a velocity profile arising from combined pressure and electroosmotic flow driving mechanisms.

In this study, an extended version of a Graetz problem with axial diffusion in an infinite domain is considered. Two heating conditions are analyzed, these being a step change in wall temperature, and a step change in wall heat flux (the lower value corresponding to an insulated condition). The adopted solution strategy is based on orthogonal Sturm-Liouville eigenfunctions expansions of the sought solutions, following the formalism of the nowadays called Generalized Integral Transform Technique [Cotta, 1993]. Closed-form analytical solutions are obtained, and the solution process involves a single numerical methods step, which involves the calculation of numerical matrix eigenvalues and eigenvectors.

Problem Formulation

The studied convective heat transfer problem is an extension of the Graetz problem, including the effects of axial diffusion, for an infinite circular channel with a sudden change in boundary condition at the wall at the origin of the axial coordinate. The problem formulation is given by two sets of equations, valid for different regions. For the upstream region ($\xi \leq 0$) the governing equation and boundary conditions in the transversal direction are given by:

$$u^* \frac{\partial \Phi}{\partial \xi} = \mathrm{Pe}^{-2} + \frac{\partial^2 \Phi}{\partial \xi^2} + \frac{1}{\eta} \frac{\partial}{\partial \eta}\left(\eta \frac{\partial \Phi}{\partial \eta}\right), \quad \text{for } \xi \leq 0, \tag{1a}$$

$$\beta\left(\frac{\partial \Phi}{\partial \eta}\right)_{\eta=1} + (1-\beta)\Phi(\xi,1) = 0, \quad \text{for } \xi \leq 0 \tag{1b}$$

$$\left(\frac{\partial \Phi}{\partial \eta}\right)_{\eta=0} = 0, \quad \text{for } \xi \leq 0, \tag{1c}$$

whereas for the downstream region ($\xi \geq 0$), the equations are given by:

$$u^* \frac{\partial \Theta}{\partial \xi} = \mathrm{Pe}^{-2} + \frac{\partial^2 \Theta}{\partial \xi^2} + \frac{1}{\eta} \frac{\partial}{\partial \eta}\left(\eta \frac{\partial \Theta}{\partial \eta}\right), \quad \text{for } \xi \geq 0, \tag{2a}$$

$$\beta\left(\frac{\partial \Theta}{\partial \eta}\right)_{\eta=1} + (1-\beta)\Theta(\xi,1) = 0, \quad \text{for } \xi \geq 0 \tag{2b}$$

$$\left(\frac{\partial \Theta}{\partial \eta}\right)_{\eta=0} = 0, \quad \text{for } \xi \geq 0, \tag{2c}$$

where Θ and Φ represent the dimensionless temperatures in the two different regions. The β coefficient will lead to different wall heating conditions: $\beta = 0$ for isothermal wall and $\beta = 1$ for constant heat flux.

Part of the required boundary conditions in the axial direction are the inlet and outlet conditions at infinity, defined as:

$$\left|\frac{\partial\Phi}{\partial\xi}\right|_{\xi\to-\infty} < \infty, \qquad \left|\frac{\partial\Theta}{\partial\xi}\right|_{\xi\to\infty} < \infty, \qquad (3a)$$

which reflect the fact that the temperature gradient cannot increase indefinitely for a large value of $|\xi|$. This form for the boundary is chosen since both prescribed wall temperature and wall flux conditions are used. If only a prescribed wall temperature was used, the actual temperature at $|\xi|\to\infty$ become bounded. Nonetheless, for the employed solution strategy the currently adopted form suffices. The remaining boundary conditions in this direction are coupling conditions between the two regions, at the axial coordinate origin ($\xi = 0$):

$$\lim_{\xi\to0^-}\Phi(\xi,\eta) = \lim_{\xi\to0^+}\Theta(\xi,\eta) \qquad (3b)$$

$$\lim_{\xi\to0^-}\frac{\partial\Phi}{\partial\xi} = \lim_{\xi\to0^+}\frac{\partial\Theta}{\partial\xi} \qquad (3b)$$

The employed dimensionless variables are defined by the following equations:

$$\Phi(\xi,\eta) = \frac{T(x,y)-T_0}{\Delta T} \quad \text{for} \ x \le 0, \qquad \Theta(\xi,\eta) = \frac{T(x,y)-T_0}{\Delta T} \quad \text{for} \ x \ge 0 \qquad (4a)$$

$$\eta = \frac{y}{D/2}, \qquad \xi = \frac{x}{D/2\,\text{Pe}} \qquad \text{with} \quad \text{Pe} = \frac{\bar{u}\,D/2}{\alpha}, \qquad (4b)$$

where D is the channel diameter and T_0 is a reference temperature, and the temperature difference ΔT can have two different definitions according to the wall heating condition in the downstream region:

$$\Delta T = T_w - T_0 \quad \text{for} \ \beta = 0, \qquad \Delta T = \frac{D}{2}\frac{\dot{q}_w''}{k} \quad \text{for} \ \beta = 1 \qquad (5)$$

where T_w and $\dot{q}_w''$ represent the wall temperature and wall heat flux in the downstream region.

The presented governing equations allow for a variety of dynamically-developed velocity profiles. However, for illustration purposes, a Hagen-Poiseuille profile is used:

$$u^* = 2\left(1-\eta^2\right) \qquad (6)$$

In spite of this choice, as it will be seen, the solution methodology is practically independent of the choice of u^*, such that any other expressions for u^* as a function of η can be used.

Finally, the Nusselt number in the downstream region can be calculated from In the upstream region, special cases exist:

$$\text{Nu} = \frac{h\,D}{k} = \frac{2}{\Theta_w - \Theta_m}, \quad \text{for} \quad \beta = 1, \qquad (7a)$$

$$\text{Nu} = \frac{h\,D}{k} = \frac{2(\partial\Theta/\partial\eta)_{\eta=1}}{1-\Theta_m}, \quad \text{for} \quad \beta = 0, \qquad (7b)$$

where Θ_w is the dimensionless wall temperature, and the dimensionless bulk temperature Θ_m is calculated form:

$$\Theta_m = 2\int_0^1 u^* \Theta \, \eta \, \mathrm{d}\eta, \tag{7c}$$

Proposed Solution Scheme

The adopted solution methodology is based on the Generalized Integral Transform Technique (GITT). As usual among this type of methodology, filter problems are proposed for removing non-homogenities from the original system. Since the upstream portion of the problem is homogeneous, only the downstream formulation needs to be filtered, which is carried out based on the following solution separation:

$$\Theta(\xi,\eta) = \theta(\xi,\eta) + F(\xi,\eta) \tag{8}$$

in which θ represents the filtered variables, whereas F represent the filter function, which is commonly obtained from simpler versions of the original problem.

For the isothermal walls condition ($\beta = 0$), the selected downstream filter function is simply a constant:

$$F(\xi,\eta) = 1, \tag{9a}$$

whereas for the isoflux wall condition ($\beta = 1$) a polynomial filter is used:

$$F(\xi,\eta) = 2\xi - \frac{7}{24} + \eta^2 - \frac{\eta^4}{4} \tag{9b}$$

whereas, both filters correspond to the solution of the problem as $\xi \to \infty$.

If the previously presented filters are employed, the resulting filtered problem for the downstream region is given by:

$$u^* \frac{\partial \theta}{\partial \xi} = \mathrm{Pe}^{-2} + \frac{\partial^2 \theta}{\partial \xi^2} + \frac{1}{\eta} \frac{\partial}{\partial \eta}\left(\eta \frac{\partial \theta}{\partial \eta}\right), \quad \text{for } \xi \geq 0, \tag{10a}$$

$$\beta\left(\frac{\partial \theta}{\partial \eta}\right)_{\eta=1} + (1-\beta)\theta(\xi,1) = 0, \quad \text{for } \xi \geq 0, \tag{10b}$$

$$\left(\frac{\partial \theta}{\partial \eta}\right)_{\eta=0} = 0, \quad \text{for } \xi \geq 0, \tag{10c}$$

and the boundary and coupling conditions in the axial direction are given by:

$$\left|\frac{\partial \theta}{\partial \xi}\right|_{\xi \to \infty} < \infty, \tag{11a}$$

$$\Phi(0,\eta) = \theta(0,\eta) + F(0,\eta) \tag{11b}$$

$$\frac{\partial \Phi}{\partial \xi}(0,\eta) = \frac{\partial \theta}{\partial \xi}(0,\eta) + \frac{\partial F}{\partial \xi}(0,\eta) \tag{11c}$$

Integral Transformation

The solution of the considered problem is accomplished employing the Generalized Integral Transform Technique []. The solution process is started by defining the transformation pairs:

$$Transform \Rightarrow \bar{\phi}_n = \int_0^1 \eta \ \Phi(\xi,\eta)\Lambda_n(\eta) \ \mathrm{d}\eta, \tag{12a}$$

$$Inversion \Rightarrow \Phi(\xi,\eta) = \sum_{n=0} \frac{\phi_n(\xi)\Lambda_n(\eta)}{N_n}, \tag{12b}$$

$$Transform \Rightarrow \bar{\theta}_n = \int_0^1 \eta \ \theta(\xi,\eta)\Lambda_n(\eta) \ \mathrm{d}\eta, \tag{12c}$$

$$Inversion \Rightarrow \theta(\xi,\eta) = \sum_{n=0}^{\infty} \frac{\bar{\theta}_n(\xi)\Lambda_n(\eta)}{N_n}, \tag{12d}$$

where Λ_ns are orthogonal solution of a Sturm-Liouville problem. For the current application, a one-dimensional Helmholtz problem is selected:

$$\frac{1}{\eta}\frac{\mathrm{d}}{\mathrm{d}\eta}(\eta \ \Lambda_n'(\eta)) + \lambda_n^2\Lambda_n(\eta) = 0, \ \text{for } 0 \leq \eta \leq 1, \tag{13a}$$

$$\Lambda_n'(0) = 0, \ \beta \ \Lambda_n'(1) + (1-\beta)\Lambda_n(1) = 0, \tag{13b}$$

which leads to infinite nontrivial solutions in the form:

$$\Lambda_n(\eta) = J_0(\lambda_n\eta), \tag{14a}$$

for $= 1, 2, \ldots, \infty$. For the constant heat flux condition ($\beta = 1$) problem (13) admits a non-trivial constant solution for $\lambda_0 = 0$. As a result, one must also include definitions for this situation:

$$\Lambda_0(\eta) = \beta, \tag{14b}$$

such that for isothermal wall conditions ($\beta = 0$) trivial solutions are obtained.

The eigenvalues are calculated numerically, being obtained from the roots of the following equations:

$$J_0(\lambda_n) = 0, \ \text{for } \beta = 0, \tag{14c}$$

$$J_1(\lambda_n) = 0, \ \text{for } \beta = 1, \tag{14d}$$

Finally, the norms are obtained from:

$$N_n = \int_0^1 \eta \, \Lambda_n^2(\eta) \, \mathrm{d}\eta, \text{ for n} = 0,1,2,\ldots, \tag{15}$$

noting that for isothermal walls there is no need for calculating N_0 as the trivial solution is obtained for n = 0.

The transformation of the given problem is accomplished by multiplying equations (1a) and (10a) by $\Lambda_n \eta$, integrating within $0 \leq \eta \leq 1$, and applying the inversion formulas (12b) and (12d) to the non-transformable terms. This process yields the following coupled system of ODEs:

$$\frac{\mathrm{d}^2 \bar{\phi}_n}{\mathrm{d}\xi^2} - \sum_{m=0}^{\infty} A_{n,m} \frac{\mathrm{d}\bar{\phi}_m}{\mathrm{d}\xi} - \mathrm{Pe}^2 \lambda_n^2 \bar{\phi}_n(\xi) = 0, \text{ for } \xi \leq 0, \tag{16a}$$

$$\frac{\mathrm{d}^2 \bar{\theta}_n}{\mathrm{d}\xi^2} - \sum_{m=0}^{\infty} A_{n,m} \frac{\mathrm{d}\bar{\theta}_m}{\mathrm{d}\xi} - \mathrm{Pe}^2 \lambda_n^2 \bar{\theta}_n(\xi) = 0, \text{ for } \xi \geq 0, \tag{16b}$$

$$\left| \frac{\mathrm{d}\bar{\phi}_n}{\mathrm{d}\xi} \right|_{\xi \to -\infty} < \infty, \qquad \left| \frac{\mathrm{d}\bar{\theta}_n}{\mathrm{d}\xi} \right|_{\xi \to \infty} < \infty, \tag{16c}$$

$$\bar{\phi}_n(0) = \bar{\theta}_n(0) + b_n, \tag{16d}$$

$$\frac{\mathrm{d}\bar{\phi}_n}{\mathrm{d}\xi}(0) = \frac{\mathrm{d}\bar{\theta}_n}{\mathrm{d}\xi}(0) + d_n, \tag{16e}$$

for $n = 0,1,2,\ldots,\infty$. However, one should note that for $\beta = 0$ there is no need to calculate $\bar{\phi}_0$ and the summations in equations (16a) and (16b) should start from $m = 1$, such that there is no need for an equation for $n = 0$.

The coefficients $A_{n,m}, b_n$, and d_n are given by:

$$A_{n,m} = \frac{\mathrm{Pe}^2}{N_{m}} \int_0^1 \eta \, u^*(\eta) \Lambda_m(\eta) \Lambda_n(\eta) \, \mathrm{d}\eta, \tag{17a}$$

$$b_n = \int_0^1 \eta \, F(0,\eta) \, \Lambda_n(\eta) \, \mathrm{d}\eta, \tag{17b}$$

$$d_n = \int_0^d \eta \left(\frac{\partial F}{\partial \xi} \right)_{\xi=0} \Lambda_n(\eta) \, \mathrm{d}\eta, \tag{17c}$$

In order to solve system (16), the infinite system representation must be truncated to a finite number of terms, which is denoted the truncation order. Once truncated, the resulting system can be solved numerically using a commercially or publicly available ODE system solver. Nevertheless, an analytical alternative to the ODE integration can be achieved if the truncated systems are reduced to first order forms by writing it in terms of new unknown vectors x and y:

$$x(\xi) = (\bar{\phi}_0, \bar{\phi}_1, \bar{\phi}_2, \ldots, \bar{\phi}_{n_{\max}}, \bar{\phi}_0', \bar{\phi}_1', \bar{\phi}_2', \ldots, \bar{\phi}_{n_{\max}}' (\xi)), \tag{18a}$$

$$y(\xi) = (\bar{\theta}_0, \bar{\theta}_1, \bar{\theta}_2, \ldots, \bar{\theta}_{n_{\max}}, \bar{\theta}_0', \bar{\theta}_1', \bar{\theta}_2', \ldots, \bar{\theta}_{n_{\max}}' (\xi)), \tag{18b}$$

Where a condition of a prescribed wall flux in both upstream and downstream regions has been considered. For isothermal walls x needs not include $\bar{\phi}_0$ nor $\bar{\phi}_0'$, and y needs not include $\bar{\theta}_0$ nor $\bar{\theta}_0'$.

By employing the new unknown vectors, equations (16a) and (16b) are rewritten in the following forms:

$$\frac{d\mathbf{x}}{d\xi} = \mathbf{M}\, \mathbf{x}, \quad \frac{d\mathbf{y}}{d\xi} = \mathbf{M}\, \mathbf{y} \tag{19}$$

Where $\mathbf{M}$ is a block matrix defined as:

$$\mathbf{M} = \left(\begin{array}{c|c} \mathbf{0} & \mathbf{I} \\ \hline \mathbf{D} & \mathbf{A} \end{array} \right), \tag{20}$$

in which the sub-matrix $\mathbf{A}$ is given by the coefficients $A_{n,m}$ and $\mathbf{D}$ is given by:

$$D_{n,n} = \mathrm{Pe}^2 \lambda_n^2. \tag{21}$$

Once the transformed ODEs have been written in the modified form given by equations (19), a solution can be obtained using analytical integration. If the eigenvalues and eigenvectors of $\mathbf{M}$ are calculated, the solution for the components of x and y can be written in the following form:

$$x_k(\xi) = \sum_{l=1}^{k_{max}} Q_{k,l}\, c_l^- \exp(\mu_l \xi), \qquad \text{for} \qquad l = 1, 2, \cdots, k_{max}, \tag{22a}$$

$$y_k(\xi) = \sum_{l=1}^{k_{max}} Q_{k,l}\, c_l^+ \exp(\mu_l \xi), \qquad \text{for} \qquad l = 1, 2, \cdots, k_{max}, \tag{22b}$$

in which $k_{max} = 2\, n_{max}$ for constant wall temperature and $k_{max} = (2\, n_{max} + 1)$ for constant wall flux. The coefficients $Q_{k,l}$ originate from matrices $\mathbf{Q}$ that contain the eigenvectors of $\mathbf{M}$ as columns. The coefficients μ_l are the eigenvalues of $\mathbf{M}$, while c_l^- and c_l^+ are arbitrary constants, to be determined from boundary conditions. The matrix $\mathbf{M}$ yields about n_{max} positive eigenvalues and n_{max} negative eigenvalues, such that roughly half of the arbitrary coefficients can be directly eliminated to satisfy the boundary conditions at $\xi \to \infty$ and $\xi \to -\infty$:

$$c_k^+ = 0 \quad \text{if} \quad \mu_k > 0, \quad \text{for} \quad k = 1, 2, \cdots, k_{max,} \tag{23a}$$

$$c_k^- = 0 \quad \text{if} \quad \mu_k < 0, \quad \text{for} \quad k = 1, 2, \cdots, k_{max,} \tag{23b}$$

The remaining c_k^- and c_k^+ values are calculated directly from solving the linear algebraic system that stems from the coupling conditions at $\xi = 0$. These equations can be directly combined into a single vector equation:

$$\mathbf{Q}\, \mathbf{c}^- = \mathbf{Q}\, \mathbf{c}^+ + \mathbf{h}, \tag{24}$$

where the vectors $\boldsymbol{c}^-$ and $\boldsymbol{c}^-$ are given by the coefficients c_n^- and c_n^+, the vector $\boldsymbol{h}$ is defined as:

$$h = \left(b_0, b_1, b_2, \rightleftharpoons, b_{max}, d_0, d_1, d_2, \rightleftharpoons, d_{max} \right) \tag{25}$$

After using equations (23), equation (24) is directly solved to yield the remaining unknown coefficients, and the solution of the transformed potentials in both upstream and downstream regions is complete. Once $\overline{\phi}$ and $\overline{\theta}$ are determined, the temperature profiles are obtained directly from the inversion formulas (12b) and (12d) combined with the separation formula defined by equation (8).

Results and Discussion

After describing the problem formulation and solution methodology, numerical results are presented. The first set of results are dedicated to examining the solution convergence for the two-different boundary condition cases. Table 1 presents the calculated values of the Nusselt number for different axial positions (all within the downstream region) and different values of the Péclet number for the isothermal walls configuration. As can be seen, the convergence behavior follows a pattern in which notably better convergence rates are obtained as one moves downwards with the flow. For positions in this region, 40 to 60 terms (depending on the value of Péclet) are sufficient for ensuring a Nusselt values with six converged digits. For positions near the boundary condition discontinuity ($\xi = 0$) a worse convergence rate is seen, and the convergence becomes highly dependent on the Pe value. For Pe = 1 and $\xi = 10-3$ not even a single converged digit is obtained with 100 terms; however, for Pe = 100, 100 terms in the series leads to four converged digits.

Also on this table, the converged solution of the case with no axial diffusion (labeled as Pe = ∞) is presented for verification purposes. As one can observe from these values, it is clear that as Pe is increased, the Nu values gradually approach the solution with no axial diffusion.

The next table (Tab. 2) presents similar Nusselt number convergence results for the isoflux walls situation. As one can observe from the presented results, again, better convergence rates are seen in regions further downstream; however, when compared to the isothermal walls case, much better convergence rates are obtained. In fact, in several occasions, as much as five terms in the series yield six converged Nu-digits. When looking into the convergence dependence on the Péclet number, one notices that the convergence in positions near the channel entrance is notably less dependent on Pe than for the case with isothermal walls. Finally, when comparing the Nusselt number values with axial diffusion with the case with no axial diffusion, one notices that with increasing Pe number, the converged Nusselt values approach the Pe = ∞ case, as expected.

The last set of results examine the behavior of the Nusselt number for different values of the Péclet number. Fig. 1 shows the distribution of Nu in the thermal developing region for different Péclet values. As can be seen, the Péclet number has a much more pronounced effect on the Nusselt values for positions near the channel entrance. In fact, although hard to see in the presented scale, different Péclet values also yield different Nu values for the isothermal walls conditions (as expected) in the thermal developed region ($\xi > 1$), which can be seen from the data in table 1.
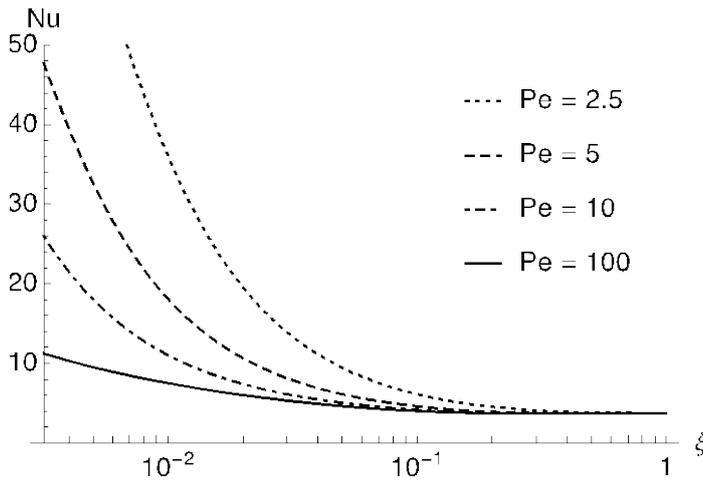
Figure 1 – Nusselt distribution for isothermal wall

Subsequently, figure 2 depicts the Nu distribution in the thermally developing region for different Péclet values, for the case with an insulated downstream and a uniform heated upstream. As one can infer from the presented results, a somewhat different behavior is seen when compared to the previous case. Firstly, the Nu values are independent of the Péclet number in the developed region (as expected, and seen in table 2). Secondly, although the dependence on the Péclet number becomes very notable as one moves further into the thermal entrance region, the Nu values have a general tendency of decreasing with increasing Péclet for this case, rather than increasing (which was observed for the isothermal walls case).
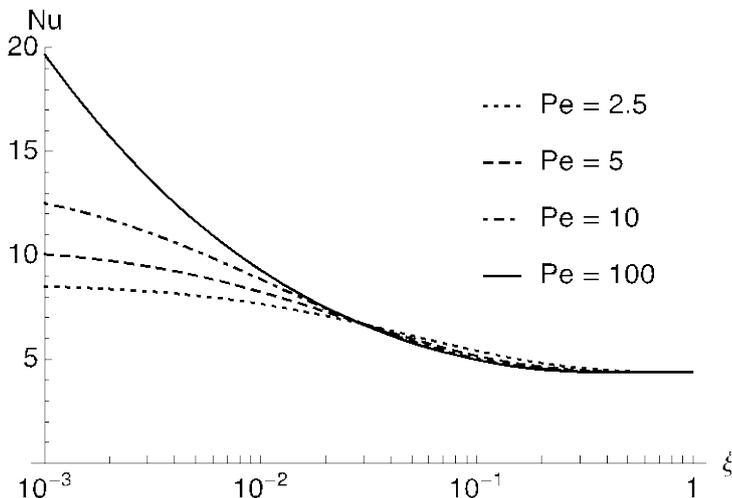
Figure 2 Nusselt distribution for insulated/isoflux wall

Table 1

Isothermal walls

n_{max}	ξ=0.001	ξ=0.01	ξ=0.1	ξ=1	ξ=10
Pe=2					
10	28.8358	22.875	7.07583	3.8087	3.80203
20	54.1035	34.0065	7.07748	3.80826	3.8016
40	100.116	43.107	7.07682	3.8082	3.80154
60	140.698	45.6955	7.07676	3.8082	3.80153
80	176.489	46.4317	7.07674	3.8082	3.80153
100	208.054	46.641	7.07674	3.8082	3.80153
conv.	443.795	46.7239	7.07673	3.8082	3.80153
Pe=5					
10	24.7794	15.09	4.61058	3.69582	3.69581
20	43.1881	17.4615	4.60843	3.69527	3.69526
40	72.3622	18.0455	4.60815	3.6952	3.69519
60	93.6745	18.0689	4.60813	3.6952	3.69518
80	109.242	18.0694	4.60812	3.69519	3.69518
100	120.613	18.0693	4.60812	3.69519	3.69518
conv.	151.419	18.0691	4.60811	3.69519	3.69518
Pe=10					
10	23.1915	10.8647	4.16302	3.66821	3.66821
20	36.6831	11.025	4.16151	3.66763	3.66763
40	53.7184	11.0232	4.16131	3.66756	3.66755
60	62.8062	11.0223	4.16129	3.66755	3.66755
80	67.6537	11.0221	4.16129	3.66755	3.66755
100	70.2391	11.022	4.16128	3.66754	3.66754
conv.	73.1923	11.0219	4.16128	3.66754	3.66754
Pe=50					
10	20.7269	7.67478	4.01245	3.65792	3.65792
20	21.8546	7.64915	4.01115	3.65733	3.65733
40	22.0204	7.64572	4.01098	3.65725	3.65725
60	22.0114	7.64537	4.01096	3.65724	3.65724
80	22.0071	7.64529	4.01096	3.65724	3.65724
100	22.0055	7.64525	4.01096	3.65724	3.65724
conv.	22.0039	7.64522	4.01096	3.65724	3.65724
Pe=100					
10	19.8799	7.54188	4.0077	3.65758	3.65758
20	18.1528	7.51788	4.0064	3.65699	3.65699
40	17.9536	7.51471	4.00623	3.65692	3.65692
60	17.9428	7.51438	4.00622	3.65691	3.65691
80	17.9404	7.5143	4.00621	3.65691	3.65691
100	17.9395	7.51427	4.00621	3.65691	3.65691
conv.	17.9386	7.51424	4.00621	3.65691	3.65691
Pe=∞					
conv.	16.264	7.47038	4.00463	3.65679	3.65679

Table 2

n_{max}	ξ=0.001	ξ=0.01	ξ=0.1	ξ=1	ξ=10
			Isoflux walls		
			Pe=2		
10	7.9516	7.44095	5.50366	4.36881	4.36364
20	8.10429	7.5016	5.50319	4.36881	4.36364
40	8.17579	7.51492	5.50314	4.36881	4.36364
60	8.19683	7.5161	5.50313	4.36881	4.36364
80	8.20611	7.51626	5.50313	4.36881	4.36364
100	8.21101	7.51628	5.50313	4.36881	4.36364
conv.	8.22006	7.51629	5.50313	4.36881	4.36364
			Pe=5		
5	9.40303	8.15929	5.1761	4.36377	4.36364
10	9.77981	8.23411	5.17148	4.36377	4.36364
20	9.97406	8.24913	5.17099	4.36377	4.36364
40	10.0564	8.24983	5.17093	4.36377	4.36364
60	10.0764	8.24977	5.17092	4.36377	4.36364
80	10.0837	8.24975	5.17092	4.36377	4.36364
100	10.0869	8.24975	5.17092	4.36377	4.36364
conv.	10.0903	8.24974	5.17092	4.36377	4.36364
			Pe=10		
5	11.819	8.9259	5.04433	4.36365	4.36364
10	12.2099	8.86054	5.04022	4.36365	4.36364
20	12.4298	8.85146	5.03973	4.36365	4.36364
40	12.5093	8.85019	5.03968	4.36365	4.36364
60	12.5235	8.85007	5.03967	4.36365	4.36364
80	12.5272	8.85004	5.03967	4.36365	4.36364
100	12.5284	8.85003	5.03967	4.36365	4.36364
conv.	12.529	8.85002	5.03967	4.36365	4.36364
			Pe=50		
5	20.9235	9.53499	4.97964	4.36364	4.36364
10	18.9231	9.30816	4.97572	4.36364	4.36364
20	18.6708	9.29342	4.97524	4.36364	4.36364
40	18.6412	9.2916	4.97519	4.36364	4.36364
60	18.6382	9.29142	4.97518	4.36364	4.36364
80	18.6375	9.29138	4.97518	4.36364	4.36364
100	18.6372	9.29136	4.97518	4.36364	4.36364
conv.	19.6625	9.29551	4.97283	4.36364	4.36364
			Pe=100		
5	24.5256	9.53521	4.97728	4.36364	4.36364
10	20.2961	9.31245	4.97337	4.36364	4.36364
20	19.7092	9.2976	4.9729	4.36364	4.36364
40	19.6679	9.29577	4.97284	4.36364	4.36364
60	19.6641	9.29559	4.97284	4.36364	4.36364
80	19.6631	9.29554	4.97283	4.36364	4.36364
100	19.6628	9.29553	4.97283	4.36364	4.36364
conv.	19.6625	9.29551	4.97283	4.36364	4.36364
			Pe=∞		
conv.	19.9868	9.29502	4.97205	4.36364	4.36364

Conclusions

This paper presented an analytical solution to the extended Graetz problem with axial heat diffusion in an infinite domain configuration with a step change in wall boundary condition at the axial coordinate origin. The channel geometry comprises circular channels (tubes or pipes), and both isothermal walls and insulated/isoflux walls were analyzed as boundary conditions. The solution methodology was based on the Generalized Integral Transform Technique using simple eigenfunctions as a basis for the sough solution. The eigenfunctions are obtained from a Helmholtz problem, and all integral coefficients can be calculated analytically. In fact, the entire solution procedure is analytical, except for the numerical calculation of matrix eigenvalues and eigenvectors. The solution method was verified to lead to converged values which are in accordance with physically expected results. The results also match the values without axial diffusion as the Péclet number is increased. After demonstrating the convergence of the solution, the Nusselt number distribution for different Péclet values was analyzed, and the results are also in accordance with expected literature values. As final comments one should mention that the same solution procedure can be used for any dynamically developed velocity profile, as it occurs in many other occasions. Also, the methodology can be easily extended to other configurations such as other channel geometries, different wall heating conditions, and vicious and other flow heating effects.

Acknowledgements

References

Y.BayazitogluandM.N.Ozisik. On the solution of Graetz type problems with axial conduction. *International Journal of Heat and Mass Transfer,* 23(10):1399–1402, 1980.

B. Cetin, A. G. Yazicioglu, and S. I. Kakaç, Fluid flow in microtubes with axial conduction including rarefaction and viscous dissipation. *International Communications in Heat and Mass Transfer,* 35(5):535–544, 2008. ISSN 0735-1933.

Barbaros Çetin, Almila Guvenc¸Yazicioglu, and Sadik Kakac. Slip-flow heat transfer in microtubes with axial conduction and viscous dissipation – an extended Graetz problem. *International Journal of Thermal Sciences,* 48(9): 1673–1678, 2009. ISSN 1290-0729.

R. M. Cotta. *Integral Transforms in Computational Heat and Fluid Flow.* CRC Press, Boca Raton, FL, 1993.

M. A. Ebadian and H. Y. Zhang. An exact solution of extended Graetz problem with axial heat conduction. *International Journal of Heat and Mass Transfer,* 32(9):1709–1717, 1989.

M. A. Ebadian and H. Y. Zhang. Effects of heat generation and axial heat conduction in laminar flow inside a circular pipe with a step change in wall temperature. *International Communications in Heat and Mass Transfer,* 17(5):621–635, 1990.

L.Graetz. Uber die warmeleitungsfahigkeit von flussigkeiten. *Annalen der Physik*, 254(1):79–94,1882.

C.J. Hsu. Exact solution to entry-region laminar heat transfer with axial conduction and the boundary condition of the third kind. *Chemical Engineering Science,* 23(5):457–468, 1968.

P.R. Johnston. Axial conduction and the graetz problem for a bingham plastic in laminar tube flow. *International Journal of Heat and Mass Transfer,* 34(4-5):1209–1217, 1991.

H. C. Ku and D. Hatziavramidis. Chebyshev expansion methods for the solution of the extended Graetz problem. *Journal of Computational Physics,* 56(3):495–512, 1984.

C. Laohakul, C. Y. Chan, K. Y. Look, and C. W. Tan. On approximate solutions of the graetz problem with axial conduction. *International Journal of Heat and Mass Transfer,* 28(3):541–545, 1985.

M. L. Michelsen and J. Villadsen. The Graetz problem with axial heat conduction. *International Journal of Heat and Mass Transfer,* 17(11):1391–1402, 1974.

T. Min, J. Y. Yoo, and H. Choi. Laminar convective heat transfer of a bingham plastic in a circular pipei. Analytical approach—thermally fully developed flow and thermally developing flow (the Graetz problem extended). *International Journal of Heat and Mass Transfer,* 40(13):3025–3037, 1997.

R. G. Najjar and C. Laohakul. An approximate solution to the Graetz problem with axial conduction and prescribed

wall heat flux. *International communications in heat and mass transfer,* 13(3):315–324, 1986.

S. Olek. Heat transfer in duct flow of non-newtonian fluids with axial conduction. *International Communications in Heat and Mass Transfer,* 25(7):929–938, 1998.

A. Sharma and S. Chakraborty. Semi-analytical solution of the extended Graetz problem for combined electroosmotically and pressure-driven microchannel flows with step-change in wall temperature. *International Journal of Heat and Mass Transfer,* 51(19-20):4875–4885, 2008.

B. Vick and M. N. Ozisik. An exact analysis of low peclet number heat transfer in laminar flow with axial conduction. *Letters in Heat and Mass Transfer,* 8(1):1–10, 1981.

B.Vick, M.N.Ozisik, and Y. Bayazitoglu. A method of analysis of low peclet number thermal entry region problems with axial conduction. *Letters in Heat and Mass Transfer,* 7(4):235–248, 1980.

B.Vick, M.N. Ozisik, and D.F. Ullrich. Effects of axial conduction in laminar tube flow with convective boundaries. *Journal of the Franklin Institute,* 316(2):159–173, 1983.

Hybrid contour method/eigenstrain model to predict residual stress in glass

B.A. Balan[1] and *M. Achintha[1]

[1]Faculty of Engineering and Environment, University of Southampton, SO17 1BJ, UK.

*Corresponding author: Mithila.Achintha@soton.ac.uk

Abstract

Using finite element (FE) models the residual stress present in a glass sample was constructed using the knowledge of surface deformation resulted in from the stress relaxation along a newly cut plane. The residual stress profile, validated with a scatter light polariscope, was then used to accurately establish the misfit strain (i.e. eigenstrain) of the original glass specimen. The paper shows that once the underlying eigenstrain distribution has been determined, the complete residual stress distribution can simply be determined by incorporating the eigenstrain profile as a misfit strain in an appropriate FE model. The results show that stress depth profile generated in float glass is parabolic. It is also shown that the hybrid model enables modelling residual stress in new geometries (e.g. stress concentrations around the hole) and/or during subsequent loading application, by simply using the knowledge of eigenstrain depth profile.

Keywords: Contour Method, Eigenstrain, Finite Element, Glass, Residual Stress

Introduction

Over the last decades, owing to its unique properties, designers began to use glass as a load bearing structural material in buildings [IStructE (1999)]. One of the greatest difficulties that inhibits accurate prediction of the structural response of commercial glass is the lack of comprehensive analytical/numerical tools to predict residual stress distributions generated due to manufacturing cooling process. The current design guidelines in the UK are lacking in complete design methodology, most of them being based on rules of thumb [IStructE (1999)]. *The Institution of Structural Engineers* design guidelines [IStructE (1999)] provide design permissible tensile stress values (e.g. 28 N/mm^2 for float glass with a thickness up to 6 mm) but recommend that: "In the absence of code-based allowable tensile stresses it is left to the judgment of the engineer what figures to adopt". This often results in over designed conservative structures.

The misfit strains developed during the cooling of glass in the manufacturing process generate residual stresses. As it is known the effect of residual stresses can be significant in the strength prediction of materials [Withers (2007)]; thus negligence of this can lead to premature failure of the structures. Glass, in particular, is susceptible to brittle fracture failures due to the presence of inevitable surface defects.

The non-crystalline microstructure of the glass makes it impossible to determine the residual stress using conventional experimental methods such as X-ray or neutron diffraction. A scattered light method [Aben et. al. (2008)], which uses the magnitude of birefringence in glass, may be used to determine residual stresses. However, the method is limited by not being able to accurately determine the through thickness stress profiles at certain depths, especially in annealed glass. The current paper presents a validated hybrid contour /eigenstrain method to characterise the full field residual stresses in commercially available float and tempered glass.

Methodology

The contour method, originally developed by Prime (2001) to model residual stresses in metallic components, was used in the present study to determine the residual stress in glass. Although the contour method provides information about the residual stresses, it is not able to advance any details about the cause of the residual stresses (i.e. eigenstrain). Using the knowledge of the residual stress distribution constructed from the contour method, the eigenstrain distribution in the sample was computed from an inverse eigenstrain analysis developed by Korsunsky (2009). The knowledge of the eigenstrain profile was then used to compute the stress concentrations around a hole in plate with new geometry, and/or under applied loads. Fig. 1 presents the step-by-step procedure of the current method.

Original glass sample

Cut sample in half and
measure surface displacements

Displacement contour of the cut plane

Finite element analysis

Residual stress

Inverse eigenstrain analysis

Eigenstrain distribution

Finite element model

Residual stress in new geometries/during applied loadings
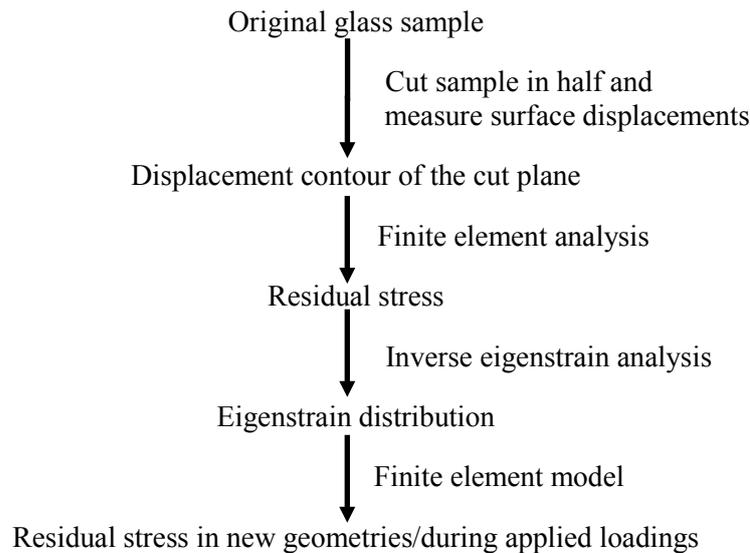
Figure 1. Step-by-step procedure of contour/eigenstrain hybrid method

Construction of residual stress using contour method

In the contour method, the residual stress distribution is determined by incorporating the surface contour perpendicular to the cut plane as a displacement boundary condition in a static finite element (FE) model representing an initially stress-free half sample of the original specimen. Therefore, the determination of the displacement contour of the cut plane is a prerequisite in the present analysis. The contour method principle illustrated in Fig. 2 shows that after cutting the sample in half (Fig. 2b), the surface deformation (average of surface deformation of both sides is used to account for possible shear effects) is used to construct the residual stress developed in one half of the specimen as a response of forcing back the deformed surface to the initially planar state (Fig. 2c). The contour method proved its success in modelling residual stress in steel and metal alloys in different applications e.g. bent elements [Pagliaro (2008)], welds [Hosseinzadeh (2011)]. A comprehensive discussion of theory of the contour method can be found elsewhere [Pagliaro (2008); Johnson (2008)].
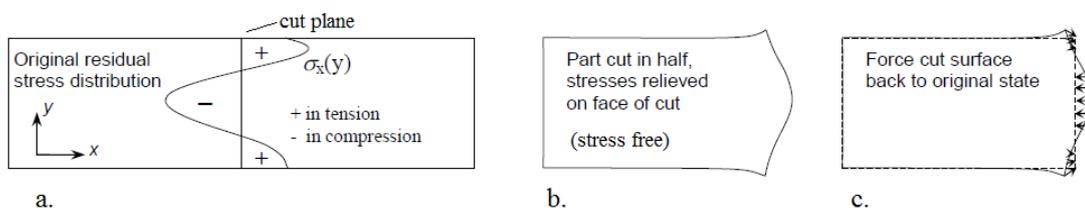
Figure 2. Contour method principle [Prime (2001)]

The analysis of a comercially available float glass sample (Pilkington Glass) of 150 x 100 x 10 mm is discussed below to demonstrate the application of the method to construct residual stress field. Making a cut which is free from defects is not trivial, in particular due to the brittle nature of glass. The most suitable cutting technique used for metal specimens is the wire electric discharge machine (EDM) [Prime (2001)], where the material removal is done by spark erosion. However the method cannot be used to cut glass and different alternatives such as diamond disk cutting and water-jet cutting were investigated. The results revealed that water-jet cutting (done by a comerical contractor) with a jet diameter of ~1 mm and a 80 mesh garnet grade provides a good "cut plane". It should be appreciated that in the experiments sacrificial glass pieces, as shown in Fig. 3a, were used on either side of the main specimen to eliminate the "edge effects".

After the sample was cut into two halves, the displacement contour of both cut surfaces was measured using a 3-D micro-coordinate system which offers great accuracy being able to achieve a vertical resolution of up to 10 nm [Alicona, 2.1.5]. It was observed that the GFM G4 10x objective offers the optimum magnification for determining the displacement contour profile of the sample. The 3-D contour presented in Fig. 3b was done with a sampling distance of 1.75 μm. The measurement was made along the thickness of the sample (10 mm) and accros a width of 1 mm, chosen to eliminate the effects due to potential local defects (e.g. micro-cips).
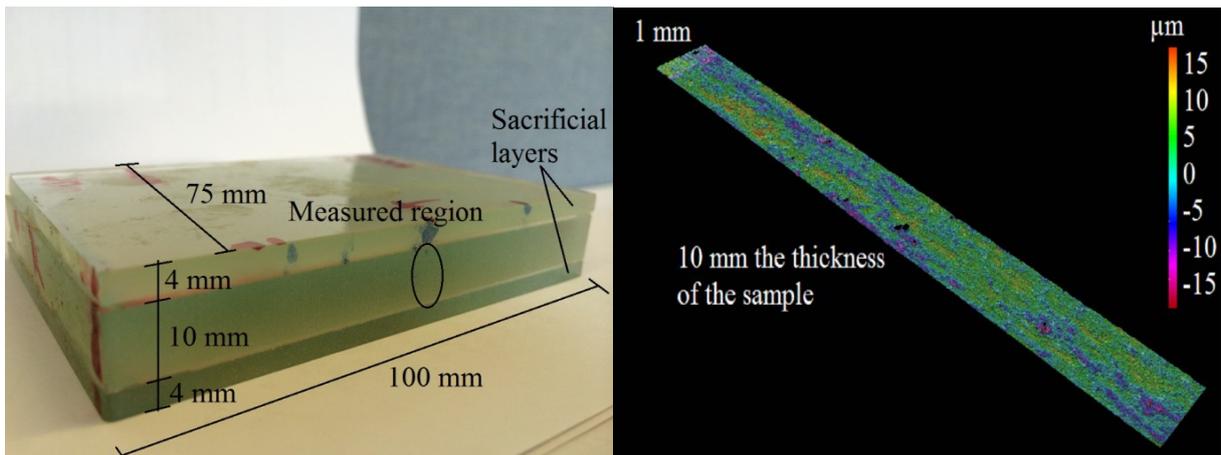
Figure 3. a. Half of the sample after cutting (left), b. 3-D contour of the cut-plane (right)

It should be appreciated that the deformations due to stress relaxation in this specimen are ~1-2 μm. However, the current work is a feasibility study to demonstrate the application of the present hybrid model to predict stresses in tempered glass, which is widely used in commercial applications. Because of the significantly high stresses in tempered glass it is expected that displacement along a cut plane in tempered glass will be more significant than that in an annealed glass specimen.

The surface deformations of each side of the cut shown in Fig. 4a are very similar. Separate measurements across the length of the sample were taken and it was concluded that the surface deformation is mostly uniform in the lateral direction and varies only along the thickness. For instance, Fig.4b shows the measured contour depth profile at three different locations within the left cut plane; all three profiles are very similar. Therefore, it is appropriate to incorporate the surface contour of the cut-plane into a FE model as a polynomial curve that varies in thickness. The average depth profile presented in Fig. 4a was represented by a best fit 2nd order polynomial with coefficients: -0.045, 0.419, -0.616.
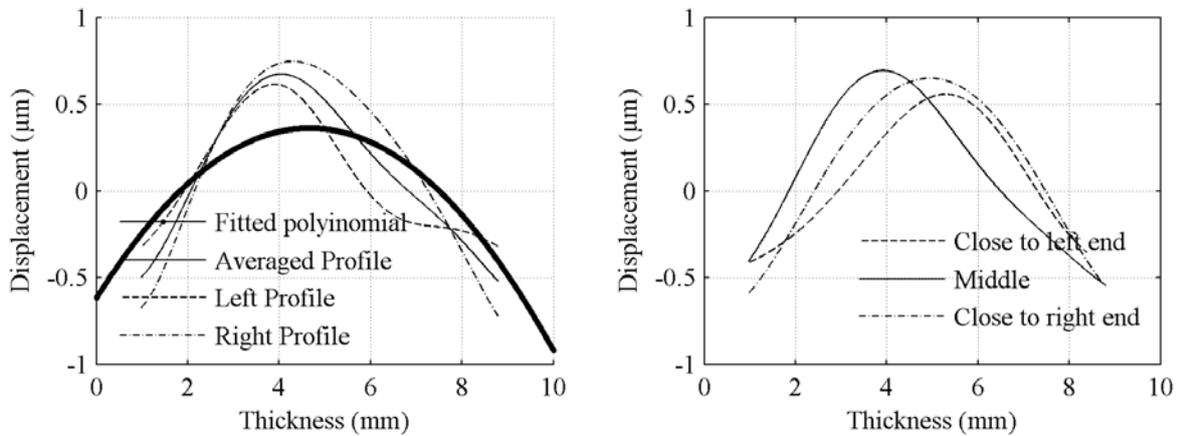
Figure 4. a. Averaged middle profile and fitted polynomial (right), b. Displacement profiles at various locations (left)

ABAQUS/Standard [Abaqus, 6.9-3] was used in the present study to model the residual stress distribution. A 3-D model with 8-noded, linear brick stress elements (3D8R) was used in the simulation. With a linear-elastic behaviour, it was appropriate to assume a material model with Young modulus =70 GPa and a Poisson's ratio =0.23 to characterise material properties of glass. The residual stress distribution was conveniently determined by incorporating the approximate polynomial curve of the displacement contour (Fig. 4a) in a FE model representing the initially stress-free half sample of the original glass specimen.

The residual stress distribution computed using the above FE simulation is presented in Fig. 5a (only the stress component normal to the cut surface is shown here since it is the most relevant stress distribution). The results show a parabolic stress distribution, with tension at the outer surface (~8 MPa) and compression (~7 MPa) at the mid-thickness. The depth of the tension zones on each side is 2 mm (~20% of the specimen thickness) and is balanced by a middle compression zone of 6 mm thick (~60% of the overall thickness). It is worth noting that the "edge effects" due to FE simulation could not be avoided, thus the surface stress predicted from the present FE model may be slightly overestimated. Thus, the above quoted surface tension value is actually the value slightly below/above (~1 mm) the actual surface. Fig. 5b shows that the cumulative force along the depth is zero when integrating the stresses along the depth profile.
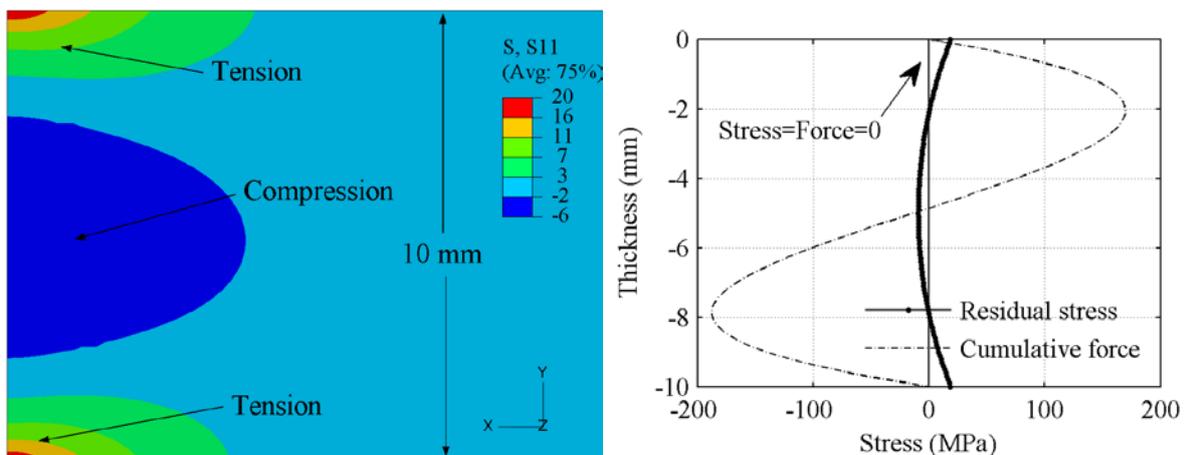
Figure 5. a. Residual stress (left), b. Cumulative force distribution (right)

The unique characteristic of the current method is that the full stress field can be predicted fulfilling the overall equilibrium, compatibility and boundary conditions. It should be noted that the magnitudes of the predicted residual stresses agree with the expected stresses in practice [Geandier et. al (2003)]. Experiments using a scatter light polariscope are currently being undertaken to validate the present results.

Eigenstrain profile estimation using the constructed residual stress distribution

Although, the contour method can be used to model the residual stress in a given glass specimen, a separate experiment programme is required to predict the stress distribution in a new specimen. The residual stress is a response to the eigenstrain developed in the specimen during glass manufacturing process. Thus, as shown hereafter, once the knowledge of the eigenstrain distribution is available the stress state in real-life practical structural elements can be determined in a computationally efficient manner.

The eigenstrain method [Korsunsky (2009)], which is used here to determine the eigenstrain profile, uses a least square approach to determine the unknown eigenstrain distribution based on the residual elastic stresses measured at a finite number of locations. The technique was successfully used [Achintha and Nowell (2011)] to reconstruct the full residual stress field in alloy materials due to laser shock peening.

Assume eigenstrain distribution
(A Chebyshev series)

FE model for each
polynomial of the series

Resultant residual stress distribution

Minimisation of the mismatch
between the measured
and calculated stress

Accurate eigenstrain profile of the sample

Finite element model

Residual stress in new geometries/during applied loadings
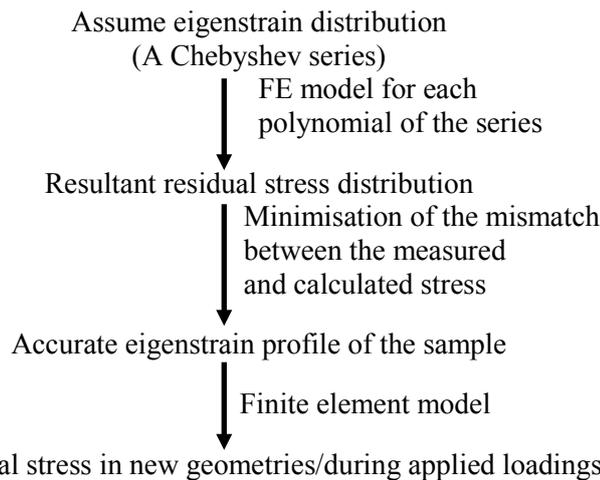
Figure 6. Step-by-step procedure of an inverse eiegnstrain analysis of the hybrid method

For the glass sample discussed previously, the eigenstrain distribution was considered to be uniform in the lateral direction, varying only with thickness. Initially an eigenstrain profile represented as a series of Chebyshev polynomials [Mason and Handscomb (2003)] was assumed (although alternative polynomial choices are possible). The number of polynomials in the series is to some extent arbitrary, but it should be large enough to capture the exact form of the eigenstrain distribution accurately. In the present study the analysis was done for a different number of polynomials to ensure that the result is independent of the value chosen. On separate FE models each polynomial of the assumed Chebyshev series was implemented individually and the respective residual stress in the specimen was determined. The response of the specimen to the applied eigenstrain is elastic, thus the resultant residual stress distribution caused by the original assumed eigenstrain distribution is the sum of each individual residual stress. Using a least-square analysis between the predicted stress and the corresponding measured data, the accurate estimate of the actual eigenstrain distribution was established. It should be appreciated that the stress values determined previously using the contour method were used as experimental data in this analysis. Once the best estimate of the eigenstrain distribution has been established, the residual stress

distribution can be determined in the usual way by incorporating this eigenstrain distribution in a FE model. The step-by-step procedure of the analysis technique is presented in Fig. 6.

Fig. 7a shows the "best estimate" of the eigenstrain depth profile and Fig. 7b shows the comparison between the residual stress depth profile determined from the earlier contour analysis and that from the eigenstrain analysis. From Fig. 7b it is evident that, as expected, the predictions from the eigenstrain method agree with that constructed from the contour method. The small mismatch between stress profiles is related to the procedure in which the residual stresses were determined. The contour method uses boundary displacement to obtain residual stress and thus the residual stress distribution is correctly predicted only at that edge. In the case of eigenstrain method the full residual stress was determined as an overall response of the model to the eigenstrain profile.
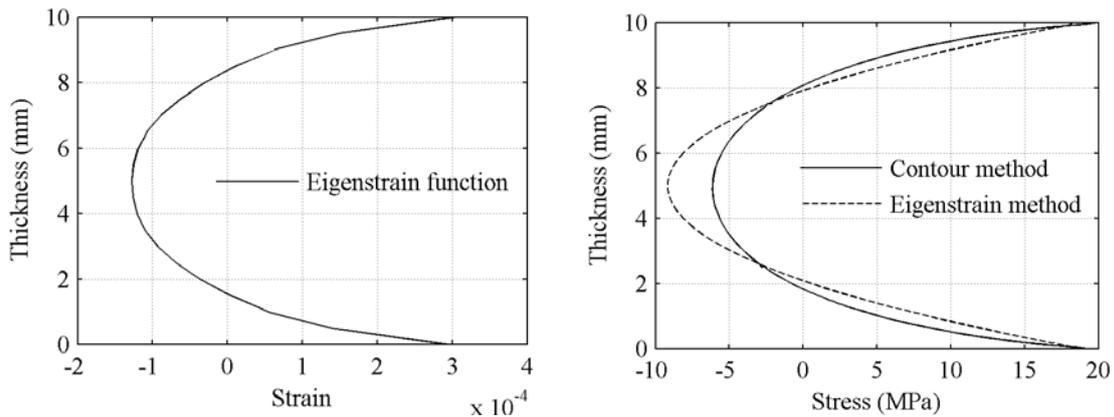
Figure 7. a. Computed eigenstrain distribution (left) b. Comparison of stress profile (right)

Prediction of stress distribution in different structural elements

The knowledge of the eigenstrain distribution allows determining the structural response of real-life structural glass elements of practical geometries and/or under applied loading.

For instance, the knowledge of the eigenstrain profile was used to study the effects due to geometry in a practical glass element. The results of a glass plate (150 x 100 x 10 mm) with a 20 mm diameter central hole, under uni-axial tensile loads of 10 and 20 MPa (X direction) are discussed. Symmetry conditions are used to model only a quarter of the specimen. As it can be seen in Fig. 8a, after the eigenstrain distribution was implemented to the FE model, a full residual stress field (only the middle principal stress is shown in Fig. 8a) was achieved in the sample (e.g. B) fulfilling the equilibrium conditions at the boundaries (e.g. A, C, D).
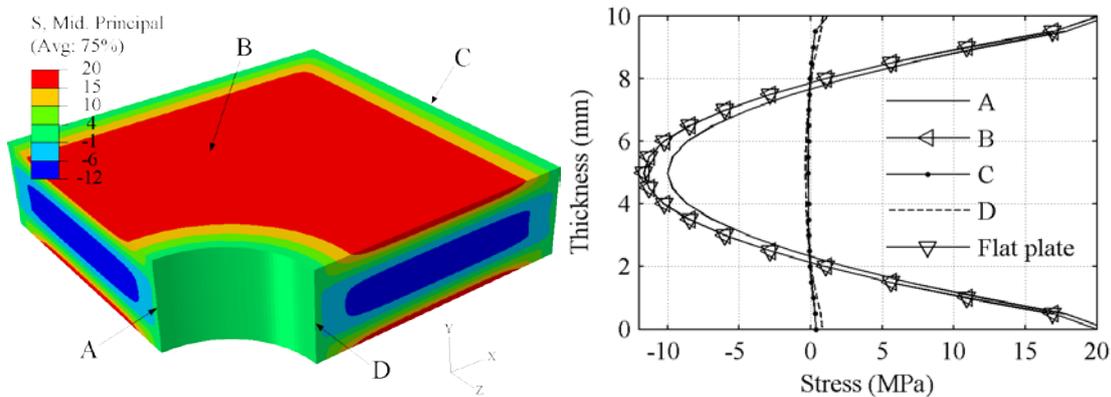
Figure 8. a. Residual stress field (Mid. Principal) b. Stress profile in X direction (S11) at different locations in the sample

Initially the sample has no applied load and the presence of the hole does not influence the residual stress distribution. The results presented in Fig. 8b show that the residual stress in X direction (i.e. along the direction of the applied uni-axial tension) across path A of the hole, and across path B (through thickness) in the plate matches the one of the flat sample without the hole.

Under tensile load, the stress distribution in the sample is no longer uniform; stress concentrations around the hole have developed as can be seen in Fig. 9a. Fig. 9b presents the same sample as before without any initial residual stress. As expected, the magnitude of the stresses in the sample if there was no initial residual stress is lower than that in the sample incorporating residual stresses. From both models it is evident that path A represents a locus where stress concentrations arise, but only the model incorporating initial residual stress is able to provide a comprehensive analysis of stress distribution and evolution in the sample during loading.

It is expected that residual stresses distribution affects the failure of a structural element. If, for example, the exampled considered here was to have an ultimate tensile load of 40 MPa, the FE glass model not incorporating the initial residual stress distribution would result in a satisfying structural design. Whereas, as shown in Fig. 10a, in the case of the FE glass model, in which the residual stress distribution was considered, it is clear that the ultimate limit was reached and the structural element might unexpectedly fail.

Considering the stress distribution along the path A (Fig. 9) was a favourable case, because both models were presenting stress concentration there. If one is to consider a random path away from the edges (path B) then accounting for stress distribution in the analysis makes a significant difference. As can be seen in Fig. 10a the results show that the residual stress (RS) completely changes the stress profile and the magnitude of the surface tension is more than double the value that was computed using a FE model without any initial residual stress. This proves that by incorporating residual stresses into analysis facilitates the modelling of the full stress field generated during subsequent loading. The results presented in Fig. 10b show that even though the step size of the load was constant (10 MPa) the magnitude by which the residual stress distribution (across path A) for each step increased was not constant.
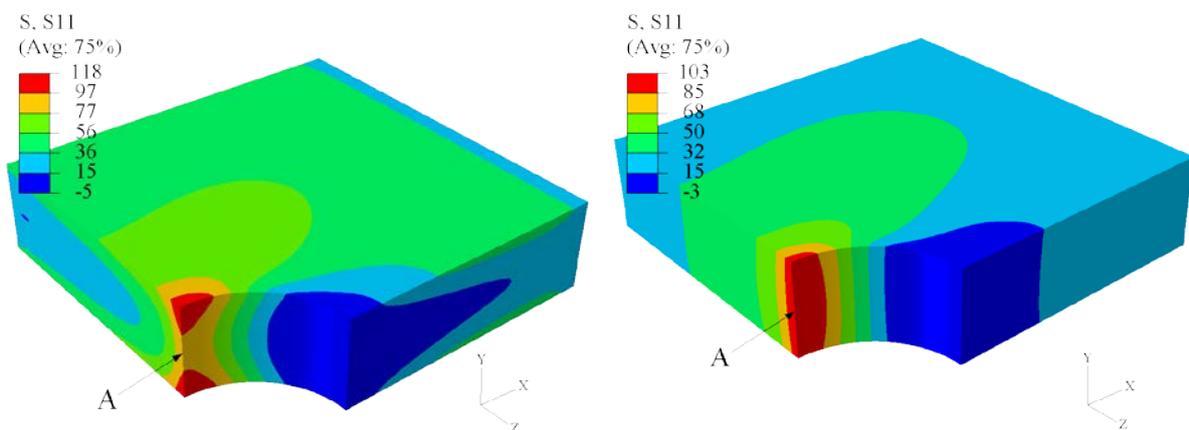
Figure 9. a. Glass sample with residual stress under 20 MPa (X direction) tensile load (left), b. Glass sample without residual stress under 20 MPa (X direction) tensile load (right)
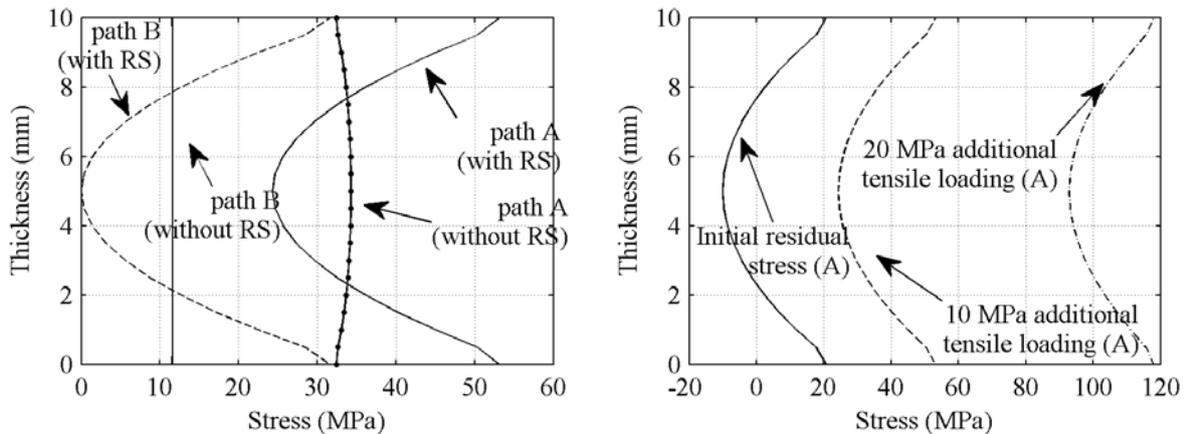
Figure 10. a. Stress profile at 10 MPa tensile loading (left) b. Stress profile along path A for different load steps (right)

Conclusions

This paper presents a hybrid validated experimental/numerical-modelling tool to characterise the residual stress present in commercially available glass. The study shows that the hybrid modelling approach works well to model the residual stresses. It has been shown that, by applying contour-method based finite element models, the full residual stress distribution that satisfies overall equilibrium, compatibility and boundary conditions can be accurately determined. An inverse eigenstrain analysis has been developed to estimate the eigenstrain distribution of glass. The eigenstrain distribution depends only on the glass manufacturing process, therefore, once the knowledge of the eigenstrain depth profile in a given glass type is available, structural response of practical glass elements under service loads can be determined in a computationally efficient manner.

Acknowledgements

References

ABAQUS (6.9-3) ABAQUS/Standard. 6.9-3 ed., Simulia

Aben, H., Anton, J. and Errapart, A. (2008) Modern photoelasticity or residual stress measurement in glass, *Strain* **44**, 40-48.

Achintha, M. and Nowell, D. (2011) Eigenstrain modelling of residual stresses generated by laser shock peening, *Journal of Materials Processing Technology* **211**, 1091-1101.

Alicona (2.1.5) Alicona Infinite Focus 2.1.5 IFM, Alicona Imaging Gmbh.

Geandier, G., Denis, S. and Mocellin, A. (2003) Float glass fracture toughness determination by Hertzian contact: experiments and analysis, *Journal of non-crystalline solids* **318**, 284-295.

Hosseinzadeh, F. and Bouchard, P. (2011) Application of the contour method to validate residual stress predictions, *2nd International Conference on Advances in Nuclear Materials,* Mumbai, India.

IStructE (1999) *Structural use of glass in buildings*, 1st ed, Institution of Structural Engineers, London.

Johnson, G. (2008) *Residual stress measurements using contour method*, PhD, University of Manchester.

Korsunsky, A.M. (2009) Eigenstrain analysis of residual strains and stresses, *Journal of Strain Analysis for Engineering Design* **44**, 29-43.

Mason, J.C. and Handscomb, D.C. (2003) *Chebyshev polynomials,* Chapman Hall/CRC, Boca Raton, FL

Pagliaro, P. (2008) *Mapping multiple residual stress components using the contour method and superposition*, PhD, Universita Degli Studi di Palremo.

Prime, M.B (2001) Cross-sectional mapping of residual stresses by measuring the surface contour after a cut, *Journal of Engineering Materials and Technology* **123**, 162-168.

Whithers, P.J. (2007) residual stress and its role in failure, *Reports on progress in physics* **70**, 2211-2264.

Numerical Simulation on Cavitation in a Vane Pump with Moving Mesh

*Qunfeng Zhang[1,2], X.Y. Xu[3]

[1]School of Civil Engineering, Beijing Jiaotong University, Beijing, 100044, China
[2]Department of Chemical Engineering, Imperial College London, South Kensington Campus, London SW7 2AZ, United Kingdom
[3]Department of Chemical Engineering, Imperial College London, South Kensington Campus, London SW7 2AZ, United Kingdom

*Corresponding author: zhangqunfeng@263.net

Abstract

Cavitation can occur in vane pumps when the local fluid static pressure falls below the vapor pressure at the operating conditions. Cavitation can reduce the flow rate and erode the internal surface of a vane pump. Because direct measurement is difficult to implement, predictive tools based on Computational Fluid Dynamic (CFD) has become a valid alternative approach to determine the position and size of cavitation in a vane pump. In this study, a purpose-built code has been developed for use with Star-CD to capture the gap shape between the vane and stator with a moving mesh. The internal flow of a vane pump has been simulated with the volume of fluid method, together with Rayleigh cavitation model and RANS-based Mender's SST $k - \omega$ turbulence model. The simulation results demonstrate the initiation, development and collapse of cavitation during the rotation of vane, which leads to accumulation or release of the mass of oil, resulting in a reduced or increased area of cavitation respectively. The transient flow rate at the inlet is stable while the outlet flow rate fluctuates significantly. The radiuses of vane tip and gap size are found to influence the size of cavitation and flow rate of the vane pump. The computational model presented here can provide the basis for optimal design of vane pumps.

Keywords: Vane pump, Numerical simulation, Moving mesh, Volume of fluid method, Cavitation model, Turbulence model

Introduction

The physical phenomenon of cavitation can occur in vane pumps when the local fluid static pressure falls below the vapor pressure at the operating conditions. The vane pump operating performance is greatly affected by the cavitation. Usually, the cavitation can reduce the working life, efficiency of vane pump and can generate vibration, noise and solid surface erosion in vane pump. Vane pump cavitation are mainly blade surface cavitation and clearance cavitation. The blade surface cavitation is generated at the position with low static pressure induced by the secondary flow near blade surface while the clearance caviation is generated for the local high speed leakage velocity near the clearance between tip of vane and stator wall. Researchers [Shcherbin and Smolyanskii (1994); Cho and Han (1998);Wu (2005); Antonio and Rosario (2005), L.Wang and Quan (2006)] have utilized theoretical analysis and empirical formulas to investigate the effect factors like profile curve of the stator, flow rate fluctuation of pump, and indicated the important effect factors of the performance and service life of vane pump. In the past, theoretical analysis and empirical formulas have been mainly relied on during the process of pump design, however it is not sufficient. It is important to follow a combined strategy between simulation, experiment and analysis to design a vane pump with the characteristics of large flow rate, high pressure and low noise. Solving the three

dimensional Navier-Stokes equations based on finite volume method plays an important role in predicting the performance of the vane pump.

The key factors of accurately simulating the vane pump internal flow are grid deformation and reconstruction as the chamber volume changes. The selection of cavitation model and turbulence model also plays an important role. The common CFD softwares utilized in the simulation of vane pump flow include FLUENT, Pumplinx and Starccm+, etc. But simplifications were made such as simplifying the geometry and fitting the profile curve of vane tip to the profile curve of stator. Sometimes artificially enlarged leakage gap was performed in order to avoid numerical convergence problems. All these modifications led to less accurate simulations of the leakage and cavitation phenomena of the vane pump.

In this paper, a purpose-built code for the computational fluid dynamics software Star-CD has been completed. This program can generate the grids in the cases of different vanes number, different profile curves of stator or different gap sizes in a fast way. It is especially capable of capturing the gap shape between the vane and the stator. In the present simulation, this feature was used combined with the Menter SST k-ω turbulence model [Menter (1994)], the Rayleigh cavitation model [Ahuja et al (2001); Bakirand Gerber (2004)] and VOF [Hirt and Nichols (1981); Ubbink and Issa (1999)] method. The flow characteristics together with the generation, development and collapse of cavitation were analyzed. Influences of the radius of vane tip on the pump performances were investigated. The comparisons of the computed results are consistent with the experimental data, indicating the appropriate selections of numerical method and models.

Governing Equations and Cavitation Model

To simulate the generation, development and collapse of cavitation in a vane pump, the three dimensional averaged Navier-Stokes equations combined with the VOF method and Rayleigh cavitation model were solved. The governing volume-fraction equations for the vapor and liquid lubricant oil are as follows:

$$\frac{\partial \rho_v \beta}{\partial t} + \frac{\partial \rho_v (u_j - v_{gj})\beta}{\partial x_j} = \rho_v S_\beta \tag{1}$$

$$\frac{\partial \rho_l \alpha}{\partial t} + \frac{\partial \rho_l (u_j - v_{gj})\alpha}{\partial x_j} = -\rho_v S_\beta \tag{2}$$

where α and β are volume fractions of liquid and vapor lubricant oil respectively. The sum of the two volume fractions must satisfy: $\alpha + \beta = 1$. The source term S_β accounts for mass exchange between the liquid and vapor phases during cavitation generating. ρ_l and ρ_v are densities of liquid and vapor of oil respectively. It is S_β. is calculated from the Rayleigh cavitation model.

The governing momentum equation for the vapor-liquid mixture is:

$$\frac{\partial \rho u_i}{\partial t} + \frac{\partial \rho (u_j - v_{gj})u_i}{\partial x_j} = -\frac{\partial p}{\partial x_i} + \rho g_i + f + \frac{\partial \tau_{ij}}{\partial x_j} \tag{3}$$

where density ρ is defined through the volume fractions as $\rho = \rho_l \alpha + \rho_g \beta$. f is the source term of momentum transport at the interface between the liquid and vapor phase.

$$\tau_{ij} = 2\mu S_{ij} - \frac{2}{3}\mu \frac{\partial u_k}{\partial x_k}\delta_{ij} - \rho \overline{u_i' u_j'} \qquad S_{ij} = \frac{1}{2}(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}) \tag{4}$$

$$\mu = \alpha \mu_l + \beta \mu_v \tag{5}$$

τ_{ij} is the deformation tensor. δ_{ij} is the Kronecker operator. S_{ij} is the deformation rate tensor. And the dynamic viscosity μ is defined through the volume fractions.

Due to the grid deformation and reconstruction during simulation, an additional equation has to be satisfied simultaneously with the other conservation equations. The space conservation law (SCL) [Demirdzic and Peric (1998)] is used to avoid generating artificial mass sources during numerical solution procedure. The conservation equation is described by:

$$\frac{d}{dt}\int_{V_P} dV - \sum_{k=1}^{n_f} \int_{S_k} v_{gj} \cdot dS_j = 0$$

(6)

where v_{gj} is the grid velocity, n_f the number of grid cells on the surface of the control volume, and the surface vector dS_j.

Turbulence Model

The eddy viscosity model is used for the Reynolds stress term to close the equation set.

$$-\rho\overline{u_i' u_j'} = \tau_{tij} = 2\mu_t S_{ij} - \frac{2}{3}(\mu_t \frac{\partial u_k}{\partial x_k} + \rho k)\delta_{ij}$$

(7)

where k is the turbulent kinetic energy and μ_t is the turbulent viscosity.

The shear stress transport (SST) k-ω turbulence model is developed by Menter, who used the SST k-ω model often merit it for its good behavior in adverse pressure gradients and separating flow. The SST model can be switched to a k-ε turbulence model in the free-stream and thereby avoids the common k-ω problem that the model is too sensitive to the inlet free-stream turbulence properties. The turbulence kinetic energy k and the specific dissipation rate ω are obtained from the following transport equations:

$$\frac{\partial}{\partial_t}(\rho k) + \frac{\partial}{\partial x_j}[\rho(u_j - v_{gj})k] = \mu_t P - \rho\beta^* \omega k + \frac{\partial}{\partial x_j}\left[(\mu + \sigma_k \mu_t)\frac{\partial k}{\partial x_j}\right]$$

(6)

$$\frac{\partial}{\partial_t}(\rho\omega) + \frac{\partial}{\partial x_j}[\rho(u_j - v_{gj})\omega] = \gamma\rho P - \beta\rho\omega^2 +$$

(7)

$$\frac{\partial}{\partial x_j}\left[(\mu + \sigma_\omega \mu_t)\frac{\partial \omega}{\partial x_j}\right] + (1 - F_1)2\rho\sigma_\omega \frac{1}{\omega}\frac{\partial k}{\partial x_j}\frac{\partial \omega}{\partial x_j}$$

$$\mu_t = \rho\frac{k}{\omega} \quad P = S_{ij}\frac{\partial u_i}{\partial x_j}$$

(8)

where $\sigma_k, \beta^*, \gamma, \beta, F_1$ are the empirical constants obtained from theoretical analysis combined with experimental data.

Grid Generation and Numerical method

Figure 1 is the section structure of the vane pump. The regions without vane rotating have been generated hexahedral grids using the mesh generator software ICEM-CFD while the region with vane rotating has been generated hexahedral grids with a purpose-built code developed by author. The purpose-built code can control the key nodes on the profile curve of the stator, and generate the moving grids by an algebraic grid generation method. So the gaps shape between the vane and stator can be captured well. The total amount of grid cells is about 560,000. Presented in Figure 2 is the grid of the rotating region (chamber), deforming with the rotation of the shaft. The thin layer grids on top surface and bottom surface are the gaps in axial direction. Figure 3 shows the good orthogonal grids near the gap with different radius (R=17mm, 34mm and 51mm) of vane tip. In the

process of calculations, the grids deform at each time step. Grids were generated in the chamber cells at current time step and the three-dimensional coordinates of the nodes were output to a file. Then a commercial finite-volume-based CFD software Star-CD reads coordinate data from the file by subroutine and complete the iteration. The spatial discretization was achieved via a second-order difference scheme MARS (Monotone Advection and Reconstruction Scheme), which based on the TVD scheme. The SIMPLE algorithm is used for the pressure-velocity coupling for the unsteady calculations. In order to improve the convergence rate and stability, a solution based on single-phase without cavitation model is first computed. Then initialized from this non-cavitating solution, the two-phase mixture solutions with activating cavitation model were calculated.
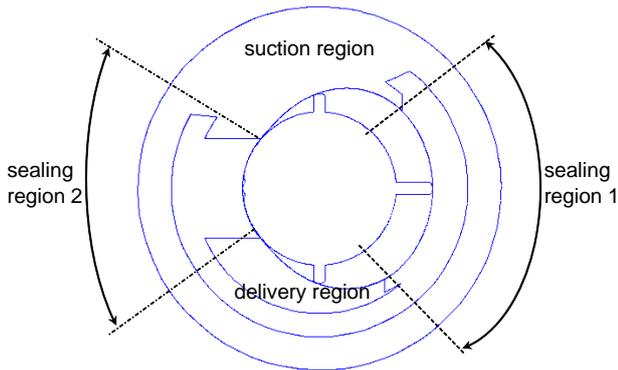
Figure 1 Section structure of vane pump
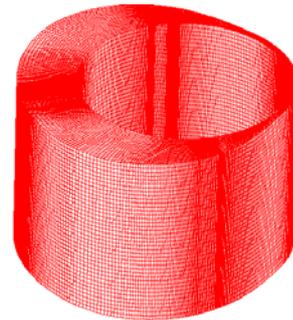
Figure 2 Grid of rotating region

| R=17mm | R=34mm | R=51mm |
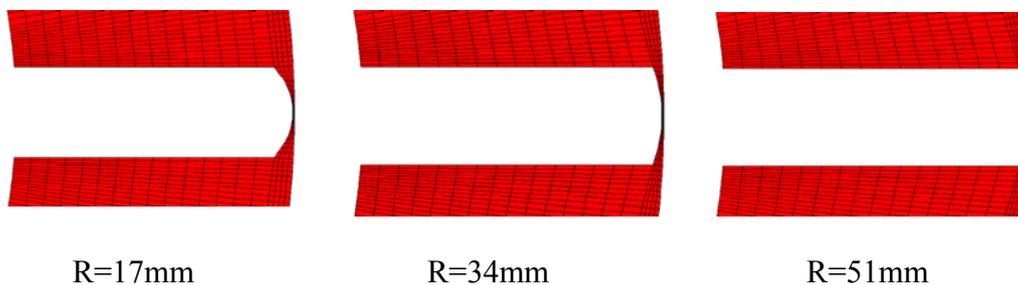
Figure 3 Grid near the gap with different radius of vane tip

Boundary Condition and Computational Cases

Often the conditions at the boundaries should represent exactly the physical conditions. For the vane pump, inlet and outlet pressure of the pump have fluctuations in reality, and setting variable pressures at inlet and outlet boundaries which needs to be measured by tests is more accurate in computation. But it is difficult to match the transient pressures with the corresponding rotor angle even if the variable pressures are obtained. So a constant values of pressure were specified at the inlet and outlet of the pump. All wall surfaces were set in non-slip and non-permeable conditions. The grid-interfaces between vanes and stator were defined as the sliding boundary condition.

Firstly, the original pump model with 4 vanes was simulated under rotation speed of 3268 RPM. For the original model, the gap between the vane tip and stator wall is 0.05mm, the radius of vane tip is 17mm, and the absolute pressures at the inlet and outlet boundary were given as 47KPa and 501KPa respectively. After comparing the calculated results and experimental results of the case, the cases with different radius of vane tip (R=34mm and 51 mm) were simulated. The influences of different radius of vane tip to performances of vane pump were analysized.

Results and Discussion

Figure 4 presents the cavitation distributions on a section of the vane pump under different rotator angle. A volume fraction 1 represents the vapor of oil while zero stands for the liquid phase and the values between 0 and 1 represent transition region of liquid oil and vapor of oil. The cavitation region in the duct of inlet is relative stable during the vane rotating while the cavitation region along the rotor is unstable, which is generated before the sealing region is closed, developed in the sealing region 1 and collapsed during the sealing region being opened. The generation, development and collapse of cavitation lead to accumulation and release of the mass of oil. The instantaneous flow rates at inlet and outlet versus rotor angle are showed in Figure 5. Corresponding to the variation of cavitation region, the instantaneous flow rates at inlet is stable while the one at outlet fluctuates significantly. For the pressure in the delivery region is higher than the value in sealing region 1, when the vane rotates from sealing region 1 to the delivery region, some lubricant oil in delivery region flows back to sealing region 1, resulting in the phenomenon of rapid decrease and subsequent recovery of flow rate at outlet. The averaged values of flow rate at inlet and outlet over a period are 36.53L/min and 37.66 L/min. When the cavitation model is involved, numerical errors are introduced in the process of dealing with the vapor/liquid interface. As a consequence, there exists the difference between inlet and outlet values of average flow rate but the mass conservation law is still satisfied in general. The averaged values of flow rate at inlet and outlet is close to flow rate obtained from experimental test, which is 38.1 L/min.

The caviatation distributions of pump with different radius of vane tip on a section depicted in
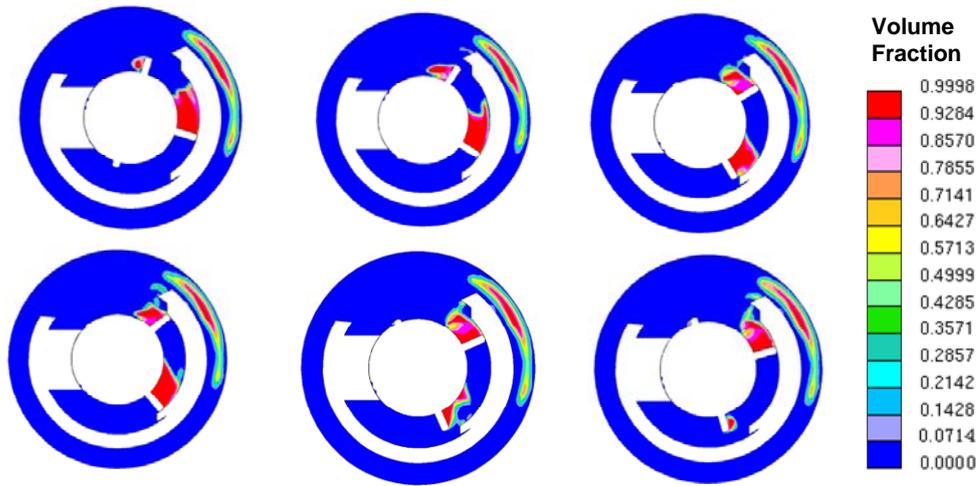
Figure 4 Cavitation distributions on a section of the vane pump

Figure 6 are presented in Figure 7. With the increase of radius of vane tip from R=17mm to R=51mm, the cavitation areas are expanded a little when the sealing region is closed. At the same rotor rotating angle, the leakage velocity from delivery region to sealing region is showed in Figure 9. Although the minimal size of the gap is the same, which is 0.05mm, the area of inlet and outlet of the gap is reduced and the
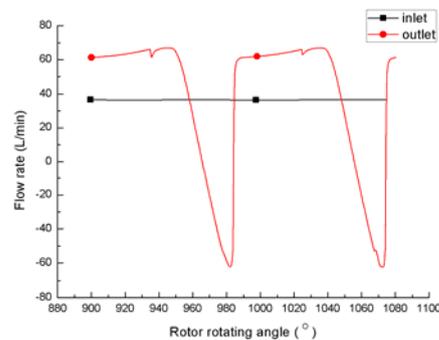
Figure 5 Instantaneous flow rates at inlet and outlet
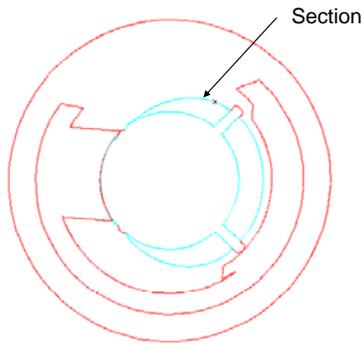
Figure 6 Section position
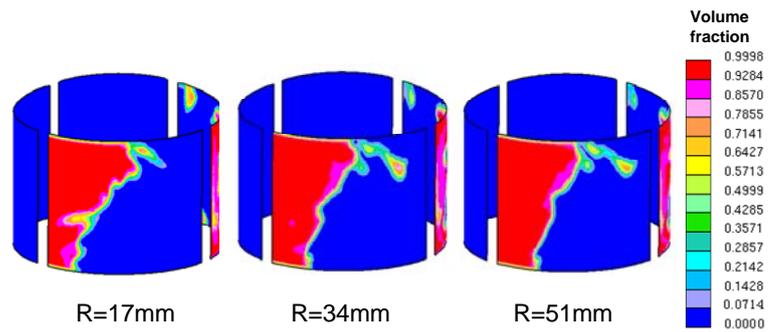
Figure 7 Cavitation distributions with different radius of vane tip

pressure loss is enhanced with increase of radius of vane tip. So the leakage velocity is somewhat larger for the case R=17mm.

Corresponding to the differences of cavition distributions and velocity magnitudes discussed above, Figure 10 shows the instantaneous flow rate during the sealing region being closed is a little larger with R=51mm while the back flow rate during the sealing region is being opened is increase slightly. The averaged flow rate of outlet is given in the Table 1 below.

Table 1 Averaged flow rate at outlet

Radius of vane tip (mm)	17	34	51
Flow rate (L/min)	37.66	36.5	36.2
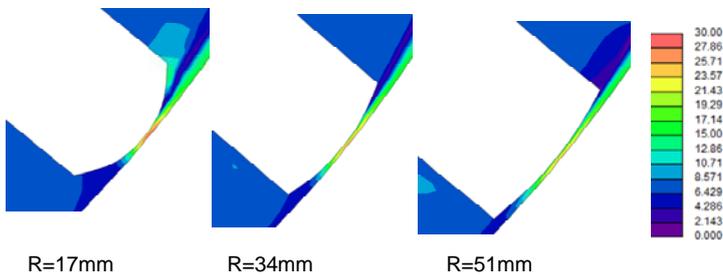
Figure 9 Leakage velocity distributions from delivery region to sealing region with different radius of vane tip
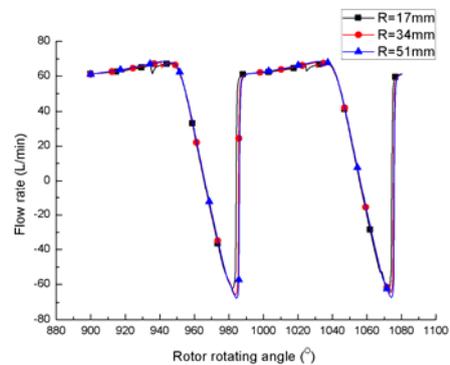
Figure 10 Instantaneous flow rates at outlet with different radius of vane tip

Conclusions

Cavitation is generated, developed and collapsed during the rotation of vane, which leads to accumulation and release of the mass of oil, corresponding to a reduced and increased area of cavitation respectively, and the transient flow rate of inlet is stable while the one at outlet fluctuates significantly. The radius of vane tip influences the cavitation area and leakage velocity during the sealing region being closed and opened; the averaged flow rate at outlet of pump over a period is a little larger in the pump with small radius of vane tip.

References

Shcherbin, V. D. and Smolyanskii, B. G. (1994) Optimization of geometry of working chamber of sliding-vane pump with circular stator, *Chemical and Petroleum Engineering,* **31,154-157.**

Cho, M. R. and Han, D. C. (1998) Vane tip detachment in a positive displacement vane pump, *KSME International Journal,* **12,881-887.**

Wu, Y.L. (2005) Fluid machinery engineering. Beijing: China Environmental Science Press.

Antonio, G. and Rosario, L. (2005) Cam shape and theoretical flow rate in balanced vane pumps, *Mechanism and Machine Theory,* **40,353-369.**

Wang, L. and Quan, L. (2006) Transient flux analysis and calculation on single-acting vane pump. *Journal of Vibration, Measurement & Diagnosis*, **26,188-191.**

Menter, F. R. (1994) Two-equation eddy-viscosity turbulence models for engineering applications, *AIAA Journal,* **32, 1598-1605.**

Ahuja,V., Hosangadi, A. and Arunajatesan, S. (2001) Simulation of cavitating flows using hybrid unstructured meshes, *Journal of Fluids Engineering,* 123,331–340.

Bakir, F., Rey, R. and Gerber, A. G. (2004) Numerical and experimental investigations of the cavitating behavior of an inducer, *International Journal of Rotating Machinery*, **10,15–25.**

Hirt, C. W. and Nichols, B. D. (1981) Volume of fluid （VOF）method for the dynamics of free boundary, *Journal of Computational Physics*, 39,201-225

Ubbink, O. And Issa, R. I.(1999) A method for capturing sharp fluid interfaces on arbitrary meshes, *Journal of Computational Physics,* 153,26-50

Demirdzic, I. and Peric, M. (1998) Space conservation law in finite volume calculations of fluid flow, *International Journal for Numerical Methods in Fluids,* 8,1037-1050.

Long-term analysis of crown-pinned concrete-filled steel tubular arches

*K. Luo, Y.L. Pi, W. Gao and M.A. Bradford

The School of Civil and Environmental Engineering, The University of New South Wales,
Sydney, NSW 2052, Australia.

*Corresponding author: kai.luo@student.unsw.edu.au

Abstract

Concrete-filled steel tubular (CFST) arches with crown-pins are widely used in engineering practice. Investigation on the long-term features of CFST arches is essential to the stability and serviceability of CFST arches. This paper presents a long-term analysis of crown-pinned CFST arches subjected to sustained central concentrated load, accounting for the coupling effects of creep and shrinkage of its concrete core and temperature change. The algebraically tractable age-adjusted effective modulus method is used to model the creep behaviour of the concrete core and the analytical solutions of time-dependent structural behaviour are obtained. It is found that the coupling effects cause significant increase of deformation and internal force, which may affect the local strength reserve and the routine service of the CFST arches.

Keywords: Crown-pinned arch, CFST, Creep, Shrinkage, Temperature

Introduction

Among composite structures, concrete-filled steel tubular (CFST) cross-section is widely employed in engineering practice. As can be seen from the cross section in Fig. 1, the composite section is composed of two components including the steel tube and the concrete core, which facilitates the resistance of the structure to the external force.
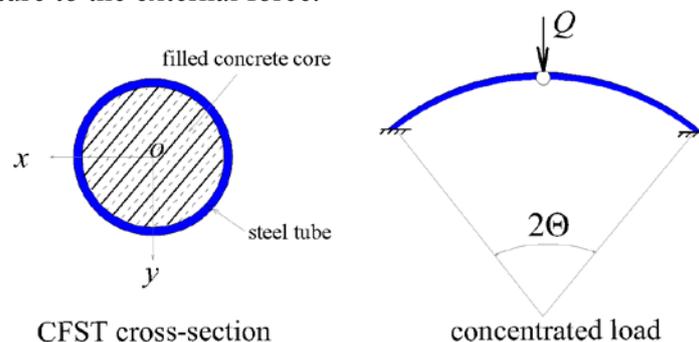
Figure 1. CFST crown-pinned arch subjected to a central concentrated load

As time goes on, the creep deformation increases when a CFST arch undergoes a sustained load, while the shrinkage strain emerges all the time even when there is no applied load. The visco-elastic effects of the concrete core exert significant effects on the structural behavior of CFST arches. Studies of the buckling behaviour have been focused on pin-ended and fixed arches [Pi et al. (2002); Bradford et al. (2002)]. In these studies, the arch is assumed to be continuous without any pins between its ends. It is known that in many cases, arches are built by joining two separate curvilinear segments together at the crown, thereby reducing the arch size to meet transport requirements and to create a pin at the arch crown. Because of the crown-pin, the structural responses and buckling behaviour of the arch are different from those of arches without the crown-

pin. However, investigations on crown-pinned shallow circular arches do not appear to be reported in the open literature. This paper, therefore, presents an analytical analysis to investigate the non-linear long-term behaviour of crown-pinned CFST shallow circular arches subjected to a central concentrated load.

Long-term effects of concrete

Due to the different Young's moduli of materials and the effects of creep and shrinkage in the concrete core, the stress σ_s in the steel tube and the stress σ_c in the concrete core are quite different with each other.

The stress in the steel tube can be expressed as

$$\sigma_s = E_s \varepsilon \tag{1}$$

in which E_s is Young's modulus of steel, while the stress in the concrete core can be obtained based on the model proposed by ACI Committee-209 and the Australian design code for concrete structures AS3600 as

$$\sigma_c = E_{ec} \left(\varepsilon + \varepsilon_{sh} + \varepsilon_T \right) \tag{2}$$

in which the age-adjusted effective modulus of concrete E_{ec} is given by [Gilbert (2010), Bazant (1972)]

$$E_{ec} = \frac{E_c}{1 + \chi(t,t_0)\phi(t,t_0)} \tag{3}$$

in which $\phi(t,t_0)$ is the creep coefficient and $\chi(t,t_0)$ is the aging coefficient and they are given by

$$\phi(t,t_0) = \left[\frac{(t-t_0)^{0.6}}{10 + (t-t_0)^{0.6}} \right] \phi_u \text{ and } \chi(t,t_0) = 1 - \frac{(1-\chi^*)(t-t_0)}{20 + (t-t_0)} \tag{4}$$

respectively, where ϕ_u is the final creep coefficient (the value when $t \to \infty$) and given by

$$\phi_u = 1.25 t_0^{-0.118} \phi_{\infty,7} \text{ and } \chi^* = \frac{k_1 t_0}{k_2 + t_0} \tag{5}$$

With $k_1 = 0.78 + 0.4e^{-1.33\phi_{\infty,7}}$ and $k_2 = 0.16 + 0.8e^{-1.33\phi_{\infty,7}}$.

ε_{sh} is the shrinkage strain of the concrete and can be expressed by ACI Committee-209 and Australian design code for concrete structures AS3600 as

$$\varepsilon_{sh} = \left(\frac{t}{t+d} \right) \varepsilon_{sh}^* \tag{6}$$

in which t is the time in days, because the egress of the moisture in the concrete core is prevented by the steel tube, $d = 35$ days for moist curing can be used for the concrete core of CFST members, and ε_{sh}^* is the final shrinkage strain (the value when $t \to \infty$). Although experimental studies of the shrinkage strain ε_{sh} and creep coefficient $\phi(t,t_0)$ of CFST columns have been reported by several researchers [Zhong (1994); Terrey (1994); Uy (2001); Han (2004)] and the empirical values for the final shrinkage strain ε_{sh}^* and the final creep coefficient ϕ_u of CFST columns were proposed, these values for the CFST member cannot be used directly for the time-dependent analysis of CFST members. Instead, they can be derived from the creep and shrinkage test results of CFST columns. From the creep and shrinkage test results [Uy (2001)], the shrinkage strain and creep coefficient at time $t = 140$ days can be derived. Based on the derived data, the empirical value of the final

shrinkage strain $\varepsilon_{sh}^* = 340 \times 10^{-6}$ (the value when $t \to \infty$) and the creep coefficient $\phi_{\infty,7} = 2.5$ are used in the following investigation.

ε_T is the thermal strain resulted by temperature change in concrete core and can be expressed as

$$\varepsilon_T = \alpha_c (T - T_0) \tag{7}$$

where α_c is the coefficient of thermal expansion, T_0 is the reference temperature and $\Delta T = T - T_0$ indicates the temperature change in the ambient environment.

Non-linear in-plane equilibrium

Assumptions adopted in this investigation are: 1. The Euler-Bernoulli hypothesis is applied so that the plain remains plane when it rotates about the neutral axis. 2. The dimension of the cross-section is much smaller than the length and radius of the arch to ensure sufficient slenderness. Because the arch and load system is symmetric, equilibrium of a half arch ($0 \le \theta \le \Theta$ and Θ is half of the included angle of the arch) is considered (Fig. 2). Based on the assumptions, differential equations of equilibrium for a crown-pinned circular arch can be derived from the principle of virtual work as

$$N' = 0 \quad \text{and} \quad \frac{\tilde{v}^{iv}}{\mu^2} + \tilde{v}'' = -1 \tag{8}$$

for the arches that are subjected to a central concentrated load Q [Pi et al. (2008)], where $(\ )' \equiv d(\ )/d\theta$, θ denotes the angular coordinates, $\tilde{v} = v/R$ and $\tilde{w} = w/R$, v and w are the radial and axial displacements with R being the radius of the arch, μ is the dimensionless axial force parameter defined by $\mu = NR^2/EI$ with E being Young's modulus and I being the second moment of area of the cross-section, and the axial compressive force N is defined by

$$N = -\left(A_s E_s + A_c E_{ec}\right)\left[\tilde{w}' - \tilde{v} + \frac{1}{2}(\tilde{v}')^2\right] - A_c E_{ec}\left(\varepsilon_{sh} + \varepsilon_T\right) \tag{9}$$
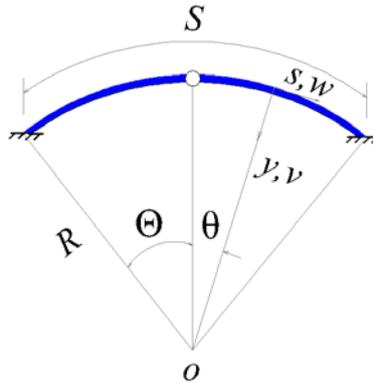
Figure 2. Arch geometry

The boundary conditions can also be derived as

$$\tilde{v}'' = 0 \quad \text{at} \quad \theta = 0 \text{ and } \tilde{w} = 0 \quad \tilde{v} = 0 \quad \text{at} \quad \theta = \Theta \tag{10}$$

Solving the equations given by Eqs. (8) and (9) simultaneously will lead to the solution of the radial displacement $\tilde{v}$ for crown-pinned arches subjected to a central concentrated load, and leads to the non-linear equilibrium equation between the internal force parameter μ and external force Q as

$$A_1 \tilde{Q}^2 + A_2 \tilde{Q} + A_3 = 0 \qquad (11)$$

where $\tilde{Q}$ is the dimensionless load defined by $\tilde{Q} = R^2 \Theta Q / 2EI$, and the expressions for coefficients for A_1, A_2, A_3 can be obtained accordingly.

Long-term analysis

The long-term deformation and internal force of a crown-pinned CFST shallow arch are illustrated in the following figures including three scenarios accounting for different long-term effects. In the example shown below, the crown-pinned arch is fixed at two ends. The geometry is given as follows: the span is $L = 15$ m, the rise-to-span ratio is $f / L = 1/6$ and the thermal expansion coefficient of concrete is assumed as $\alpha_c = 340 \times 10^{-6}$ with the temperature change being $\Delta T = -20^{\circ}C$. The dimension of the cross-section is assumed that the outer radius of the steel tube is $r_0 = 250$ mm and the inner radius of the concrete core is $r_i = 240$ mm.
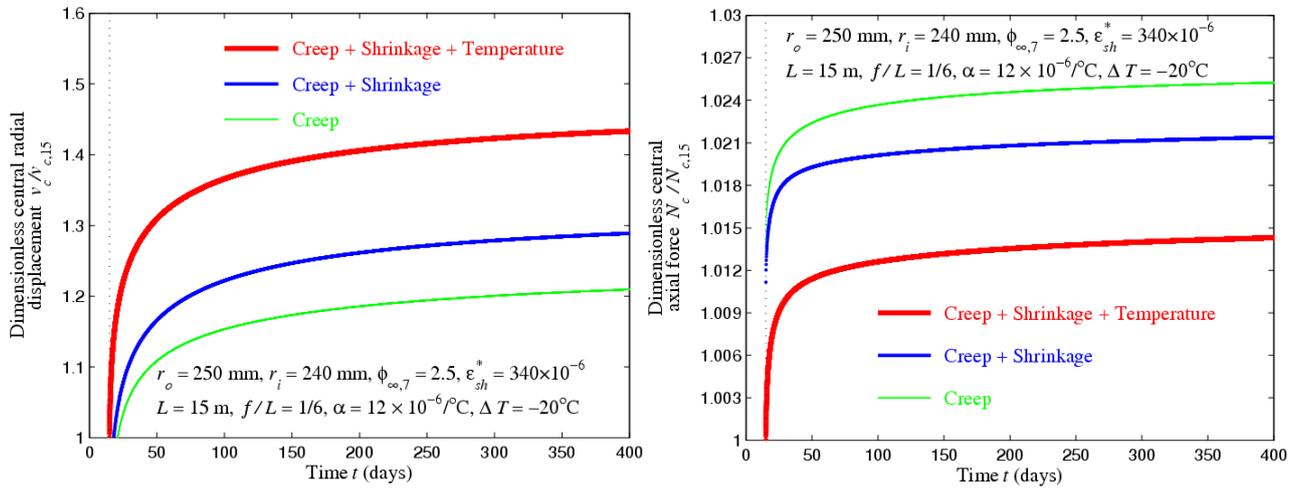
Figure 3. Long-term deformation and internal force

In Fig. 3, the central radial displacement of the arch is shown with respect to the increased time. $v_{c,15}$ and $N_{c,15}$ represents the central radial displacement and central axial force at time $t = 15$ days, due to the coupling effects of creep, shrinkage and temperature change. Among these three curves, one situation that all of these non-mechanical factors are combined together will lead to a drastic increase of displacement in the long-term. So the coupling effects of creep, shrinkage and temperature change influence the structural behaviour most significantly. In this case, the large deformation may exceed the maximum limit of amount referred in design practice. The routine serviceability cannot be assured in the long-term either. For a crown-pinned CFST shallow arch, the long-term internal axial force will also increase but only slightly in the life time. It can be seen that the change scope is the greatest when the coupling effects of creep, shrinkage and temperature change is accounted for, although the axial forces are initially different at the first loading time $t = 15$ days. The temperature change is assumed to be decreased in this example so the displacement is increased in radial direction, while the axial force is less than that of the other situations in which the temperature change is not considered.
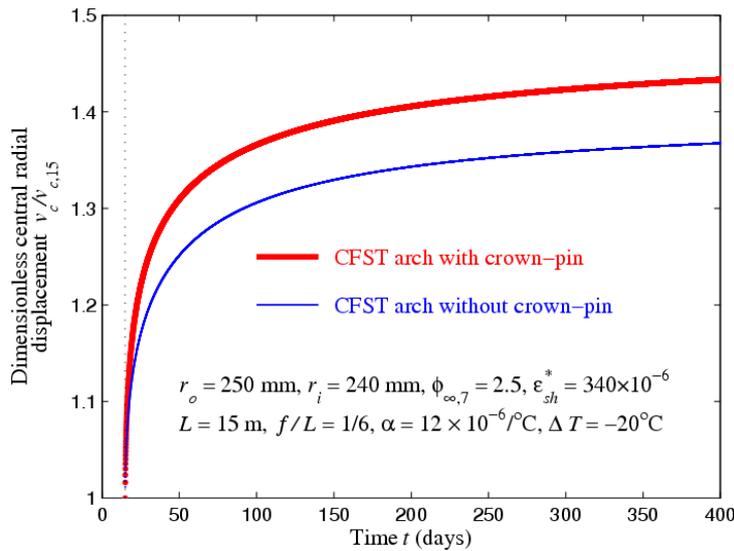
Figure 5. Long-term displacement

Following the same geometry and the other parameters of the arch above, the effects of the crown-pin can be also investigated by comparing the structural behaviour of two CFST arches: one of them is crown-pinned and the other one is continuous along the arch body. The boundary conditions are different when solving the differential equations of equilibrium. As can be found from Fig. 5, the comparison reveals an obvious gap of increasing rates between the crown-pinned CFST arch and the continuous CFST arch. The long-term central displacement of the CFST arch with the crown-pin is greater than that of the CFST arch without the crown-pin. The crown-pinned CFST arch will deflect within a broader range in the long-term and may become unsafe and cannot remain functional when it cannot withstand large deformation. The crown-pin acts as a degrader to the resisting ability of CFST arches to the external load in the long-term.

Conclusions

This paper studied the non-linear long-term behaviour of crown-pinned shallow CFST arches subjected to central concentrated load. It was found that the coupling effects of the creep, shrinkage and temperature change in the concrete core influence the long-term response significantly and may affect the local strength reserve and the routine service of the CFST arches. Hence, these effects need to be accounted for in design practice of CFST arches. It was also shown that the crown-pin also plays an important role in the long-term response. As time increases, the change scopes of deformations and internal forces were found to be different from that of the continuous CFST arches. In design practice, all these factors should be considered to ensure a safe structure that can satisfy the strength requirement and also the normal serviceability.

References

Pi, Y. L., Bradford, M. A. and Uy, B. (2002), In-plane stability of arches. *International Journal of Solids and Structures*, 39(1), pp. 105–125.

Bradford, M. A., Uy, B. and Pi, Y. L. (2002), In-plane stability of arches under a central concentrated load. *Journal of Engineering Mechanics*, 128(7), pp. 710–719.

ACI Committee 209, Prediction of Creep, Shrinkage and Temperature Effects in Concrete Structures, American Concrete Institute (ACI), Detroit, 1982.

AS3600, Australian Standard: Concrete Structures, Standard Association of Australia, Sydney, 2001.

R. I. Gilbert, Time-dependent Behaviour of Concrete Structures, in press, Taylor & Francis, London, 2010

Z. P. Bazant, Prediction of concrete creep effects using age-adjusted effective modulus method, *ACI Structural Journal*. 69 (1972) 212–217.

S. T. Zhong, Concrete-filled Steel Tubular Structures, Heilongjiang Science and Technology Press, 1994

P. J. Terrey, M. A. Bradford and R. I. Gilbert, Creep and shrinkage of concrete in concrete-filled circular steel tubes, Proceeding of 6th International Symposium on Tubular Structures, 1994, pp. 293–298.

B. Uy, Static long-term effects in short concrete-filled steel box columns under sustained loading, ACI Structural Journal, 98 (2001) 96–104.

L. H. Han, Y. F. Yang, and W. Liu, The behaviour of concrete filled steel tubular columns with rectangular section under long-term loading, Journal of Civil Engineering, 37 (2004) 12–18.

Pi, Y. L., Bradford, M. A. and Francis, T. L. (2008), Non-linear in-plane buckling of rotationally restrained shallow arches under a central concentrated load. *International Journal of Non-linear Mechanics*, 43, pp. 1–17.

Size and Scale Dependence of the Behavior of Truss Models: A Gradient Elasticity Approach

O. T. Akintayo[1, 2]

[1] Superurban Technologies Limited, P. O. Box 73463, Victoria Island, Lagos, Nigeria.
[2]Department of Civil Engineering, Division of Structural Engineering, Laboratory of Mechanics and Materials, Faculty of Engineering, Aristotle University of Thessaloniki, 54124, Greece.

femiakint@superurbangroup.com

The effect of size and scale dependence on the behavior of truss models is investigated, by introducing higher order strain gradient terms and an internal length parameter (the gradient elasticity coefficient) into the constitutive stress-strain relation of the bar element of the truss model. Instead of an algebraic equation (1D classical elasticity stress-strain relation), a differential equation governs the response of the elastic bar element and extra boundary conditions are required at the nodes. The displacement fields are obtained by deriving the associated stiffness matrix directly from the governing differential equation of the gradient elastic bar element.

Moreover by establishing a ratio between the micro-scale internal length and the macro-scale length of the bar elements the effect of changes in the truss bar element size and the microstructure internal length of the bar element material is revealed. A numerical example is presented as illustration.

Keywords: Size and Scale dependence, Gradient elastic coefficient, Gradient elastic truss bar, Gradient elasticity stiffness matrix

1. Introduction

In developing new engineering structures the analysis of their mechanical properties and behavior is often based on qualitative analysis carried out to enable a reasonable evaluation of the feasibility before embarking on any elaborate expensive research. In this process, computational analysis and simulation play a major role in developing these new structures.

Computationally, in order to study the nonlinear mechanical behavior of materials, a first choice is to use the standard Finite Element Method (FEM) with complex elements. While this method allows the computational stress analysis of a continuum with any boundary conditions and any loading, several problems and complications arise mainly due to the inclusion of a nonlinear constitutive law when updating the stiffness matrix of the finite element Argyris (1978, 1981, and 1984). However, from a geometrical standpoint the simplest finite elements are one-dimensional or a line element which is the two-node bar element. One-dimensional models can be very accurate and very cost effective in the proper applications. Hence, in search of a less complicated and efficient computational tool for testing constitutive equations, a truss model can be used for the linear and nonlinear analysis of a continuum, since the bars of the truss are the simplest possible finite elements.

In the nonlinear analysis of concrete, truss models have been used in Bazant et al. (1990) and Bazant (1997) and in Goel et al (1997) it was used for the analysis of steel structures. Akintayo et al. (1998, 2000), and Papadopoulos & Xenidis (1998, 1999) studied the response of concrete computationally using the plane truss model by considering coarse truss structures. Kiousis et al. (2010) using the model of Papadopoulos & Xenidis (1999) also studied concrete columns in compression. A random particulate model for fracture of aggregate and fiber composites was used

by Bazant et al. (1990) assuming the particles to be elastic and having only axial interaction as in a truss. In Salem (2004), a fine micro truss model was proposed for the analysis of reinforced concrete using isotropic truss members by a generalization of the strut and tie model. Nagarajan et al. (2010) studied the mesoscopic numerical analysis of reinforced concrete using a modification of the micro truss model of Salem (2004). Hence from these works it becomes apparent that the size of the truss bar element to be used could directly be related to the scale of interest and detailed analysis required in any simulation which may have an adverse effect on the result.

In the investigation of new materials and structure, the ultimate goal of the engineer is to be able to obtain qualitative and reasonably quantitative analytic results. In order to achieve this, simple but effective models need be adopted to accommodate the micro deformation mechanisms of the structure. However, with the truss model increasing the fineness of the truss bars usually increases the computational cost. Therefore the effect of the size of the truss bar element and its relation to the scale of interest become a vital consideration in the simulation and analysis process.

The interpretation of size and scale effects can be approached in different ways. Several theoretical models have been developed to interpret size effects such as the strain gradient theories of Aifantis, 1999a and 1999b, Gao et al. 1997 and Fleck and Hutchinson 1997 with the formulation of the latter in Fleck and Hutchinson 2001. Other works include models with dislocation confined in thin films (Freund 1987, Thompson 1993 and Nix 1998); Theories on discrete dislocation dynamics include the works of Zbib and Aifantis 2003 and Needleman and Van der Giessen 2003. Atkins, 1999 and Bazant 1999 presented their study on fracture mechanics theories (especially for concrete) and statistical models was initially proposed by Irwin 1964 and later by Liu and Zenner 1995 and Seifried, 2004.

The theory of gradient elasticity is a simple approach to include microstructure deformation in the analysis of a material/structure, since it becomes particularly useful for small volumes, where the internal length introduced by the gradient coefficients is comparable to the characteristic dimension of the system. Mindlin (1964) showed that by isolating a typical unit cell element from the grid of say a crystal lattice (local representative volume), the modeling of a continuum with micro deformation can be developed. Ben-Amoz (1976) by assuming a particulate composite material as consisting of the matrix and inclusion (unit cell) also showed it is possible to classify composite media by the degree of inhomogeneity with the ratio of the length of the local representation and the length of the unit cell. Hence a relation is established between the deformation within the local representative volume and the unit cell by using higher-order strain/stress gradients to represent the micro deformation within the macro structure/material. In line with this same concept, but in a different manner, the theory of gradient elasticity proposed by Aifantis (1984) included the higher order strain gradient directly into the constitutive relations and introduced an internal length parameter, which relates to the micro unit of the material. A recent review of this theory is given in Askes and Aifantis [2011].

In recent years this theory has gained more increasing interest amongst researchers and the engineering community due to its ability to provide additional information, which the classical elasticity theory is incapable of providing. The failure of the classical elasticity theory to include higher order strain gradient contributions can lead to underestimates of stresses and inadequacy in capturing any scale and size dependent behavior in small-scale structures: since classical elasticity theory possesses no characteristic length (i.e. material parameter with internal length scale), which consider the interaction between macro and micro length scales in the constitutive response and the corresponding interpretation of associated scale and size effect.

Many authors have studied the gradient theories using various computational methods. Amongst others include: In the framework of gradient plasticity Pamin and de Borst (1998) used the finite element method to simulate the crack spacing problem with a reinforced concrete model. Chang, et

al. (2002) applied higher-order strain/higher order stress gradient models derived from a discrete microstructure to fracture and related their constitutive equations to that of Aifantis (1984). By using the gradient elasticity theory of Aifantis (1984), Dessouky et al. (2003) presented a finite element model for the microstructure analysis of asphaltic materials. In a similar manner, Akarapu and Zbib (2006) also considered the analysis of plane cracks in elastic materials using the finite element methods.

Motivated by this and using the constitutive stress-strain relation of the gradient elasticity model of Aifantis [1984], the gradient truss model was first studied in Akintayo (2011) and later presented in Akintayo et al. (2012). Subsequently, a more detailed study was presented in Akintayo 2014, in which different boundary conditions were imposed at the support of the bar element and the corresponding force-displacement relations were derived for a robust application in the proposed gradient truss model. It is shown that the gradient elastic bar element is able to support strain gradient along its length such that simulation of the micro-scale deformation is included and a means of relating the macrostructure bar length to the microstructure internal length is established.

In this paper based on these findings, in a simple manner by using the gradient bar element stiffness expression and by considering the ratio between the size of the truss bar and the internal length of the material being simulated, the effect of changes in this ratio on the simulation is investigated. As an illustration, the gradient enhanced bar element is used to simulate a simple truss structure. Subsequently, by considering different bar length to internal length ratios, the response of the truss structure to scale and size dependence behavior can be examined.

Brief Review of the Classical Elasticity Bar Element and its Local Stiffness Matrix

Consider the generic truss element shown in Figure 1(a). The force and displacement components are linked by the element stiffness relations

$$\bar{\mathbf{f}} = \bar{\mathbf{K}}\bar{\mathbf{u}} \tag{1}$$

which written out in full is

$$\begin{bmatrix} \bar{f}_{xi} \\ \bar{f}_{yi} \\ \bar{f}_{xj} \\ \bar{f}_{yj} \end{bmatrix} = \begin{bmatrix} \bar{K}_{xixi} & \bar{K}_{xiyi} & \bar{K}_{xixj} & \bar{K}_{xiyj} \\ \bar{K}_{yixi} & \bar{K}_{yiyi} & \bar{K}_{yixj} & \bar{K}_{yiyj} \\ \bar{K}_{xjxi} & \bar{K}_{xjyi} & \bar{K}_{xjxj} & \bar{K}_{xjyj} \\ \bar{K}_{yjxi} & \bar{K}_{yjyi} & \bar{K}_{yjxj} & \bar{K}_{yjyj} \end{bmatrix} \begin{bmatrix} \bar{u}_{xi} \\ \bar{u}_{yi} \\ \bar{u}_{xj} \\ \bar{u}_{yj} \end{bmatrix} \tag{2}$$

There are several ways to construct the stiffness matrix $\bar{\mathbf{k}}$ in terms of the bar length l_o, modulus of elasticity E and bar cross-sectional area A. The most straightforward technique is the unit displacement method.

By viewing the truss element in Figure 1(a) as a spring in Figure 1(b), we can set the element stiffness $k_s = \bar{K}_{ijij}$, with

$$k_s = \frac{AE}{l_O} \tag{3}$$

Consequently the force-displacement equation is

$$F = \frac{AE(u_j - u_i)}{l_O} = k_s d = \frac{AE}{l_O} d \tag{4}$$

where *F* is the internal axial force and *d* the relative axial displacement, which physically denote the bar elongation.

(a)   (b)

(c)   (d)
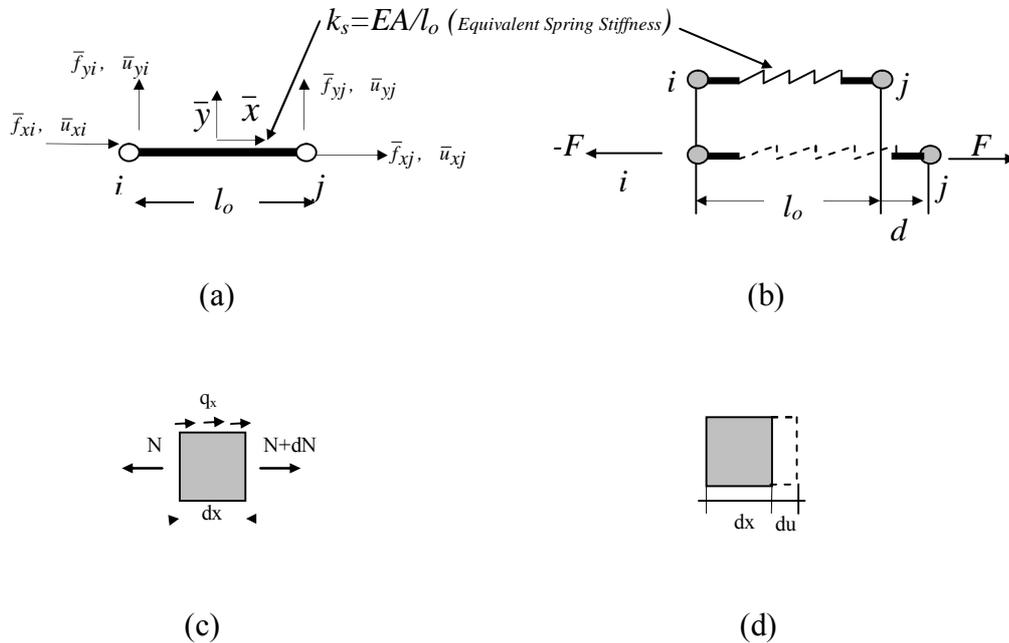
Fig.1.(a). Generic 2-node truss bar element referred to its local coordinate system $\{\bar{x},\bar{y}\}$ with nodal forces f_{ij} and displacements u_{ij}. (b) Interpretation of a Generic truss element as equivalent spring (c) Equilibrium for infinitesimally small truss element. (d) Kinematics for infinitesimally small truss element.

By assuming the displacement is of equal magnitude and direction at each node as well as within the element (i.e. constant along the bar), the strain takes the form

$$\varepsilon = \frac{du}{dx} = \frac{\Delta L}{l_o} = \frac{\left(u_j - u_i\right)}{l_o} \tag{5}$$

and on the basis of the one dimensional Hooke's law the stress-strain relation is

$$\sigma = E\varepsilon \tag{6}$$

where σ is the stress, ε is the strain and E is the Elastic or Young Modulus.

The elastic 2-node bar element of Fig. 1a is prismatic, weightless, and isotropic; the Poisson's effect is not considered and the axial load is applied at the centroid. Then equilibrium in the x-direction for the infinitesimally small length of the truss bar element, shown in Fig 1c. gives

$$q_x = -\frac{dN}{dx} \tag{7}$$

The normal constant stress σ of a one-dimensional truss is the force F applied on the truss per unit cross-sectional area;

$$\sigma = \frac{F}{A} = E\varepsilon \tag{8}$$

By considering the infinitesimal element Fig. 1c and denoting the infinitesimal elongation, by du, the relationship between the strain and the displacement is obtained as

$$\varepsilon = \frac{du}{dx} \tag{9}$$

and the governing differential equation for truss members reads:

$$q_x = -EA\frac{d^2u}{dx^2} \tag{10}$$

which in terms of the force applied per unit cross-sectional area can be rewritten as

$$N = EA\frac{du}{dx} \tag{11}$$

The element stiffness is given by Eq. 3 and since equilibrium suggests that $f_i = -f_j$, hence the force-displacement relation for the bar is given by

$$\underbrace{\begin{Bmatrix} f_i \\ f_j \end{Bmatrix}}_{f_e} = \underbrace{\frac{AE}{l_o}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}}_{K_e}\underbrace{\begin{Bmatrix} u_i \\ u_j \end{Bmatrix}}_{u_e} \tag{12}$$

where $\mathbf{f_e}$, $\mathbf{K_e}$ and $\mathbf{u_e}$ are the force, stiffness and displacement matrices respectively.
However, in plane classical elasticity structural analysis, the simple local stiffness matrix of an isolated bar, with respect to reference axes (x,y), is usually split into the elastic or material stiffness k_e and the geometric stiffness k_g and is written as

$$k_\ell = k_e + k_g$$
$$= \frac{EA}{l_o}\begin{pmatrix} c_x^2 & c_xc_y \\ c_xc_y & c_y^2 \end{pmatrix} + \frac{N}{l}\begin{pmatrix} c_y^2 & -c_xc_y \\ -c_xc_y & c_x^2 \end{pmatrix} \tag{13}$$

Where (l_o , l) are the undeformed and deformed length of the bar, and (c_x , c_y) are direction cosines of the bar.

Gradient Elastic Bar

   Consider an elastic bar length l_o shown in Fig. 2, with modulus of elasticity E, and a cross-sectional area of A, fixed at one end and subject to an axial tensile force F at the right end x = l_o. The one-dimensional gradient elasticity stress-strain constitutive equation presented in Aifantis 1984 and used in Altan & Aifantis 1997, Akintayo 2011, Akintayo et al. 2012 is given by

$$\sigma_x = E\left(\varepsilon_x - c\frac{d^2\varepsilon_x}{dx^2}\right) \tag{14}$$
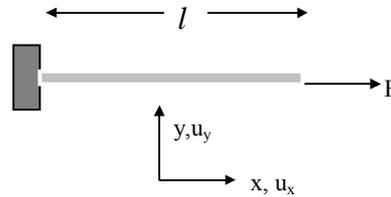
`Fig.2. A bar under axial tensile load F

where σ_x is the stress and $\varepsilon_x = du/dx$ is the strain, while c is the gradient elastic coefficient ($c \equiv \ell^2$ i.e. a microstructure-dependent internal length).

From the stress-strain relation of Eq. 14, in relation to the force applied per unit cross-sectional area of Eq. 8, the following is obtained:

$$EA\left(\frac{du}{dx} - c\frac{d^3u}{dx^3}\right) = F \tag{15}$$

and from the equilibrium condition of Eq. 11 the corresponding governing differential equation for the gradient elastic case takes the form

$$EA\left(\frac{d^2u}{dx^2} - c\frac{d^4u}{dx^4}\right) + q_x = 0 \tag{16}$$

The solution of Eq. 16 is easily obtained as

$$u = C_1 + C_2 x + C_3\left(\cosh(\frac{x}{\sqrt{c}}) + \sinh(\frac{x}{\sqrt{c}})\right) + C_4\left(\cosh(\frac{x}{\sqrt{c}}) - \sinh(\frac{x}{\sqrt{c}})\right) \tag{17}$$

where C_1, C_2, C_3 and C_4 are constants to be determined from the boundary conditions.

Boundary Conditions and Determination of Constants

In order to solve for the constants C_1, C_2, C_3 and C_4, classical and extra non-classical boundary conditions need be determined. The following classical boundary condition is used

$$u\big|_{x=l_O} = 0 \; ; \qquad\qquad EA\left(\frac{du}{dx}\right)_{x=l_O} = F \tag{18}$$

A general discussion of the extra boundary conditions for gradient elastic bar was provided in Akintayo (2011, 2012a). Hence this is not recapitulated, or expanded upon here, as it is out of the scope of the present paper. We proceed instead, with the consideration of the following simple case where zero strain is imposed at the free end and zero strain gradient at the fixed end given below

$$u\big|_{x=0}=0; \qquad EA\left(\frac{du}{dx}-c\frac{d^3u}{dx^3}\right)\bigg|_{x=l_o}=F; \qquad \frac{du}{dx}\bigg|_{x=l_o}=0; \qquad \frac{d^2u}{dx^2}\bigg|_{x=0}=0 \qquad (19)$$

Using these boundary conditions, the constants C_1-C_4 are determined as follows:

$$C_1=0; \qquad C_2=\frac{F}{EA}; \qquad C_3=-C_4=-\frac{Fe^{\frac{l_o}{\sqrt c}}\sqrt c}{\left(1+e^{\frac{2l_o}{\sqrt c}}\right)EA} \qquad (20)$$

Displacement, Strain and Axial Force

With the obtained constants the displacement and axial force within the assumed gradient elasticity framework are obtained as

$$u=\frac{Fx}{EA}-\frac{F\sqrt c}{EA}\left(\sinh\left(\frac{x}{\sqrt c}\right)\text{sech}\left(\frac{l_o}{\sqrt c}\right)\right); \quad \varepsilon^g=\frac{F}{EA}-\frac{F}{EA}\left(\cosh\left(\frac{x}{\sqrt c}\right)\text{sech}\left(\frac{l_o}{\sqrt c}\right)\right); \quad F=\frac{EAu}{x-\sqrt c\left(\sinh\left(\frac{x}{\sqrt c}\right)\text{sech}\left(\frac{l_o}{\sqrt c}\right)\right)} \qquad (21)$$

It is easily seen that the axial displacement, strain and force expressions contain the classical term and the gradient elastic contribution (the micro-scale deformation) which includes the internal length $\sqrt c$. Hence, at the free end of a bar $x = l_o$, the following relations are obtained for the displacement and axial force:

$$u=\frac{Fl_o}{EA}-\frac{F\sqrt c}{EA}\left(\tanh\left(\frac{l_o}{\sqrt c}\right)\right); \qquad \varepsilon=0; \qquad F=\frac{EAu}{l_o-\sqrt c\,\tanh\left(\frac{l_o}{\sqrt c}\right)} \qquad (22)$$

In Eq. (22) the classical displacement and strain values can be retrieved in the absence of the gradient elastic contribution.

Gradient Elastic Bar Element Stiffness

In this section, the gradient elastic bar element stiffness is derived. With this the simulation of a bar with a gradient enhanced bar element can be analysed. From Eq. 4 and Eq. 22, the corresponding gradient elastic bar stiffness is obtained as

$$k_i^g=\frac{EA}{l_o-\sqrt c\,\tanh\left(\frac{l_o}{\sqrt c}\right)} \qquad (23)$$

It is readily seen that unlike the classical case for the gradient elasticity case the bar stiffness does not only depend on the conventional geometric and material properties of the bar, but also on the introduced micro-scale parameter $\sqrt c$ of the underlying microstructure. From Eq. 23 in the absence of the gradient elastic contribution the classical bar element stiffness is retrieved. By considering the following arbitrary properties of the bar element: Area A = 1cm^2, Elastic Modulus E=100MPa; bar length l_o=10cm, the stiffness values for the classical case and the gradient elastic cases are given in Table 1 for different bar length and internal length ratios l_o /$\sqrt c$. The different bar stiffness are represented as k^c (classical elastic case) and k^g_1 - k^g_5 (gradient elastic cases) and are given in Table 1 for the changes in l_o/$\sqrt c$ from 20 to 4.

Table 1. The following properties are used arbitrarily to obtain the bar stiffness k^c (classical elastic case) and k$^g_1$ - k$^g_5$ (gradient elastic cases): A = 1cm, E=100MPa and l_o=10cm

Cases	Bar Stiffness Values for $l_o/\sqrt{c}=20$	Bar Stiffness Values for $l_o/\sqrt{c}=10$	Bar Stiffness Values for $l_o/\sqrt{c}=6.7$	Bar Stiffness Values for $l_o/\sqrt{c}=5$	Bar Stiffness Values for $l_o/\sqrt{c}=4$
k^c	10	10	10	10	10
k^g_{1a}	10.526	11.111	11.765	12.4997	17.6259

This numerical example reveals that as the bar length to internal length ratio $l_o/\sqrt{c}$ changes from 20 to 4 i.e. the bar length becomes comparable to the internal length the simulation gives higher stiffness values. Moreover since the simulation values of the displacement at the nodes of a truss model depend on the stiffness of the bar, hence from Table 1, for any particular specimen size, bar elements sizes comparable to its material micro-scale characteristic length should give lower displacement values. Consequently, it is implied that the simulation result of a specimen of particular size for any scale of interest will be different for different bar element sizes used. Smaller bars will be stiffer than those with larger ones, hence finer truss models would give lower displacement values than coarse models. Thus it is required first that the specimen size and a particular reference scale of interest be identified in order to choose the appropriate bar length that will adequately simulate the material or structure.

Numerical Example

In this numerical example, the gradient truss model is applied to the truss structure of Fig 3 simply supported. The vertical and horizontal bar length l_o = 10cm and the diagonal bars length $l_o\sqrt{2}$ = $10\sqrt{2}$cm, and all bars have the same Young's modulus E =100 Mpa.
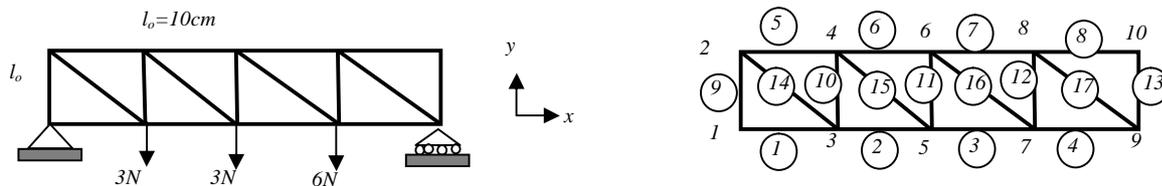
Fig. 3. (a). Loaded Seventeen element truss structure (b) Numbering of nodes and elements.

Nodal coordinate
{{0,0},{0,10},{10,0},{10,10},{20,0},{20,10},{30,0},{30,10},{40,0},{40,10}}

Applied node forces

{{0,0},{0,0},{0,-3},{0,0},{0,-3},{0,0},{0,-6},{0,0},{0,0},{0,0}}

Boundary Conditions
u = {0,0, 0,0, 0,-0.10, 0,-0.10, 0,-0.15, 0,-0.15, 0,-0.10, 0,-0.10, 0,0, 0,0}

For the gradient truss model the gradient elastic coefficient values considered are $\sqrt{c}$ = 0.5cm ($l_o/\sqrt{c}$ = 20); 1.0cm ($l_o/\sqrt{c}$ = 10); 1.5cm ($l_o/\sqrt{c}$ = 6.7); 2.0cm (($l_o/\sqrt{c}$ = 5); 2.5cm ($l_o/\sqrt{c}$ = 4), $\sqrt{c}$ = 3.0cm ($l_o/\sqrt{c}$ = 3.3); 3.5cm ($l_o/\sqrt{c}$ = 2.85); 4.0 cm ($l_o/\sqrt{c}$ = 2.5); 5.0cm (($l_o/\sqrt{c}$ =2); 6.0cm ($l_o/\sqrt{c}$ = 1.7).The

cross-sectional area is taken as the same for all the bars: A = 1cm^2. We proceed to obtain the element stiffness matrix in global coordinates.

By using the Direct Stiffness Method, a simple Mathematica program is used for this analysis. It comprises of three major processing stages: (1) the pre-processing, (2) processing, and (3) post processing.

The pre-processing stage is implemented by the driver program which puts the data structure in place by defining the model and directly setting the data structures.
The processing stage involves three major stages and the in-built Mathematica function *LinearSolve* is used: Firstly the master stiffness matrix is assembled with a subroutine element stiffness module; secondly, the master stiffness matrix and the node force vector are modified for the displacement boundary conditions; thirdly, the solution of the modified displacement equations is then obtained. Once these three processing stages are executed and the displacements made available the post processing stage follows.

At the post processing stage, through a **Ku** matrix multiplication, the forces are recovered to include the reactions. The internal (axial) forces in the truss elements are computed, and then the deflected shapes can be plotted.

Below the deflected shapes indicating the displacement are given for the classical and gradient elastic cases.

Classical Elastic Case
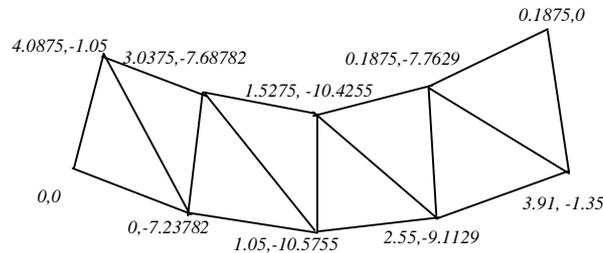
Fig. 4. Classical Elastic deformed configuration of truss structure.

Gradient Elastic Cases

For the truss structure of Fig. 3 the global stiffness matrices for Cases 1 – 10 and the deformation configurations are shown in Fig. 5.

Here also it can be observed that as the $l_o/\sqrt{c}$ ratio changes from 20 to 1.7 (i.e. the bar length becomes comparable to the internal length), the structure increases in stiffness significantly and the displacement values are much less than the classical ones.

Consequently, the truss model simulation result of a specimen of a particular size can be related to the particular scale of interest by identifying a ratio between a characteristic length scale of the specimen and the truss bar length to be used. This result is indicative that finer truss models of a particular specimen size would give lower displacement values than coarse models of the same specimen.
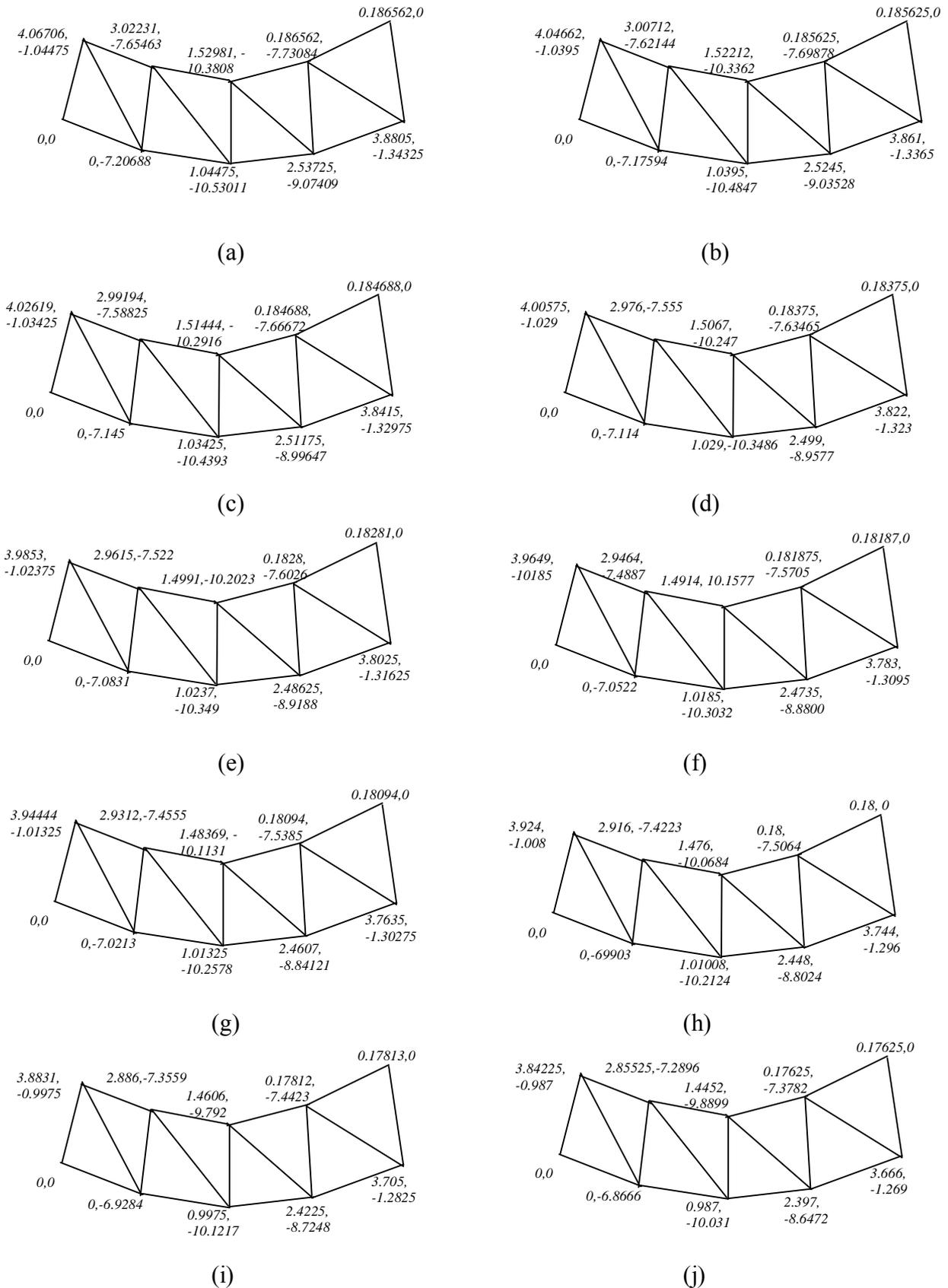
Fig. 5. Gradient Elastic deformed configuration of structure. (a). (l_o / $\sqrt{c}$ = 20); (b). (l_o / $\sqrt{c}$ = 10); (c). (l_o / $\sqrt{c}$ = 6.7); (d). ((l_o / $\sqrt{c}$ = 5; (e). (l_o / $\sqrt{c}$ = 4); (f). (l_o / $\sqrt{c}$ = 3.3); (g). l_o / $\sqrt{c}$ = 2.85); (h). (l_o / $\sqrt{c}$ = 2.5); (i). ((l_o / $\sqrt{c}$ =2); (j). (l_o / $\sqrt{c}$ = 1.7).

Conclusion

In this exploratory study size and scale dependent behaviour of truss models was examined using the gradient elasticity theory. With the introduction of a characteristic internal length parameter of a specimen at a defined scale, a ratio can be established with the truss bar element length.

It is shown that the simulation result of a truss structure of fixed size differ according to the ratio of the truss bar element length to the internal length of the material. The simulation results with the bar length comparable to the chosen internal length are shown to be stiffer. Hence smaller bars will be stiffer than larger ones and finer truss models comparable to the characteristic length scale will give lower displacement values at the nodes than coarser models.

Consequently, in order to adequately simulate a material or structure, a particular characteristic length scale of interest need be identified in relation to the specimen to be simulated and thus the appropriate corresponding truss bar element size can be identified and used.

References

Argyris, J.H. (1978,1981, 1994) Editor, Conferences on Finite Elements in Nonlinear Mechanics (FENOMECH), Stuttgart.

Bažant, Z. P., Tabbara, M. R. and Kazemi, M. T., and Pijaudier-Cabot, G. (1990). "Random particle model for fracture of aggregate or fiber composites." J. Eng. Mech., 116, (8), 1686–1705.

Bazant, Z. (1997) Fracturing truss model: Size effect in shear failure of reinforcement, J. of Engrg. Mech. ASCE. 123, 1276-1288.

Bazant Z.P and Planas, J. (1998) Fracture and size effect in concrete and other quasi brittle materials. CRC press, Boca Raton, FL.

Goel, H. C., Stojadinovic, B. and Lee, K. H., (1997) Truss analogy for steel moment connection, Aisc Engrg. J. 2nd quarter, 43-53.

Akintayo F., Papadopoulos, P.G. and Aifantis, E.C. (1998) *Simulation of Uniaxial Compression of a Concrete Column by a Truss Model with Instability*, Proc. of 5th National Congress on Mechanics, Ioannina, Greece, Vol. 2, pp.899-906.

Akintayo, F., Papadopoulos, P.G. and Aifantis, E.C. (2000) Influence of strain gradient on the ductility of a reinforced concrete column: Abnormal loading on structures: experimental and numerical modelling, Ed. K. S. Virdi 162(10). Spon Press.

Papadopoulos P.G and Xenidis H.C. (1997) Confinement of an R/C column which prevents Structural Instability of Concrete, Journal of Engineering Mechanics ASCE, vol. 123, January

Papadopoulos, P.G. and Xenidis, H.C. (1998) Amount of Confinement Preventing Global Instability of a Concrete Column, Proc. 11th European Conference of Earthquake Engineering, Paris, Vol. 3.

Kiousis, P. D., Papadopoulos, P.G. and Xenidis, H. (2010) Truss Modeling of Concrete Columns in Compression, Journal of Engineering mechanics, Vol. 136, 1006-1013.

Salem, H.M. (2004) The micro truss model: an innovative rational design approach for reinforced concrete. Journal of Advanced Concrete Technology. 2(1), 77-87.

Nagarajan, P., Jayadeep U.B. and Madhavan Pillai, T.M., (2010) Mesoscopic numerical analysis of reinforced concrete beams using a modified micro truss model, Interaction and Multiscale Mechanics, Vol. 3, No. 1 23-37.

Aifantis E. C. (1999a) Strain gradient interpretation of size effect. *International journal of fracture* 95(1-4): 299 – 314.

Aifantis, E. C. (1999b) Gradient deformation at nano, micro and macro scale. *Journal of engineering material and technology Transaction of the ASME*, 121 (2): 189-202.

Gao, H., Huang, Y., Nix, W. D. and Hutchinson, J.W. (1999) Mechanism based strain gradient plasticity – I Theory, *J. Mech. Phys. Sol.* **47**, 1239 – 1263.

Fleck, N. A., Muller, G. M., Ashby, M. F. and Hutchinson, J. W. (1994) Strain gradient plasticity: theory and experiment, *Acta Metall. Mater.* 42, 475 – 487.

Fleck, N. A. and Hutchinson J. W. (2001) A reformulation of strain gradient plasticity. *Journal of the mechanics and physics of solids*, 49 (10): 2245 – 2271.

Freund, L. B. (1987) The stability of a dislocation threading a strained layer on a substrate. *Journal of applied mechanics – Transaction of ASME*, 54 (3): 553 – 557.

Thompson, C. V. (1993) The yield stress of polycrystalline thin films. *Journal of materials research*, 8 (2): 237 – 238.

Nix W. D. (1998) Mechanical properties of thin films. *Metallugical transaction A- physical metallurgy and metals science*, 20 (11): 2217 – 2245.

Zbib, H. M. and Aifantis E. C. (2003) Size effects and length scales in gradient plasticity and dislocation dynamics. *Scripta materialia* 48 (2): 155-160.

Needleman A., and Van der Giessen, E., (2003). Discrete dislocation plasticity. In engineering plasticity from macroscale to nanoscale, volume 233-2 of key Engineering materials pages 13-24.

Atkins A. G. (1999) Scaling laws of elastoplastic fracture. *International Journal of Fracture*, 95 (1-4): 51-65.

Irwin, G. R. (1964) Dimensional and geometric aspects of fracture. In fracture of engineering material, American society for metals, pages 211-230.

Liu, J., and Zenner H. (1995) Estimation of s-n curves under consideration of geometrical and statistical size effect. *Materialwissenschaft und werkstofftechnik*, 26 (1): 14-21.

Seifried, A. (2004) About statistics in fatigue strength. *Materialwissenschaft und werkstofftechnik*, 35(2): 93-111.

R. D. Mindlin. (1964) Microstructure in linear elasticity, *Arch. Ration. Mech. Anal.*, 16, 51-78.

Ben-Amoz, M. (1976) A dynamic theory for composite materials. *J Appl. Math. Ph. (ZAMP)* 27, 83–99.

Aifantis, E. C. (1984) On the microstructural origin of certain inelastic models, *Transactions of ASME, J. Engng. Mat. Tech.* 106, 326-330.

Askes H. and Aifantis, E. C. (2011) Gradient elasticity in statics and dynamics: An overview of formulations, length scale identification procedures, finite element implementations and new results, *International Journal of Solids and Structures*, **48**, 13 1962-1990.

Pamin, J. and de Borst, R. (1998) Simulation of crack spacing using a reinforced concrete model with an internal length parameter, Arch. Appl. Mech., 68, 9, 613-625.

Chang, C. S., Askes, H. and Sluys, L.J. (2002) Higher-order strain/higher order stress gradient models derived from a discrete microstructure, with application to fracture, Engineering Fracture mechanics, 69 1907-1924.

Dessouky, S., Masad, E. Zbib, H., and Little, D. (2003) Gradient elasticity finite element model for the microstructure analysis of asphaltic materials. Proceedings Second MIT Computational Fluid and Solid Mechanics 228-233

Akarapu, S. and Zbib, H. M. (2006) Numerical analysis of plane cracks in strain-gradient elastic materials International Journal of Fracture, Vol. 141 Nos. 3-4, 403-430.

Altan. B.S. and Aifantis, E.C. (1997) On some aspects in the special theory of gradient elasticity, J. Mech. Behavior Mats. 8, 231-282.

Akintayo, O. T. (2011) Analytical and Numerical Study of the Behavior of Materials and Structures in Gradient Elasticity, Ph.D. dissertation, Gen. Dept. Eng. Sch., Aristotle University Thessaloniki.

Akintayo, O. T., (2012) Papadopoulos P. G., and Aifantis E. C., A note on gradient truss models, *International Review of Mechanical Engineering* Vol. 6, N. 4, 691 – 697, May.

Akintayo, O. T. (2014) Towards a gradient truss model Part I: Bar element displacement-force relations, *International Review of Civil Engineering* Vol. 5, N. 1, 32 – 42, January.

Geometrically exact beam dynamics, with and without rotational degree of freedom

*Tien Long Nguyen[1], Carlo Sansour[2], and Mohammed Hjiaj[1]

[1]Department of Civil Engineering, Institut National des Sciences Applquées de Rennes (INSA de Rennes), Rennes, France
[2]Faculty of Engineering, Nottingham University, Nottingham, United Kingdom

*Corresponding author: tien-long.nguyen@insa-rennes.fr – tien.nguyen@nottingham.ac.uk

Abstract

Non-linear rod dynamics is the focus of research in many engineering areas such as structural, aerospace and petroleum engineering as well as multibody dynamics. Also in non-classical areas such as biomechanics, micro- and nano-mechanics, geometrically exact formulations for rod dynamics are of importance. Rod formulations can be distinguished in regard to the basic kinematic assumption underlying the formulation. In the so-called Timoshenko-type beams, shear effects are taken into account and so rotational degrees of freedom describing the rotation of the cross section are considered. These are highly non-linear in nature. In contrast the Euler-Bernoulli assumption of zero shear deformation can be carried over into the non-linear regime resulting in displacement-only formulation but with highly non-linear expressions for the strain tensor incorporating higher gradients. In either formulation, the integration of the time dependent equations is challenging. It has been recognised that energy conservation is key for stable integration in long term dynamics. The so-called energy-momentum methods is a class of integrators, which, by design, conserve the momentum, angular momentum and the energy in the discrete case, if the same conservation properties are present in the continuous case. While for the Timoshenko beam some progress has been made and specific energy-momentum methods are known in the literature, the same is not true far the higher-gradient beam formulation of the Bernoulli beam.

In this paper, we are going to develop a unified formulation of an energy-momentum integration scheme for both geometrically exact Bernoulli and Timoshenko beams. We will show that the stable integration in either case is achievable with excellent results. Further important novel aspect of the models are the full incorporation of the rotational inertia. A range of applications from structural dynamics to flexible multibody dynamics will show the excellent performance of the new energy-momentum integration scheme.

Keywords: Non-linear dynamics, Computational method, Euler-Bernoulli rod, Timoshenko rod, Energy-momentum method, Multi-body dynamics.

Introduction

Dynamics of beam as well as in many new emerging areas of applications such as nano, bio mechanics, remains a very active topic of research. In the geometrically exact beam theories, a popular approach is the one so-called Timoshenko kinematics which still makes use of the assumption of planar cross sections in the deformed configurations but allows for shear to be considered by dropping the assumption that the vector normal to the centre line remains normal after the deformation [Ibrahimbegovic and Fray (1993); Iura and Iwakuma (1991)]. An example is the formulation of Simo and Vu Quoc [Simo and Vu-Quoc (1986)] which is based on a previous intrinsic formulation by Reissner [Reissner (1972)]. The Timoshenko type beam is well known to be suitable for short beams, sandwich composite beams and high-frequency excitation beams. On the contrary, in many applications, we desire to have a displacement-only formulation for example

in mechanisms, nano and bio mechanics where the Euler-Bernoulli model is the best choice dispite of the complexities involved due to the kinematics assumptions.

Beside the kinematics descriptions and the strain measures, the time integration of the dynamical equations has also been the focus of research for decades. It is now generally accepted that classical time integration methods such as Newmark, standard midpoint rule do suffer severe shortcomings [Newmark (1959); Chung and Hulbert (1993)]. Especially the lack of stability is a major issue. It has been soon recognized that the conservation of energy is the key for the stability of the time integration scheme. Moreover, an efficient time integrator so-called energy-momentum method has been developed which conserves not only the energy but also the momentum and the angular momentum of the system. This method can provide good accuracy and stability in long-term dynamics. The first attempt to an energy-momentum method was proposed by Simo and Tarrow [Simo and Tarnow (1993)] but this algorithm is only valid for quadratic-nonlinearities. The method has many applications in Timoshenko beams but the treatment of rotation is anything but trivial, especially when it comes to incorporate the inertial term. Due to the highly complex non-linearity as a result of the kinematics assumptions, such this formulation is not as common for Euler-Bernoulli beam model.

In this paper, we aim to develop an energy-momentum integration scheme for geometrically exact Bernoulli and Timoshenko beams. Numerical examples will be provided to show the excellent performance of the method.

Kinematics, dynamics equation and finite discretization

Euler-Bernoulli beam model and kinematics description

Let $\mathcal{B} \subset \mathbb{R}^3$, with $\mathbb{R}$ denoting the real numbers, define a reference configuration of the body. Without loss of generality we want to identify the reference configuration with the body itself. The actual configuration is denoted by $\mathcal{B}_t \in \mathbb{R}^3$. We assume that our body is thin in two dimensions such that it is rod-like with a cross section A at the reference configuration. The material particles are identified by their position vectors $\in \mathcal{B}$, the corresponding placement at the actual configuration by $\mathbf{x} \in \mathcal{B}_t$. A deformation is a map $\mathbf{x} = \varphi(\mathbf{X})$, the gradient of which defines the deformation gradient $\mathbf{F} = \frac{\partial \mathbf{x}}{\partial \mathbf{X}}$. We want to restrict ourselves to plane deformations and assume that the deformation lies in the $\mathbf{e}_1 - \mathbf{e}_2$ plane. For any material point in the cross section a suitable curvilinear coordinate system which we consider to be convected, is then given by the triple $s, z, \mathbf{e}_3$. z is the coordinate in the direction of the normal vector in the cross section. The relation holds

$$\mathbf{X}(s, z) = \mathbf{X}_0(s, z) + z\mathbf{N}(s), \tag{1}$$

where $\mathbf{X}_0(s, z)$ is the placement of the central line at the reference configuration. Correspondingly, $\mathbf{G}_0 = \frac{\partial \mathbf{X}_0}{\partial s}$, is a tangent vector. Similarly, we can introduce $\mathbf{G} = \frac{\partial \mathbf{X}}{\partial s}$, $\mathbf{G}_0 = \frac{\partial \mathbf{X}}{\partial s}\Big|_{z=0}$, $\mathbf{N} = \frac{\partial \mathbf{X}}{\partial z}$. The triple $(\mathbf{G}, \mathbf{N}, \mathbf{e}_3)$ defines a local curvilinear bases. The relations also hold

$$\mathbf{G}_0 \cdot \mathbf{N} = 0, \qquad |\mathbf{N}| = 1, \qquad \mathbf{N} = \mathbf{e}_3 \times \frac{\mathbf{G}}{|\mathbf{G}|}, \tag{2}$$

where $\times$ denotes the cross product of vectors, a dot denotes the scalar product of vectors. The corresponding tangent vectors at the deformed configuration are defined as $(\mathbf{g}, \mathbf{n}, \mathbf{e}_3)$ with $\mathbf{g}_0 = \mathbf{g}|_{z=0}$, $\mathbf{n}$ is the normal vector in the deformed configuration and

$$\mathbf{g} = \frac{\partial \mathbf{x}}{\partial s}, \qquad \mathbf{n} = \mathbf{e}_3 \times \frac{\mathbf{g}}{|\mathbf{g}|} = \mathbf{e}_3 \times \frac{\mathbf{X}_{0,s} + \mathbf{u}_{,s}}{|\mathbf{X}_{0,s} + \mathbf{u}_{,s}|}. \tag{3}$$

To derive the rod theory we adopt the Bernoulli hypothesis which assumes rigid cross sections and that the deformation can be completely characterized by the assumption

$$\mathbf{x} = \mathbf{X}(s) - z\mathbf{N}(s) + \mathbf{u}(s) + z\mathbf{n}(s) = \mathbf{X}_0(s) + \mathbf{u}(s) + z\mathbf{n}(s), \tag{4}$$

where $\mathbf{u}(s)$ is the displacement at the curvilinear coordinate s.

In the context of an in-plane Bernouilli beam, the right Cauchy deformation tensor has only one single non-trivial component which is C_{11} which reads

$$C_{11} = \left(\mathbf{X}_{0,s} + \mathbf{u}_{,s} + z\mathbf{n}_{,s}\right) \cdot \left(\mathbf{X}_{0,s} + \mathbf{u}_{,s} + z\mathbf{n}_{,s}\right), \tag{5}$$

where a comma denotes a derivative. With $\mathbf{E} = \frac{1}{2}(\mathbf{C} - 1)$ as the Green strain tensor, one is then left with one sigle non-trivial component E_{11} which is given by (the term in z^2 can be neglected since our thickness of the beam is small compared to its length)

$$E_{11} = \mathbf{u}_{,s} \cdot \mathbf{X}_{0,s} + \frac{1}{2}\mathbf{u}_{,s} \cdot \mathbf{u}_{,s} + z\left(\mathbf{n}_{,s} \cdot \left(\mathbf{X}_{0,s} + \mathbf{u}_{,s}\right) - \mathbf{N}_{,s} \cdot \mathbf{X}_{0,s}\right). \tag{6}$$

By defining ε_{11} as the axial strain, κ as the change of curvature, their expressions read

$$\varepsilon_{11} = \mathbf{u}_{,s} \cdot \mathbf{X}_{0,s} + \frac{1}{2}\left(\mathbf{u}_{,s} \cdot \mathbf{u}_{,s}\right) \tag{7}$$

$$\kappa = \mathbf{n}_{,s} \cdot \left(\mathbf{X}_{0,s} + \mathbf{u}_{,s}\right) - \mathbf{N}_{,s} \cdot \mathbf{X}_{0,s}. \tag{8}$$

Timoshenko beam model and kinematics description

Timoshenko beam model and Bernoulli one differ only in the assumption about the cross section which is still rigid but no longer perpendicular to the central line. Therefore, the kinematics should be described differently. The triple $(\mathbf{G}, \mathbf{N}, \mathbf{e}_3)$ defines the local curvilinear bases, $\mathbf{N}$ is the normal vector to the tangent space of the rod. $\mathbf{N}$ is given by the relation $\mathbf{N} = \mathbf{G} \times \mathbf{e}_3$. Altogether, the relations hold

$$\mathbf{G} = -\sin\alpha\ \mathbf{e}_1 + \cos\alpha\ \mathbf{e}_2, \qquad \mathbf{N} = \cos\alpha\ \mathbf{e}_1 + \sin\alpha\ \mathbf{e}_2, \tag{9}$$

Where α is a function of s and determines the angle closed between $\mathbf{e}_1$ and $\mathbf{N}$. Accordingly, the displacement field is defined by

$$\mathbf{u} = \mathbf{x} - \mathbf{X}.$$

Furthermore, one has $\mathbf{g} = \mathbf{G} + \mathbf{u}_{,s}$ and correspondingly we obtain $\mathbf{F} = (\mathbf{G} + \mathbf{u}_{,s}) \otimes \mathbf{G}$.

We consider now a rotational field $\mathbf{R} \in SO(3)$, where $SO(3)$ is a group of orthogonal tensors with positive determinant. Since we remain in plane $(\mathbf{e}_1 - \mathbf{e}_2)$, the rotation vector is fixed to vector $\mathbf{e}_3$. Therefore, we obtain the following expression of the rotation tensor

$$\mathbf{R} = \cos\omega\ (\mathbf{e}_1 \otimes \mathbf{e}_1 + \mathbf{e}_2 \otimes \mathbf{e}_2) - \sin\omega\ (\mathbf{e}_1 \otimes \mathbf{e}_2 - \mathbf{e}_2 \otimes \mathbf{e}_1) + \mathbf{e}_3 \otimes \mathbf{e}_3.$$

We denote the corresponding axial vector by $\mathbf{k}$. To get the strain measures, we apply the direct method of a Cosserat line. Accordingly, we get the first Cosserat deformation tensor (the stretch tensor) $\mathbf{U} := \mathbf{R}^T\mathbf{F}$, the second Cosserat strain tensor $\mathbf{K} := -\mathbf{k} \otimes \mathbf{G}$. Because the Cosserat is assumed to be one-dimensional and in-plane deformation, we can write down $\mathbf{u}, \mathbf{U}, \mathbf{K}$ as follows $\mathbf{u} = u_1\mathbf{e}_1 + u_2\mathbf{e}_2$, $\mathbf{U} = U_{11}\mathbf{G} \otimes \mathbf{G} + U_{13}\mathbf{G} \otimes \mathbf{N}$, $\mathbf{K} = \kappa\ \mathbf{e}_3 \otimes \mathbf{e}_3$, where the expressions of the components are defined as

$$U_{11} = \cos\omega - \sin(\alpha + \omega)\ u_{1,s} + \cos(\alpha + \omega)\ u_{2,s}, \tag{10}$$

$$U_{13} = \sin\omega + \cos(\alpha + \omega)\ u_{1,s} + \sin(\alpha + \omega)\ u_{2,s}, \tag{11}$$

$$\kappa = \omega_{,s}. \tag{12}$$

In this model, the shear deformation is included in this model and is not assumed to be zero.

Dynamics equations

Starting from Hamilton Principle for our conservative mechanical system, the dynamics equation for our Bernoulli beam is written down as follows

$$\rho A \int_L (\ddot{\mathbf{u}} \cdot \delta\mathbf{u})ds + \rho I \int_L \ddot{\mathbf{n}} \cdot \delta\mathbf{n}ds + \int_0^L (EA\varepsilon_{11}\delta\varepsilon_{11} + EI\kappa\delta\kappa)ds - \int_0^L (\mathbf{p} \cdot \delta\mathbf{u} + M\delta\omega)ds +$$
$$(\mathbf{p} \cdot \delta\mathbf{u} + M\delta\omega)|_0^L = 0, \tag{13}$$

where E is Young's Modulus of the material, $\mathbf{p}$ is external force and M is external moments, I is the moment of inertia of the section and L is the length of the beam, ρ is the material density.
The dynamics equation for Timoshenko beam has the following form

$$\rho A \int_L (\ddot{\mathbf{u}} \cdot \delta\mathbf{u})ds + \rho I \int_L \ddot{\omega}\delta\omega ds - \int_L \left[\delta\mathbf{n} \cdot \boldsymbol{\epsilon}(\mathbf{n}, \omega) + \mathbf{n} \cdot \frac{\partial\boldsymbol{\epsilon}(\mathbf{u},\omega)}{\partial\mathbf{u}}\delta\mathbf{u} + \mathbf{n} \cdot \frac{\partial\boldsymbol{\epsilon}(\mathbf{u},\omega)}{\partial\omega}\delta\omega - \frac{\partial\tilde{\psi}(\mathbf{n})}{\partial\mathbf{n}}\delta\mathbf{n} + \right.$$
$$\left. \frac{\partial\psi_2(\omega)}{\partial\omega}\delta\omega \right] ds + \int_0^L (\mathbf{p} \cdot \delta\mathbf{u} + M\delta\omega)ds + (\mathbf{p} \cdot \delta\mathbf{u} + M\delta\omega)|_0^L = 0, \tag{14}$$

where $\mathbf{n}$ and $\boldsymbol{\epsilon}$ is the force and strain vectors respectively, ψ is the stored energy, $\tilde{\psi}$ is the complementary energy related to the strains $\boldsymbol{\epsilon}$ using Legendre transformation in order to avoid locking phenomena and construct robust finite elements.

Regarding to finite element approach in the Bernoulli case, given the fact that second derivatives are present in the equations (a result of the Bernoulli hypothesis), the finite element formulation must exhibit continuous first derivatives. Hence, we resort, within a finite element context to interpolation functions defined by cubic Hermite polynomials. For Timoshenko model, the finite element will be of hybrid type, a two-node elements with linear kinematical fields and constant force $\mathbf{n}$ (normal and shear components) is considered.

Energy-momentum time integration scheme

After the spatial discretisation via the finite element method, the numerical approach is completed by devising a step-by-step time integration scheme for the time dependent equations. Classical implicit schemes like the Midpoint rule or Newmark method have been very popular in the structural dynamics community. However while these are stable integration methods in the linear regime, they proved less so in the highly non-linear one, especially in long-term dynamics. They suffer from numerical instabilities like blow-ups as well documented in the literature [Bathe (1997); Sansour et al. (1997); Sansour et al. (2004)]. Energy-momentum methods proved to provide here the necessary stability. In what follows we will develop such a method tailored to our rod formulation. However, so far no such formulation was attempted for the Bernoulli beam because of the complexities involved in the kinematic assumptions. In the following we want to develop for the first time such an Energy-momentum method. In doing so, we resort to an idea developed in [Sansour et al. (1997); Sansour et al. (2004)]. The method described there is attractive because it is independent of the involved non-linearity, the source of problem in the presently considered beam. The starting point, however is the standard midpoint rule. From step n, where all kinematical fields and velocities are known, we need to find these quantities at time step n+1, Consider ξ to be a scalar which defines any position within the time interval ΔT, with $0 \le \xi \le 1$. We start with the following expressions:

$$\mathbf{x_{n+\xi}} = \mathbf{x_{n+1}} + (\mathbf{1} - \boldsymbol{\xi})\mathbf{x_n}, \tag{15}$$

$$\dot{\mathbf{x}}_{\mathbf{n+\eta}} = \frac{\mathbf{x_{n+1}} - \mathbf{x_n}}{\mathbf{\Delta T}}, \tag{16}$$

$$\ddot{\mathbf{x}}_{\mathbf{n+\eta}} = \frac{\mathbf{x_{n+1}} - \mathbf{x_n}}{\mathbf{\Delta T}}, \tag{17}$$

Where $\boldsymbol{\eta}$ is an open parameter. The first defines a convex set, the following two are true for some value of $\boldsymbol{\eta}$. The midpoint rule corresponding to $\boldsymbol{\eta} = 0.5$.

The key step is to employ strain velocity fields to define the strain fields. Let us consider the following velocity fields:

$$\dot{\varepsilon} = \dot{\boldsymbol{u}}_{,s} \cdot \boldsymbol{X}_{0,s} + \boldsymbol{u}_{,s} \cdot \dot{\boldsymbol{u}}_{,s}. \tag{18}$$

$$\dot{\kappa} = \left(\frac{\partial \kappa}{\partial \boldsymbol{u}_{,s}} \cdot \dot{\boldsymbol{u}}_{,s} + \frac{\partial \kappa}{\partial \boldsymbol{u}_{,ss}} \cdot \dot{\boldsymbol{u}}_{,ss} \right) \tag{19}$$

Given the strain field defined at time n, the strain field at step n+1 then defined as following:

$$\varepsilon_{n+\xi} = \varepsilon_n + \xi\Delta T\dot{\varepsilon}_{n+\frac{1}{2}} \tag{20}$$

$$\kappa_{n+\xi} = \kappa_n + \xi\Delta T\dot{\kappa}_{n+\frac{1}{2}} \tag{21}$$

Specifically for $\xi=1$, the relations hold

$$\varepsilon_{n+1} = \varepsilon_n + \Delta T\dot{\varepsilon}_{n+\frac{1}{2}} \tag{22}$$

$$\kappa_{n+1} = \kappa_n + \Delta T\dot{\kappa}_{n+\frac{1}{2}} \tag{23}$$

This time integration scheme is proved formally and numerically to be stable and which conserve the energy, the momentum and the angular-momentum for a dynamic nonlinear system in long-term, some example are provided in the next section.

Numerical example

Example 1: Flying beam

In the first example, to investigate not only the conservation of energy but also of momentum and angular momentum, we consider a flying beam without support, the beam is depicted in Fig. 1 The loading increases linearly to a peak and decreases at the same rate to zero, Fig. 2. We run the calculation for one million time steps with $\Delta T = 1\text{E} - 4\text{s}$.
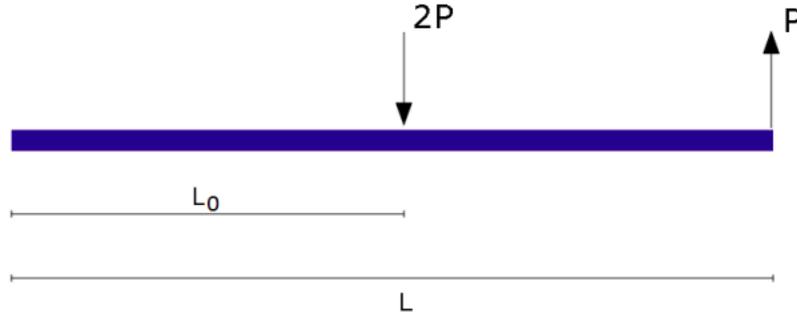
Figure 1: Beam figure

Table 1. Parameters

Parameters	symbol	value
Beam length	L	3m
Length	L_0	1.5m
Cross section area	A	200 cm^2
Cross section inertia	I	66.67 cm^4
Young's modulus	E	0.2E12 Pa
Density	ρ	48831Kg/m^3
Number of elements		4
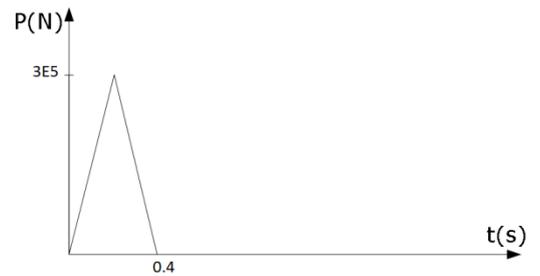
Figure 2: Loading history
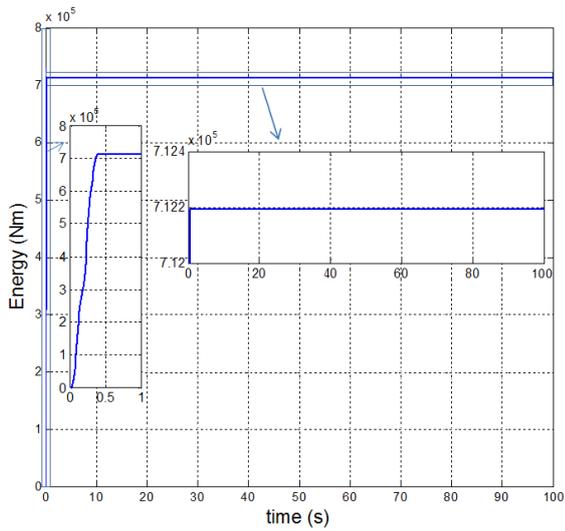
Figure 3: Energy history
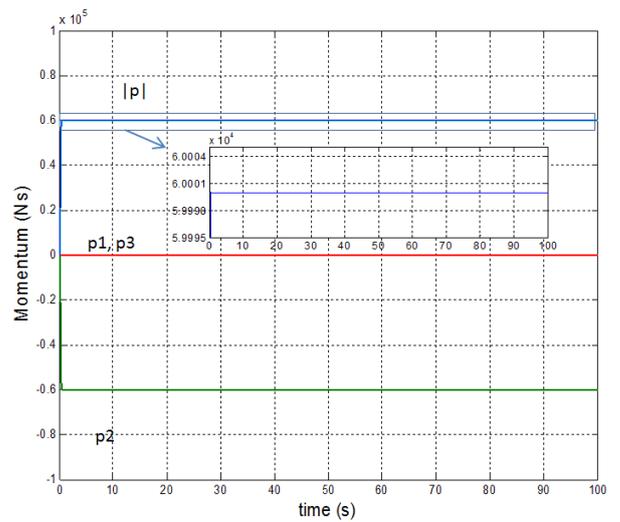
Figure 4: Momentum history
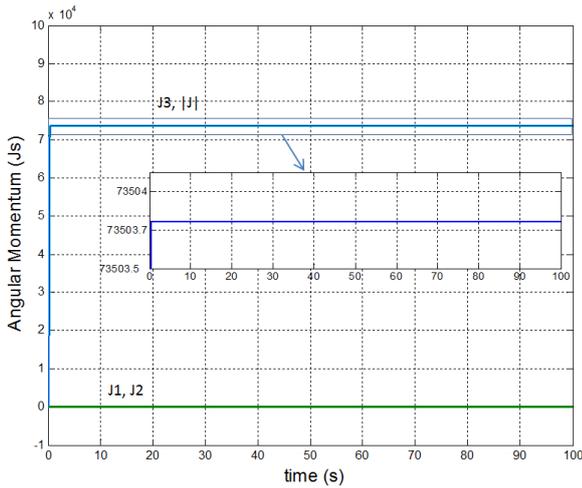
Figure 5: Angular momentum history
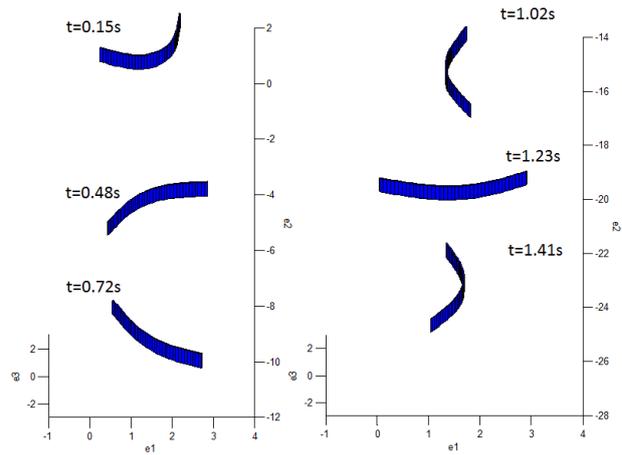
Figure 6: Beam motion snapshots

The Energy history is depicted in Fig. 3, while Fig. 4 and Fig. 5 reflect the linear momentum and angular momentum, respectively. In both figures not only the absolute value but also the components of the mentioned quantities are considered. Conservation is valid for the momentum and the angular momentum vector. Some deformations of the beam in space are captured in Fig. 6 which shows not only that the beam experiencing high deformation but also large overall displacement (24m at t≈1.4s).

Example 2: Chaotic motion of shallow arch

In this example, we investigate a chaotic motion of an arch. The arch configuration is given in Fig.~\ref{shaar}. From the configuration it can be seen that the arch is shallow and indeed can undergo a snap-through phenomena.

We consider here a system with a velocity-dependent damping with damping parameter D=2.5E-3 Ns/m. The system is subjected to a time-dependent concentrated force at its center of the form P = Fcos(ωt + Π), ω = 1000 Hz , F = 800N. The excitation can be modified either by changing its amplitude or its frequency.
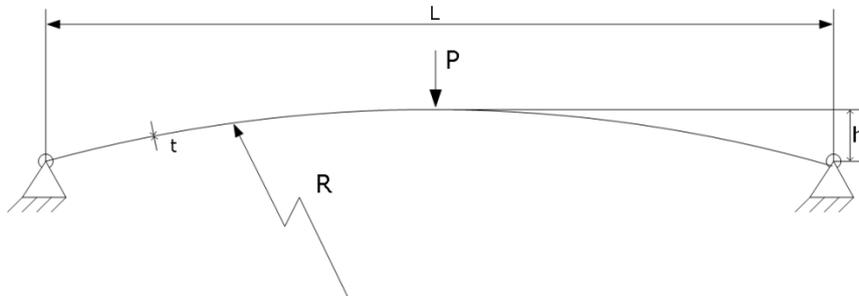
Figure 7: Shallow arch

Parameters:

Length L = 70cm
Height h = 1.53 cm
Density $\rho = 7.5E3$ kg/m^3
Time increment: $\Delta T = 1E - 4s$

Thickness t= 1 cm
Young's modulus E=0.2E12 Pa
Number of elements = 10
Number of time steps = 3E6
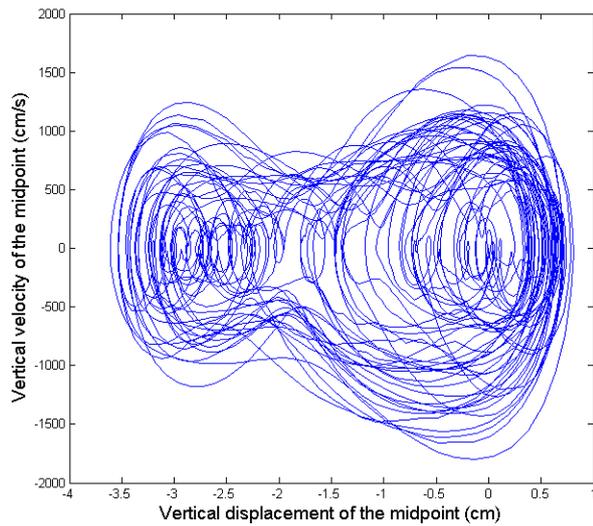
Figure 8: Phase space
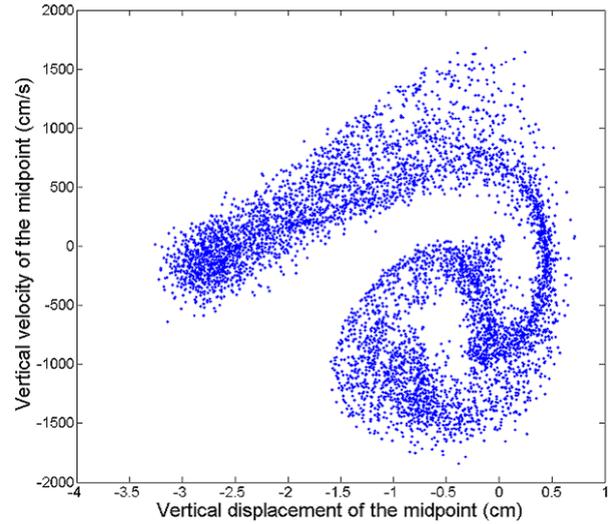
Figure 9: Poincaré section

Fig. 8 shows the phase space which plots the displacement against the velocity of the midpoint on the arch at each time step. The Poincaré section is presented in Fig.9. Those graphs show that the motion is chaotic which means long-term calculation is applicable.

Conclusion

A new time integration scheme has been presented for in-plane geometrically exact beam with/without rotational degree of freedom. The results showed an excellent performance of the method in term of accuracy and stability.

References

Bathe, K.J. (2007) Conserving energy and momentum in nonlinear dynamics: A simple implicit time integration scheme, Computers and Structure 85, 437-445.

Simo, J.C., Vu-Quoc, L. (1986) On the dynamics in space of rods undergoing large motions a geometrically exact approach, Computer *Methods in Applied Mechanics and Engineering* **66**, 125-161.

Ibrahimbegovic, A., Frey, F. (1993) Finite element analysis of linear and non-linear planar deformations of elastic initially curved beams, *International Journal for Numerical Methods in Engineering* **36**, 3239-3258.

Iura, M., Iwakuma, T. (1991) Dynamic analysis of the planar Timoshenko beam with finite displacement, *International Computers and Structures* **45**, 173-179.

Reissner, E.(1972) On one-dimensional finite-strain beam theory: the plane problem. *Journal of Applied Mathematics and Physics* **23**, 795-804.

Newmark, N.M. (1959) *ASCE Journal of the Engineering Mechanics Division.* **85**, EM3.

Chung, J., Hulbert, G.M. (1993) A time integration algorithm for structural dynamics with improved numerical dissipation: the generalized-α method. *Journal of Applied Mathematics* **60**, 371–375.

Simo, J.C., Tarnow, N. (1992) The discrete energymomentum method. Conserving algorithms for nonlinear elastodynamics . *Journal of Applied Mathematics and Physics* **43**, 757-792.

Sansour, C., Wriggers, P., Sansour, J. (2004) On the design of energymomentum integration schemes for arbitrary continuum formulations . *International Journal for Numerical Methods in Engineering* **60**, 2419-2440.

Sansour, C., Wriggers, P., Sansour, J. (1997) Nonlinear dynamics of shells: theory, finite element formulation, and intergration scheme. *Nonlinear dynamics* **13**, 279-305.

Uncertainty in non-linear long-term behavior and buckling of shallow concrete-filled steel tubular arches

*X. Shi[1], W. Gao[1], Y.L. Pi[1]

[1]School of Civil and Environment Engineering, The University of New South Wales, Sydney, NSW 2052, Australia.

*Corresponding author: shixue1985@hotmail.com

Abstract：This paper presents a theoretical analysis for the non-deterministic time-dependent non-linear behavior of shallow concrete-filled steel tubular (CFST) arches with interval parameters under a sustained uniform load. The change ranges of the final shrinkage strain and final creep coefficients of concrete core are derived from experimental results. The virtual work method is used to establish the differential equations of equilibrium for the time-dependent behavior and buckling analyses of shallow CFST arches, and the age-adjusted effective modulus method is adopted to model the creep behavior of the concrete core. Analytical solutions of the interval time-dependent displacements and internal forces of shallow CFST arches are derived. The lower and upper bounds of structural responses are determined. Comparisons of the interval analytical solutions with the interval finite element results show that the analytical solutions of the present study are accurate.

Keywords:　CFST arches, interval analysis, creep, shrinkage

1. Introduction

Applications of concrete filled steel tubular (CFST) arches are increasing in engineering structures, particularly in bridge constructions. The visco-elastic effects of creep and shrinkage of concrete core are inevitable in the long term for CFST arches. When a CFST arch is subjected to a sustained load, the creep of the concrete core will lead to the increase of its deformations with time and the deformations may be significant, while the shrinkage strain may also develop even when the arch is not subjected to any load [B.C. Chen 2000]. Hence, an investigation of significant effects of creep and shrinkage of the concrete core on the time-dependent structural behavior of CFST arches is much needed.

However, it is noted that the creep coefficient obtained from tests vary significantly from one experiment to another. Very different predictions of the time-dependent behavior of CFST columns have been reported in different studies. This shows that the uncertainties of creep and shrinkage of the concrete core do exist. To predict the long-term behavior of CFST columns reasonably, these uncertainties have to be considered.

In this paper, intervals are adopted to represent the uncertainties. This paper, therefore,

aims to investigate the interval time-dependent non-linear behavior and buckling of shallow CFST arches under a uniform radial load, to derive analytical solutions for their time-dependent non-linear deformations, internal forces and buckling, and to determine their structural life time prior to the buckling. To investigate the effects of the creep and shrinkage of concrete core, it is important to use an efficient and accurate method to describe the creep and shrinkage of concrete. [Pi et al. (2002), Pi et al. (2007)] It is known that a number of methods have been proposed and used for the creep and shrinkage of the concrete. Among these methods, the age-adjusted effective modulus method recommended by ACI Committee-209 and Australia design code for the concrete structures AS3600 are commonly considered to be efficient and accurate in evaluating the time-dependent behavior of the concrete and it could conveniently be incorporated into the structural analysis [ACI Committee 209 1982]. Algebraic formulas used in this method can be effective and practicable in modeling creep and shrink-age of concrete core, so the age-adjusted effective modulus method is used in this investigation.

2. Interval nonlinear elastic analysis of long-term behavior of shallow CFST arches

To predict the long-term performance, interval constitutive model considering creep and shrinkage of the CFST column needs to be established. The basic assumptions adopted for the interval long-term linear elastic analysis of CFST columns in this paper are: (1) deformations of CFST arch are elastic and satisfy the Euler-Bernoulli hypothesis, i.e. the cross-section remains plane and perpendicular to the arch axis during deformation; (2) the dimensions of the cross-section are much smaller than the length and radius of the arch so that the arches are sufficiently slender; and (3) the cohesion and adhesion of two different material components are fully bonded.

To account for the non-linearity resulted from creep and shrinkage of the concrete core, the derivation of the differential equations of equilibrium for shallow CFST arches need to consider non-linear longitudinal normal strain-displacement relationship and the non-linear longitudinal normal strain ε of an arbitrary point in the cross-section of shallow CFST arches can then be expressed as [Pi et al 2002]

$$\varepsilon = \tilde{w} - \tilde{v} + \frac{1}{2}(\tilde{v}')^2 - \frac{y\tilde{v}''}{R} \tag{1}$$

According to the third assumption, the deformations of each component should be compatible with each other, so their membrane strains and also the strains at the interface are identical. However, due to different Young's moduli and the effects of creep and shrinkage in concrete core, the stress σ_s in the steel tube and the stress

σ_c in the concrete core are different and they are given by

$$\sigma_c = E_{ec}(\varepsilon + \varepsilon_{sh}) = E_{ec}(u' + \varepsilon_{sh}) \tag{2}$$

and

$$\sigma_s = E_s \varepsilon \tag{3}$$

where E_{ec} is the age-adjusted effective modulus of concrete, ε_{sh} is the shrinkage strain of concrete and can be given by AS3600 [AS3600 2001]

$$\varepsilon_{sh}(t) = \frac{\varepsilon_{shfinal}}{35+t} \cdot t \tag{4}$$

where t is the loading time, $\varepsilon_{shfinal}$ is the final shrinkage strain of concrete when $t \rightarrow \infty$. E_{ec} can be calculated by

$$E_{ec}(t,\tau_0) = \frac{E_c}{1+\chi(t,\tau_0)\varphi(t,\tau_0)} \tag{5}$$

where τ_0 is the age at loading, $\chi(t,\tau_0)$ is the aging coefficient and $\varphi(t,\tau_0)$ is the creep coefficient that can be expressed as

$$\varphi(t,\tau_0) = [\frac{(t-\tau_0)^{0.6}}{10+(t-\tau_0)^{0.6}}] \cdot \varphi_{final} \tag{6}$$

where φ_{final} is the final creep coefficient when $t \rightarrow \infty$. The aging coefficient $\chi(t,\tau_0)$ can be expressed as

$$\chi(t,\tau_0) = 1 - \frac{(1-\chi^*)(t-\tau_0)}{20+(t-\tau_0)} \tag{7}$$

where

$$\chi^* = \frac{k_1\tau_0}{k_2+\tau_0} \tag{8}$$

with

$$k_1 = 0.78 + 0.4e^{-1.33\varphi_{\infty,7}} \tag{9}$$

$$k_2 = 0.16 + 0.8e^{-1.33\varphi_{\infty,7}} \tag{10}$$

$$\varphi_{\infty,7} = \varphi_{final}t_0^{0.118}/1.25 \tag{11}$$

The differential equations for the long-term analysis of a CFST arch can be obtained using a virtual work method. When the virtual work principle is used for the long-term equilibrium of the CFST arch, it can be stated as requiring that the functional

$$\int_{-\Theta}^{\Theta} [-NR(\delta\tilde{w}' - \delta\tilde{v} + \tilde{v}'\delta\tilde{v}') - M\delta\tilde{v}''] - qR^2\delta\tilde{v}\mathrm{d}\theta = 0 \tag{12}$$

where the axial compressive force N is given by

$$N = -\int_{A_s}\sigma_s\mathrm{d}A - \int_{A_c}\sigma_c\mathrm{d}A = -(A_sE_s + A_cE_{ec}{}^I)[\tilde{w}' - \tilde{v}' + \frac{1}{2}(\tilde{v}')^2] - A_cE_{ec}{}^I\varepsilon_{sh}{}^I \tag{13}$$

and the bending moment is given by

$$M = \int_{A_s} \sigma_s\, y \mathrm{d}A + \int_{A_c} \sigma_c\, y \mathrm{d}A = (E_s I_s + E_{ec}{}^I I_c)\frac{\tilde{v}''}{R} \tag{14}$$

in the axial direction:

$$N' = 0 \tag{15}$$

in the radial direction:

$$-M'' + NR\tilde{v}'' + NR - qR^2 = 0 \tag{16}$$

and leads to the static boundary conditions for pin-ended arches as
$M = 0$ at $\theta = \pm\Theta$
From Eq.(15), the axial compressive force N is constant along the arch axis. Substituting the constant axial compressive force N and the expression for M given by Eq.(14) into Eq.(16) leads to

$$\frac{\tilde{v}^{iv}}{(\mu_e{}^I)^2} + \tilde{v}'' = P \tag{17}$$

where $\mu_e{}^I$ is a time-dependent dimensionless axial force parameter defined by

$$\mu_e{}^I = \left[\sqrt{\frac{\underline{N}R^2}{E_s I_s + \overline{E_{ec}I}{}_c}}\, ,\, \sqrt{\frac{\overline{N}R^2}{E_s I_s + \underline{E_{ec}I}{}_c}}\,\right] \tag{18}$$

and P is a dimensionless load defined by

$$P^I = \frac{qR - N^I}{N^I} \tag{19}$$

By using the kinematic boundary conditions and the static boundary conditions and the kinematic boundary conditions , the solutions of Eq.(17) can be obtained as
For Pin-ended：
The radial displacement can be expressed as

$$\tilde{v}^I = \left[\frac{\overline{P}}{\underline{\mu_e}^2}\left\{\frac{\cos(\underline{\mu_e}\theta) - \cos(\underline{\mu_e}\Theta)}{\cos(\underline{\mu_e}\Theta)} + \frac{1}{2}[(\underline{\mu_e}\theta)^2 - (\underline{\mu_e}\Theta)^2]\right\}, \frac{\underline{P}}{\overline{\mu_e}^2}\left\{\frac{\cos(\overline{\mu_e}\theta) - \cos(\overline{\mu_e}\Theta)}{\cos(\overline{\mu_e}\Theta)} + \frac{1}{2}[(\overline{\mu_e}\theta)^2 - (\overline{\mu_e}\Theta)^2]\right\} \tag{20}$$

And the axial displacement can be expressed as:
Lower bound of axial displacement:

$$\underline{\tilde{w}} = \frac{(1-\overline{P})\overline{P}\theta(\theta^2 - \Theta^2)}{6} + \frac{(1+\overline{P})\overline{P}[\Theta\sin(\underline{\mu_e}\theta) - \theta\sin(\underline{\mu_e}\Theta)]}{\underline{\mu_e}^3\Theta\cos(\underline{\mu_e}\Theta)}$$

$$+ \frac{\overline{P}^2\theta}{\underline{\mu_e}^2}\left[1 - \frac{\cos(\underline{\mu_e}\theta)}{\cos(\underline{\mu_e}\Theta)}\right] + \frac{\overline{P}^2[\Theta\sin(\underline{\mu_e}\theta)\cos(\underline{\mu_e}\theta) - \theta\sin(\underline{\mu_e}\Theta)\cos(\underline{\mu_e}\Theta)]}{4\underline{\mu_e}^3\Theta\cos^2(\underline{\mu_e}\Theta)} \tag{21}$$

Upper bound of axial displacement:

$$\overline{\tilde{w}} = \frac{(1-\underline{P})\underline{P}\theta(\theta^2 - \Theta^2)}{6} + \frac{(1+\underline{P})\underline{P}[\Theta\sin(\overline{\mu}_e\theta) - \theta\sin(\overline{\mu}_e\Theta)]}{\overline{\mu}_e^{\,3}\Theta\cos(\overline{\mu}_e\Theta)}$$
$$+\frac{(\underline{P})^2\theta}{\overline{\mu}_e^{\,2}}[1 - \frac{\cos(\overline{\mu}_e\theta)}{\cos(\overline{\mu}_e\Theta)}] + \frac{\underline{P}^2[\Theta\sin(\overline{\mu}_e\theta)\cos(\overline{\mu}_e\theta) - \theta\sin(\overline{\mu}_e\Theta)\cos(\overline{\mu}_e\Theta)]}{4\overline{\mu}_e^{\,3}\Theta\cos^2(\overline{\mu}_e\Theta)} \tag{22}$$

For Fixed:

The radial displacement can be expressed as

Lower bound of radial displacement:

$$\tilde{v} = \frac{\underline{P}}{\underline{\mu}_e^{\,2}}\{\frac{(\underline{\mu}_e\Theta)[\cos(\underline{\mu}_e\theta) - \cos(\underline{\mu}_e\Theta)]}{\sin(\underline{\mu}_e\Theta)} + \frac{1}{2}[(\underline{\mu}_e\theta)^2 - (\underline{\mu}_e\Theta)^2]\} \tag{23}$$

Upper bound of radial displacement

$$\overline{\tilde{v}} = \frac{\overline{P}}{\overline{\mu}_e^{\,2}}\{\frac{(\overline{\mu}_e\Theta)[\cos(\overline{\mu}_e\theta) - \cos(\overline{\mu}_e\Theta)]}{\sin(\overline{\mu}_e\Theta)} + \frac{1}{2}[(\overline{\mu}_e\theta)^2 - (\overline{\mu}_e\Theta)^2]\} \tag{24}$$

And the axial displacement can be expressed as:

Lower bound of axial displacement:

$$\tilde{w} = \frac{(1-\overline{P})\overline{P}\theta(\theta^2 - \Theta^2)}{6} - \frac{(1+\overline{P})\overline{P}}{\underline{\mu}_e^{\,2}}[\theta - \frac{\Theta\sin(\underline{\mu}_e\theta)}{\sin(\underline{\mu}_e\Theta)}]$$
$$\frac{\overline{P}^2\theta\Theta}{\underline{\mu}_e\sin(\underline{\mu}_e\Theta)}[\frac{3\cos(\underline{\mu}_e\Theta)}{4} - \cos(\underline{\mu}_e\theta)] + \frac{\overline{P}^2\Theta^2\sin(\underline{\mu}_e\theta)\theta\cos(\underline{\mu}_e\theta)}{4\underline{\mu}_e\sin^2(\underline{\mu}_e\Theta)}$$
$$+\frac{\overline{P}^2\theta}{\underline{\mu}_e^{\,2}}[1 - \frac{\cos(\underline{\mu}_e\theta)}{\cos(\underline{\mu}_e\Theta)}] + \frac{\overline{P}^2[\Theta\sin(\underline{\mu}_e\theta)\cos(\underline{\mu}_e\theta) - \theta\sin(\underline{\mu}_e\Theta)\cos(\underline{\mu}_e\Theta)]}{4\underline{\mu}_e^{\,3}\Theta\cos^2(\underline{\mu}_e\Theta)} \tag{26}$$

Upper bound of axial displacement:

$$\overline{\tilde{w}} = \frac{(1-\underline{P})\underline{P}\theta(\theta^2 - \Theta^2)}{6} - \frac{(1+\underline{P})\underline{P}}{\overline{\mu}_e^{\,2}}[\theta - \frac{\Theta\sin(\overline{\mu}_e\theta)}{\sin(\overline{\mu}_e\Theta)}]$$
$$+\frac{\underline{P}^2\theta\Theta}{\overline{\mu}_e\sin(\overline{\mu}_e\Theta)}[\frac{3\cos(\overline{\mu}_e\Theta)}{4} - \cos(\overline{\mu}_e\theta)] + \frac{\underline{P}^2\Theta^2\sin(\overline{\mu}_e\theta)\theta\cos(\overline{\mu}_e\theta)}{4\overline{\mu}_e\sin^2(\overline{\mu}_e\Theta)}$$
$$+\frac{\underline{P}^2\theta}{\overline{\mu}_e^{\,2}}[1 - \frac{\cos(\overline{\mu}_e\theta)}{\cos(\overline{\mu}_e\Theta)}] + \frac{\underline{P}^2[\Theta\sin(\overline{\mu}_e\theta)\cos(\overline{\mu}_e\theta) - \theta\sin(\overline{\mu}_e\Theta)\cos(\overline{\mu}_e\Theta)]}{4\overline{\mu}_e^{\,3}\Theta\cos^2(\overline{\mu}_e\Theta)} \tag{27}$$

3. Interval finite element analyses

For interval buckling analysis, buckling load and equilibrium paths are sought by solving the equilibrium equation at each load increment. The equilibrium equation can be expressed as:

$$K_T^I \Delta u^I = \Delta P \tag{28}$$

At each load increment, perturbation method was employed to calculate the global

stiffness matrix. In the following, the effects of the beam axial force and bending are included. The creep coefficient and final shrinkage strain are considered as interval parameters. Using the first-order Taylor expansion and the interval arithmetic operations [Rao et al. 1997], the interval stiffness coefficients of the beam-column element can be obtained as

$$
K_{Ti}(\varphi_{final}{}^I, \varepsilon_{shfinal}{}^I) =
$$

$$
K_{Ti}(\varphi_{final}{}^c, \varepsilon_{shfinal}{}^c) + \frac{\partial K_{Ti}(\varphi_{final}{}^c, \varepsilon_{shfinal}{}^c)}{\partial \varphi_{final}{}^c}(\varphi_{final}{}^I - \varphi_{final}{}^c)
$$

$$
+ \frac{\partial K_{Ti}(\varphi_{final}{}^c, \varepsilon_{shfinal}{}^c)}{\partial \varepsilon_{shfinal}{}^c}(\varepsilon_{shfinal}{}^I - \varepsilon_{shfinal}{}^c) \tag{29}
$$

$$
= K_{Ti}(\varphi_{final}{}^c, \varepsilon_{shfinal}{}^c) + \frac{\partial K_{Ti}(\varphi_{final}{}^c, \varepsilon_{shfinal}{}^c)}{\partial \varphi_{final}{}^c}\Delta\varphi_{final} e_\Delta +
$$

$$
\frac{\partial K_{Ti}(\varphi_{final}{}^c, \varepsilon_{shfinal}{}^c)}{\partial \varepsilon_{shfinal}{}^c}\Delta\varepsilon_{shfinal} e_\Delta
$$

where i denotes the number of elements, $\varphi_{final}{}^c = \dfrac{\overline{\varphi_{final}} + \underline{\varphi_{final}}}{2}$, $\Delta\varphi_{final} = \dfrac{\overline{\varphi_{final}} - \underline{\varphi_{final}}}{2}$,

$\varepsilon_{shfinal}{}^c = \dfrac{\overline{\varepsilon_{shfinal}} + \underline{\varepsilon_{shfinal}}}{2}$, $\Delta\varepsilon_{shfinal} = \dfrac{\overline{\varepsilon_{shfinal}} - \underline{\varepsilon_{shfinal}}}{2}$, $e_\Delta = [-1,1]$. From Eq. (29), it can be easily observed that the stiffness is the function of the creep coefficient and final shrinkage strain.

The interval global stiffness matrix of the structure is assembled using the following equation

$$
K_T(\varphi_{final}{}^I, \varepsilon_{shfinal}{}^I) = \sum_{i=1}^{n} K_{Ti}(\varphi_{final}{}^I, \varepsilon_{shfinal}{}^I)
$$

$$
= K_T(\varphi_{final}{}^c, \varepsilon_{shfinal}{}^c) + \sum_{i=1}^{n}\left|\frac{\partial K_{Ti}(\varphi_{final}{}^c)}{\partial \varphi_{final}{}^c}\Delta E + \frac{\partial K_{Ti}(\varepsilon_{shfinal}{}^c)}{\partial \varepsilon_{shfinal}{}^c}\Delta E\right| e_\Delta \tag{30}
$$

Then the interval static equation of equilibrium in the interval finite element system becomes

$$
K_T(\varphi_{final}{}^I, \varepsilon_{shfinal}{}^I)\Delta(U^c + \Delta U e_\Delta) = \Delta P \tag{31}
$$

yields

$$
\Delta U = \left|K^{-1}{}_T(\varphi_{final}{}^c, \varepsilon_{shfinal}{}^c)\right|\left[\sum_{i=1}^{n}\left|\frac{\partial K_{Ti}(\varphi_{final}{}^c)}{\partial \varphi_{final}{}^c}\Delta E + \frac{\partial K_{Ti}(\varepsilon_{shfinal}{}^c)}{\partial \varepsilon_{shfinal}{}^c}\Delta E\right| e_\Delta\right]\left|U^c\right| \tag{32}
$$

To overcome the computational difficulties caused by the singularity of the tangent stiffness matrix at the limit points on the equilibrium path, the interval incremental-iterative solution is introduced for this study.

The incremental-iterative equilibrium equation at time t and iterative step i can be expressed as

$$^{t}[K_{T}]_{i}^{I}\,^{t}\Delta U_{i}^{I} = \,^{t}\Delta P(u) + \,^{t-\Delta t}\Delta P_{r} \tag{33}$$

where the subscript i denotes the iterative cycle, the superscript t denotes the load step, $^{t}\Delta P(u)$ is the incremental external forces at time t and $^{t-\Delta t}\Delta P_{r}$ is the unbalance forces at time $t - \Delta t$.

4. Model validation and discussions

4.1 Determination of intervals for the final shrinkage strain and creep coefficient

the interval of the final shrinkage strain and creep coefficient of their concrete cores can be derived from other researchers'[Uy.B 2001] test results as $\varepsilon_{shfinal}$ = [43.5, 340]

and φ_{final} = [0.5, 1.7] respectively, which are used in this study. It can be expected that the results obtained by the interval models proposed in this paper will contain these experimental results; in other words, the experimental results will fall into the interval bounds produced by the proposed models.

The creep coefficients determined by Uy [Uy.B 2001] is adopted to compare with the results obtained by the interval analytical model developed in this paper, which are illustrated in Figs. 1. Similarly, the total shrinkage strain determined by Uy is compared with the results given by the proposed interval model in Fig. 2. It can be easily seen that the results produced by the proposed model contain these experimental results as predicated. The bounds can be further updated if more experimental results available. These bounds will be useful for the future experimental investigations and design.
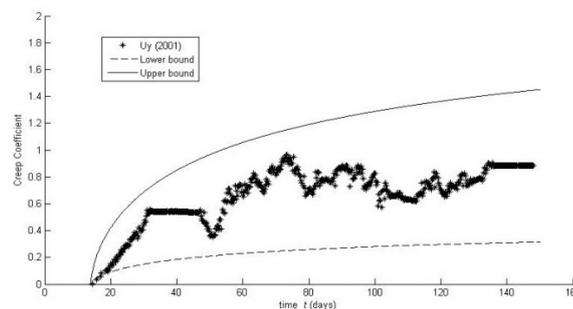
Fig.1 Comparison of creep coefficient
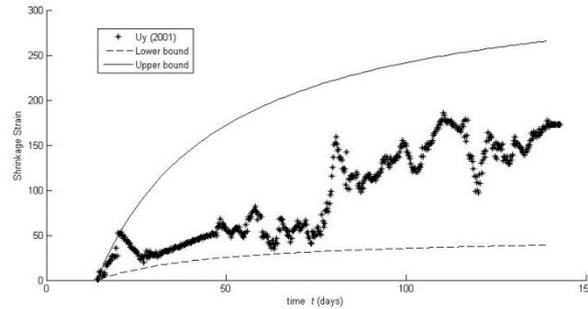
Fig.2 Comparison of shrinkage strain

The typical long-term nonlinear responses of pin-ended and fixed shallow CFST arches under a sustained uniform radial load are compared in Figs. 3 with their linear counter parts as the variations of the dimensionless central radial displacements $v_c / v_{c,15}$ and the dimensionless axial displacements is $w_{\theta/2} / w_{\theta/2,15}$ at the quarter point with time t, where the Young's moduli of the steel and concrete were assumed as Es=200 GP and Ec=30 GPa. .A circular cross-section with outer and inner radii: R=250 mm and r=240 mm was used in the investigation. The span of the arch was L=15 m and the rise-to-span ratio of the pin-ended and fixed CFST arch were f/L=1/12, respectively.
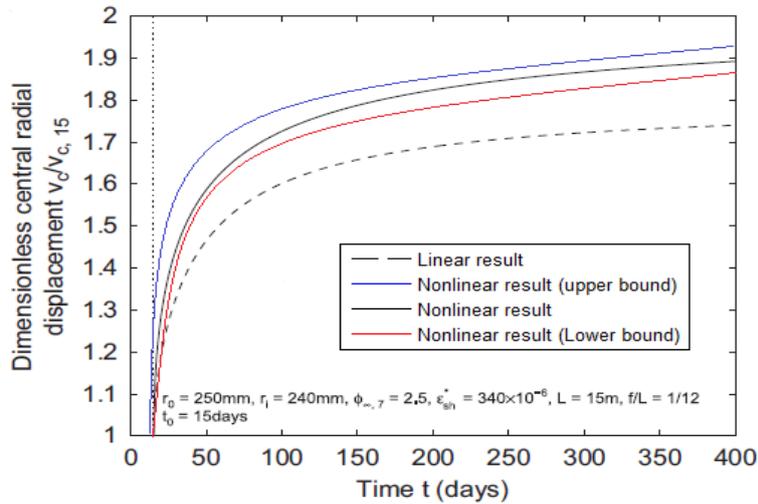
Figure 3 Interval radial displacements

5. Conclusions

This paper presents a theoretical study on the uncertain long-term and buckling analysis of shallow concrete-filled steel tubular arches subjected to a sustained uniform radial load. An interval analytical model based on the algebraically tractable age-adjusted effective modulus method is proposed to describe the time-dependent behavior of concrete in CFST arches. Non-linear analytical solutions for the time-dependent displacements and internal actions were derived. It has been found that creep and shrinkage of concrete core have significant effects on the

time-dependent non-linear deformations, internal forces and buckling behavior of shallow CFST arches.

An interval finite element was developed to describe the long-term behavior and analysis buckling. The buckling load or buckling time can be evaluated using this model. The result is compared with analytical results; it could be found it shows a good agreement.In the future, the proposed models will be further developed to analyze other types of CFST structures accounting for the uncertainties in their material and geometric properties.

Reference

ACI Committee 209, Prediction of Creep, Shrinkage and Temperature Effects in Concrete Structures, American Concrete Institute (ACI), Detroit, 1982.

AS3600, Australian Standard: Concrete Structures, Standard Association of Australia, Sydney, 2001

B.C. Chen, Experimental study on mechanical behaviors of concrete-filled steel tubular rib arch under in-plane loads, *Engineering Mechanics* **17** 44–50.

Rao SS, Berke L.Analysis of uncertain structural systems using interval analysis. *AIAA* J 1997;**35**(4):727±35.

Uy. B. 'Static long-term effects in short concrete-filled steel box columns under sustained loading.'' *ACI Struct. J.*, **98**, 96–104.

Y.-L. Pi, M.A. Bradford, B. Uy, In-plane stability of arches, *International Journal of Solids and Structures* **39** (2002) 105–125.

Y.-L. Pi, M.A. Bradford, F. Tin-Loi, Nonlinear analysis and buckling of elastically supported circular shallow arches, *International Journal of Solids and Structures* **44** (7–8) (2007) 2401–2425.

On mixed mode stress intensity factor evaluation using Interaction Integral Method for the tetrahedral finite element (for cracks with kinks)

*Y. Wakashima[1] and H. Okada[2]

[1]Department of Mechanical Engineering, Graduate School of Science and Technology, Tokyo University of Science
[2] Department of Mechanical Engineering, Faculty of Science and Technology, Tokyo University of Science

*Corresponding author: young_island_126@yahoo.co.jp

Abstract

We have been developing an interaction integral method for tetrahedral finite element, as a way to calculate the stress intensity factors for mixed mode crack problem. If cracks involve kinks, interaction integral method must have the surface integrals at the crack faces. The surface integral terms add complexities to the calculation procedures. In this paper, a method that can obviate the needs for the surface integral terms is proposed. Results that are computed by the proposed method are compared with there by the virtual crack closure-integral method. It is shown that the results of the proposed technique are more accurate than those of the conventional way.

Keywords: J-integral, Stress Intensity Factors, Virtual Crack Extension, Interaction Integral Method

Introduction

We need to calculate the stress intensity factors for the predictions of crack propagation behavior using the fracture mechanics concept. Finite element method is used for calculating the stress intensity factors. The computations of the stress intensity factors generally require the use of hexahedral finite element. However, the generations of finite element models for the cracks, consisting of the hexahedral finite elements require a large amount of manual labor and man time. On the other hand, the meshing can fully be automated if we use the tetrahedral finite elements. The tetrahedral finite elements have been commonly used in three-dimensional solid analysis in recent years.

We have been developing an interaction integral method for the quadratic tetrahedral finite element, as a way to calculate the stress intensity factors for mixed mode crack problem. And we showed that the method can calculate the stress intensity factors for mixed mode crack problem accurately, with the correction terms proposed by Daimon et al. (2014).

If cracks involve kinks, the interaction integral method has the surface integral terms at the crack faces. The surface integral terms add complexities to the calculation procedures. In this paper, a method that can obviate the needs for the surface integral terms is proposed as an extension of the interaction integral method with the correction terms. Results that are computed by the proposed method are compared with those by the virtual crack closure-integral method of Okada et al. (2008). It is shown that the results, that are obtained by proposed technique, are more accurate than those by the conventional way.

Calculation of mixed mode stress intensity factors by the interaction integral method

The J-integral can be evaluated by the domain integral method

The interaction integral method is based on the domain integral method of Nikishkov et al. (1987) that calculates energy release rate under linear elastic and small scale yielding conditions. The

formulation and the numerical evaluations of the domain integral method are briefly described, as follows.

The J-integral that was proposed by Rice (1968) evaluates the energy release rate in two-dimensional crack problems. The J-integral can also be defined in in three-dimensional problems as presented by Shivakumar et al. (1992). It is defined as a surface integral on an infinitesimally small pipe S_o surrounding the crack front, as depicted in Fig. 1. The size of Δ and ε/Δ are set to be 0 in a limit, and we can write the three-dimensional J-integral, as:

$$J = \lim_{\substack{\varepsilon/\Delta \to 0 \\ \Delta \to 0}} \frac{1}{\Delta} \int_{S_0} \left(W n_1^{\varepsilon} - n_j^{\varepsilon} \sigma_{ij} \frac{\partial u_j}{\partial x_1} \right) dS_0 \tag{1}$$

Here, W is the strain energy density, n_i are the unit outward normal vector, u_i are the displacements, and σ_{ji} are the stress tensor.
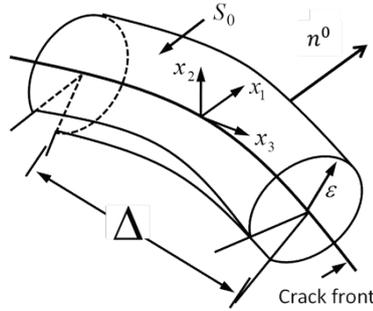
Figure 1. Definition of three-dimensional J-Integral

The J-integral can be evaluated by the domain integral method. Continuous and piecewise differentiable vector function q_i are introduced. They represent the virtual crack extension vector at the crack front and δA is the area of virtual crack extension. The three-dimensional J-integral can be represented by its domain integral form, as:

$$J = \frac{1}{\delta A} \lim_{\substack{\varepsilon/\Delta \to 0 \\ \Delta \to 0}} \int_{S_0} \left(W n^0{}_k - n^0{}_i \sigma_{ij} \frac{\partial u_j}{\partial x_k} \right) q_k dS_0 \tag{2}$$

The surface integral version of J-integral expression as shown in Eq. (2) is converted to the domain integral version by applying the Gauss divergence theorem. The domain of integration is as shown in Fig. 2. Here q_i are the components of vector function that are continuous and piece-wise differentiable in the domain of integration, the J-integral can be expressed by the following equation that is called the domain integral method.

$$
\begin{aligned}
J &= \frac{1}{\delta A} \lim_{\substack{\varepsilon/\Delta \to 0 \\ \Delta \to 0}} \int_{S_0 + S^+ + S^- + S_R + S_L + S_\varepsilon} (W n_k^0 - n_i^0 \sigma_{ji} \frac{\partial u_i}{\partial x_1}) q_k dS \\
&= -\frac{1}{\delta A} \left[\int_V \left(W \delta_{ik} - \sigma_{ij} \frac{\partial u_j}{\partial x_k} \right) \frac{\partial q_k}{\partial x_i} dV + \int_{S^+ + S^-} (W n_k^{\varepsilon} - n_i^{\varepsilon} \sigma_{ji} \frac{\partial u_i}{\partial x_k}) q_k dS \right]
\end{aligned} \tag{3}
$$

If the crack were planer, the crack surface integral terms vanish because $n_k q_k$ equals 0. But if cracks had kinks, $n_k q_k$ does not vanish, Eq. (3) contains a crack face integral.
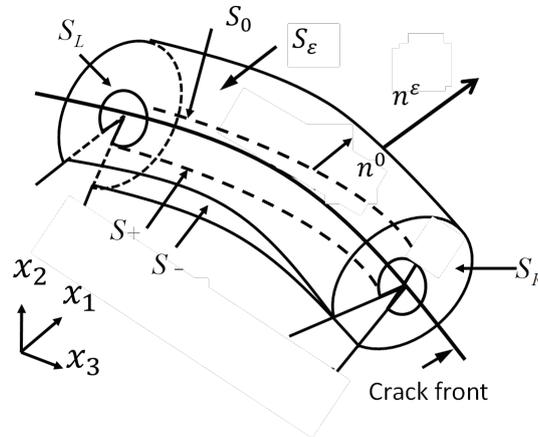
Figure 2. Definition of the domain integral method

Domain of integral method and distribution of value of q-function

To evaluate the domain integral we need to set the domain of integral and the variation of q-function. The domain is specified by using the q-function. In proposed domain integral method we set q-function to nodes by using the information of nodal locations. We set q-function values at each node by using the function $q = q(r, x_3)$ of distance r from the crack front and the coordinate x_3 in the tangential direction along the crack front. Here, r and x_3 are defined as shown in Fig. 2 (a). The q-function value linearly decreases as the distance r increases and q vanishes at the outer surface of the domain of integral. We define a trapezoid at a crack front node, as shown in Fig. 2(a) and rotate it around the crack front. The q-function values $q = q(r, x_3)$ are assigned to nodes inside the rotating body. Finite elements containing any nodes having non-zero q-function values are included in the domain of integral. As the result, the domain of integral has angular outer shape, as depicted in Fig. 2 (b). For the full details, the readers are referred to Okada and Ohata (2013).
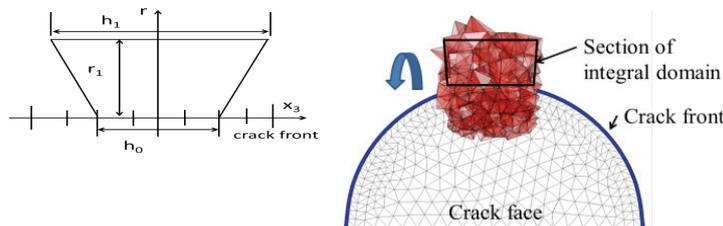
(a) Trapezoid in r - x_3 plane   (b) A typical outer shape of domain of integral
Figure 3. Domain of integral for the evaluation of J-integral

Description of interaction integral method

The interaction integral method is a technique to evaluate the mixed mode stress intensity factors based on the J-Integral [Yau et al. (1980)].

The J-integral expresses the energy release rate of a crack in linearly elastic material. The energy release rate G in an elastic body can also be expressed by the stress intensity factors (K_I, K_{II}, K_{III}). Therefore, we can write:

$$G_{total} = J = \frac{1}{E'}\left(K_I^2 + K_{II}^2\right) + \frac{1}{2\mu}K_{III}^2 = -\frac{1}{\delta A}\int_\Omega \left\{W\delta_{1i} - \sigma_{ij}\frac{\partial u_j}{\partial x_1}\right\}\frac{\partial q}{\partial x_i}d\Omega \tag{4}$$

Here, E' equals E for the plane stress and E' is $E/(1-v^2)$ for the plane strain. μ is the shear modulus and is expressed by $E/2(1+v)$. E is the Young's modulus and v is the Poisson's ratio.

The superposition of two solutions which satisfy the equilibrium can generate another equilibrium state. The independent equilibrium states are designated by superscripts (1) and (2). We can write:

$$u_j = u_j^{(1)} + u_j^{(2)}, \quad \sigma_{ij} = \sigma_{ij}^{(1)} + \sigma_{ij}^{(2)}, \quad \varepsilon_{ij} = \varepsilon_{ij}^{(1)} + \varepsilon_{ij}^{(2)} \tag{5}$$

$$W = \frac{1}{2}\left(\sigma_{ij}^{(1)} + \sigma_{ij}^{(2)}\right)\left(\varepsilon_{ij}^{(1)} + \varepsilon_{ij}^{(2)}\right) = W^{(1)} + W^{(2)} + \sigma_{ij}^{(1)}\varepsilon_{ij}^{(2)} \tag{6}$$

$$K_I = K_I^{(1)} + K_I^{(2)}, \quad K_{II} = K_{II}^{(1)} + K_{II}^{(2)}, \quad K_{III} = K_{III}^{(1)} + K_{III}^{(2)} \tag{7}$$

We can derive the following equation by substituting Eqs. (5), (6), (7) in Eq. (4).

$$\frac{2}{E'}\left(K_I^{(1)}K_I^{(2)} + K_{II}^{(1)}K_{II}^{(2)}\right) + \frac{1}{\mu}K_{III}^{(1)}K_{III}^{(2)} = -\int_\Omega \left\{\sigma_{kl}^{(1)}\varepsilon_{kl}^{(2)}\delta_{1i} - \sigma_{ij}^{(1)}\frac{\partial u_j^{(2)}}{\partial x_1} - \sigma_{ij}^{(2)}\frac{\partial u_j^{(1)}}{\partial x_1}\right\}\frac{\partial q}{\partial x_i}d\Omega \tag{8}$$

We let the equilibrium state (1) be the solution of the boundary value problem which is solved by the finite element method and the equilibrium state (2) be the asymptotic solution at the vicinity of the crack. We can set the displacements of the equilibrium (2) by using asymptotic solution of displacement near the crack front as described in text book of Broek (1986).

By setting the mode I, II or III stress intensity factor of the equilibrium state (2) be unity, we extract the respective stress intensity factor. For example, if we set $K_I^{(2)} = 1$ and $K_{II}^{(2)} = K_{III}^{(2)} = 0$,we extract $K_I^{(1)}$ of the equilibrium state (1).

The asymptotic solution of the crack with respect to the stress intensity factors can be expressed by using the local Cartesian coordinates (x_1', x_2', x_3') and cylindrical coordinates (r, θ, x_3') that are defined at the crack front as depicted in Fig. 4. The displacements can be written by:

$$u_i(r,\theta) = \sqrt{\frac{r}{2\pi}}\left\{K_I f_i^I(\theta) + K_{II} f_i^{II}(\theta) + K_{III} f_i^{III}(\theta)\right\} \tag{9}$$

If a node is above the crack face, a positive angle θ_u as shown in Fig. 4 is used in Eq. (9). If a node is below the crack face, a positive angle θ_d is used.
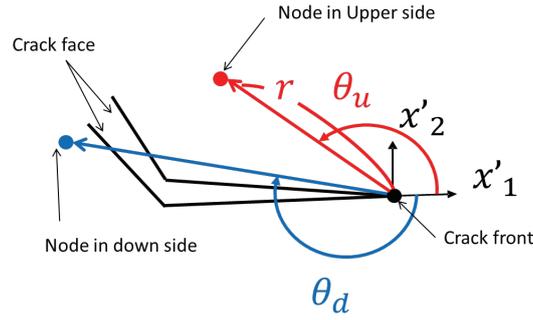
Figure 4. The local Cartesian and polar coordinate systems around the crack front

Correction terms to the asymptotic solutions (auxiliary solutions) adopted in the interaction integral method

The auxiliary solutions defined by the asymptotic solutions of crack do not satisfy the equilibrium in terms of the finite element. Therefore, a method to recover the equilibrium by adding correction terms to the asymptotic solution was proposed by Daimon et al. (2014). We write the new auxiliary solutions by the following equations.

$$\hat{u}_j^{(2)} = u_j^{(2)} + \hat{\hat{u}}_j^{(2)}, \quad \hat{\sigma}_j^{(2)} = \sigma_j^{(2)} + \hat{\hat{\sigma}}_j^{(2)} \tag{10}$$

Here, $\hat{\hat{u}}_j^{(2)}$, and $\hat{\hat{\sigma}}_j^{(2)}$ are the correction terms for the displacements and the stresses.

We write the weak form of equilibrium equations in terms of $\hat{\sigma}_j^{(2)}$, as:

$$\int w_j \left(\frac{\partial \sigma_{ij}^{(2)}}{\partial x_i} + \frac{\partial \hat{\hat{u}}_{ij}^{(2)}}{\partial x_i} \right) dV = 0 \tag{11}$$

Here, V is the domain of integral in the domain integral method, and w_j are the weight functions. We can derive a symmetric weak form by applying the Gauss divergence theorem and the generalized Hooke's law $\left(\hat{\hat{\sigma}}_{ij}^{(2)} = E_{ijkl} \frac{\partial \hat{\hat{u}}_k^{(2)}}{\partial x_l} \right)$, as:

$$\int \frac{\partial w_j}{\partial x_i} E_{ijkl} \frac{\partial \hat{\hat{u}}_k^{(2)}}{\partial x_l} dV = -\int \frac{\partial w_j}{\partial x_i} \sigma_{ij}^{(2)} dV \tag{12}$$

where E_{ijkl} are the components of fourth order tensor representing the generalized Hook's law.

We can establish a finite element formulation based on the symmetric weak form and determine the correction terms.

For cracks with kinks

If a crack involve kinks, we must account for the surface integral terms of Eq. (2). Because the crack surface and the crack propagation direction are not parallel to each other, $n_k q_k$ does not equal 0. [Okada et al. (2013)] proposed a method to vanish the surface integral terms in the domain

integral method by setting the value of q-function on the crack surface that is not parallel to crack propagation direction, as shown in Fig.5.
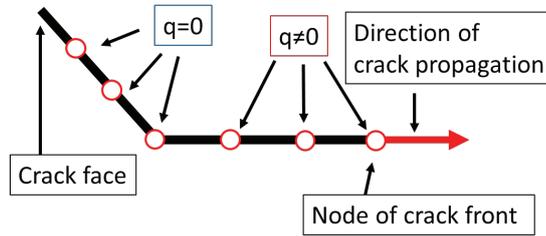
Figure 5. How to set the value of q

Numerical examples

In this section, some numerical results on the evaluation of the stress intensity factors of cracks involving kinks are presented. The stress intensity factors were also computed by VCCM proposed by Okada et al. (2008) and they are compared with those by present interaction integral method. The stress intensity factor evaluations by the interaction integral method were performed in two ways. One is to set the q-function value to be zero on the part of crack surfaces whose normal vector is not perpendicular to the direction of virtual crack extension. The other does not set q-function value to be zero on the crack surfaces.

The load conditions and the sizes of the tensile plate and the crack are as presented in Fig. 6 and Table 1. Uniformly distributed tractions are applied to the upper and lower surfaces of the plate. The crack has a kink as shown in Fig. 6. The kink angle θ is set to be 30 or 45 degrees.

The finite element analysis model is shown in Fig. 7 for the case of kink angle θ being 45 degrees. The total numbers of nodes and elements are 119152 and 82815, respectively. The size parameters of domain of integral are set to be $r_1/d_c = 64.0$, $h_0/d_c = 10.0$ and $h_1/d_c = 14.0$. The size d_c of finite element along the crack front is 0.0625 mm. The domain of integration includes the kink.

The distributions of the stress intensity factors along the crack front are presented in Figs. 8 and 9. In Figs. 8(a) and 9(a), the results of the conventional way, that does not set the q-function value to be zero at all, are presented. The results of proposed way are presented in Figs. 8(b) and 9(b). All the stress intensity factors are normalized by that of two dimensional straight edge crack subject to a remote tension (see, for example, Murakami et al. (1987)). The formula is written by:

$$K_{ref} = \sigma\sqrt{\pi a}\left\{1.12 - 0.231 \times (a/w) + 10.55 \times (a/w)^2 - 21.72 \times (a/w)^3 + 30.39 \times (a/w)^4\right\} \quad (13)$$

Here, σ, w and a in present problem are 1.0 (MPa), 20.0 (mm) and 10.0 (mm), respectively.

It is seen in Figs. 8 (a) and 9 (a) that the distributions of the stress intensity factors exhibit some oscillations. They are notable in the mode II and III components. The oscillations somewhat disappeared when the proposed method is applied to compute the stress intensity factors, as seen in Figs. 8 (b) and 9 (b).

Table 1. Sizes of the plate and crack

h_1	h_2	w	t
20.0[mm]	20.0[mm]	20.0[mm]	10.0[mm]
a$_1$	a$_2$	θ	σ
9.0[mm]	1.0[mm]	30,45°	1.0[MPa]
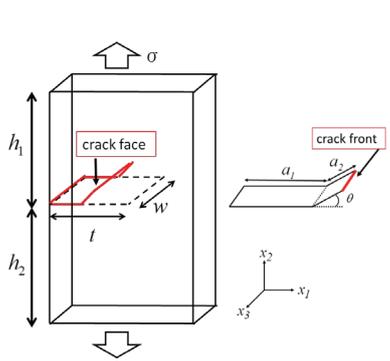
Figure 6. Plate with a through crack with a kink
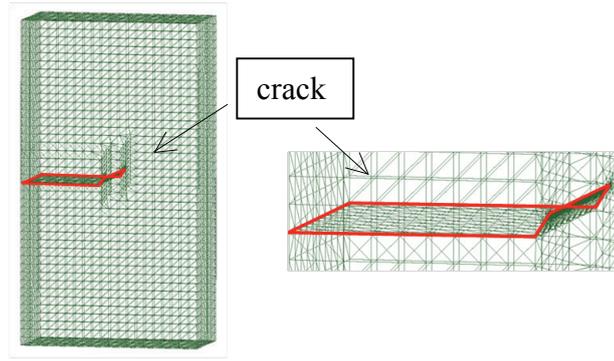
Figure 7. Finite element analysis model
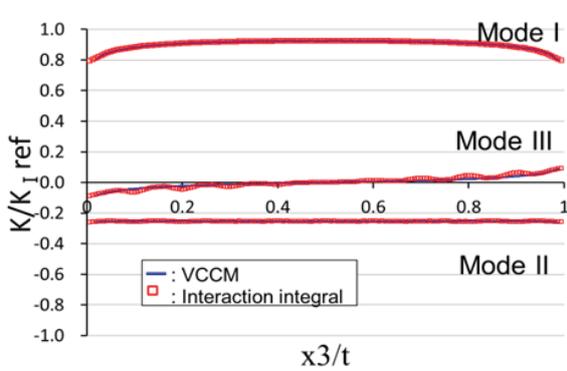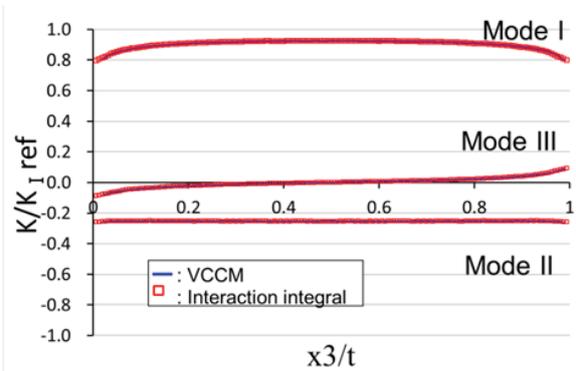
(a) Conventional way   **(b) Proposed way**

Figure 8. Distributions of the stress intensity factors for the case of kink angle (30˚)
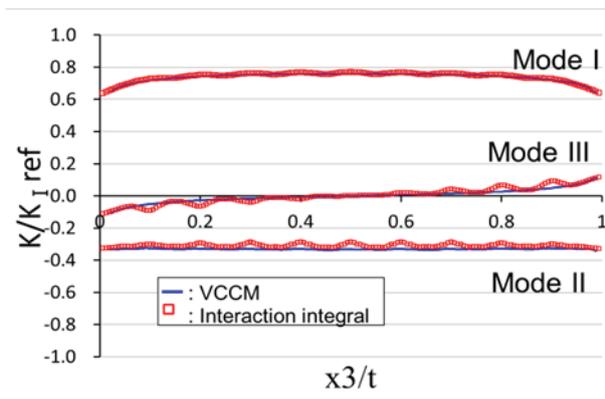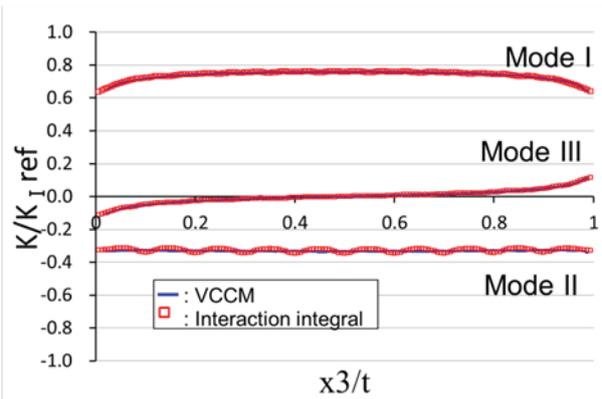
(a) Conventional way   **(b) Proposed way**

Figure 9. Distributions of the stress intensity factors for the case of kink angle (45˚)

Conclusions

The interaction integral method for tetrahedral finite element to accurately compute the stress intensity factors of mixed mode for cracks involving kinks is proposed in this paper. Comparing results of proposed method with those of conventional way, proposed method is found to serve more accurate and stable solutions than the conventional way.

References

Daimon, R. and Okada, H. (2014) Mixed-mode stress intensity factor evaluation by interaction integral method for quadratic tetrahedral finite Element with correction terms, *Engineering Fracture Mechanics* **115**, 22‑42.

Okada, H., Kawai,H. and Araki, K (2008) Virtual crack closure-integral method(VCCM) to compute the energy release rates and stress intensity factors based on quadratic tetrahedral finite elements, *Engineering Fracture Mechanics* **75**, 4466-4485, (2008).

Nikishkov, G. P. and Atluri, S.N. (1987) An equivalent domain integral method for computing crack-tip parameter in non-elastic, thermo-mechanical failure, *Engineering Fracture Mechanics* **26**, 851-867.

Rice. J. R. (1968) A path independent integral and the approximate analysis of strain concentration by notches and cracks, *Journal of Applied Mechanics* **35**, 379-386.

Shivakumar, K. N. and Raju, I. S. (1992) An equivalent domain integral method for three-dimensional mixed-mode fracture problems, *Engineering Fracture Mechanics* **42**, 935-959.

Okada, H. and Ohata, S. (2013) Three-dimensional J-integral evaluation for cracks with arbitrary curvatures and kinks based on domain integral method for quadratic tetrahedral finite element, *Engineering Fracture Mechanics* **109**, 58-77.

Yau, J , F., Wang, S, S. and Corten, H, T. (1980) A Mixed-mode crack analysis of isotropic solids using conservation laws of elasticity, *Journal of Applied Mechanics* 47, 335-341.

Broek, D (986) *Elementary engineering fracture mechanics*，Kluwer Academic Publishers.

Murakami Y, Aoki S, Hasebe N, Itoh Y, Mimata H, Miyazaki N, Terada H, Tohgo K, Toya M and Yuuki R. (1987) *Stress intensity factors handbook*, Committee on Fracture Mechanics, The Society of Materials Science, Japan.

Hyperelastic Fourth-Order Tensor Functions for Orthotropic Continua

David C. Kellermann and Mario M. Attard

School of Civil and Environmental Engineering, The University of New South Wales, Sydney, NSW 2052, Australia.

*Corresponding author: d.kellermann@unswalumni.com

Abstract

We have developed a general hyperelastic strain energy function for the modelling of orthotropic continua that is able to maintain the same logical properties of advanced isotopic hyperelastic constitutive laws. The isotropic model of Simo and Pister, due to its exceptional fit with experimental data and superior mathematical and logical features, is replicated in orthotropic form. Through the use of the proposed Intrinsic-Field Tensors yielding asymmetric strain tensors with additional degrees of freedom, orthotropic Lamé parameters where scalars are replaced by fourth-order tensors and advanced fourth-order tensor operators, the desired general form of orthotropic hyperelastic strain energy function is achieved.

Keywords: Orthotropic hyperelasticity, composites, additional degrees of freedom, intrinsic-field tensors, fourth-order tensors, Simo and Pister, strain energy function.

Introduction

Hyperelastic materials are a class of solids that can be modelled as continua with rate-independent strain energy defined purely as a function of deformation and the material parameters. The Simo and Pister model[1], like most hyperelastic Strain Energy Functions (SEFs), is restricted to isotropic materials; one of its particular benefits is that pure distortional deformation is independent of the volumetric modulus for finite strain, and that the volumetric strain is a logarithmic function of deformation. From a mathematical standpoint, the scalar strain energy function is expressed as the product of scalar deformation invariants and scalar coefficients.

In this paper we posit that there is a generalised form of the SEF that a large class of hyperelastic functions should be able to be written within, and that those that cannot *can* either be closely approximated by the general form or do not satisfy certain expected boundary conditions of finite strain hyperelasticity. This general form is an abstraction one level up of the classical SEF and is mathematically encompassing; we called it the Generalised Strain Energy (GSE).

After first demonstrating that there is an exact representation for the Simo and Pister strain energy within the GSE, we further revisit the class of orthotropic tensors that are asymmetric and of the form of Intrinsic-Field Tensors (IFTs)[2]. We also propose a natural separation of the extended form of the Hookean material tensor for stiffness, which is naturally extended for IFTs such that it utilises all free terms within a fourth order tensor having major symmetry. These are the orthotropic Lamé material tensors for stiffness and compliance.

These tools allow a new model for *orthotropic Simo and Pister hyperelasticity* that we purport to be the first of its kind and the only such model to inherit and maintain so may logical properties of isotropic hyperelasticity, structural tensors[3] and orthotropic material models simultaneously. Since the proposed model achieves these features by derivation and as pure theoretical development, the properties are ensured. Hence we do not, in this short paper, provide numerical examples or experimental correlations. The following section begins by analysing an alternate representation for the Simo and Pister model that is conducive to our subsequent transformations.

Isotropic Simo And Pister Model

Classical form of the isotropic Simo and Pister strain energy function

The isotropic hyperelastic model of Simo and Pister[1] is desirable due to various logical properties. Shown as follows,

$$\underset{S\&P}{W} = \tfrac{1}{2}\lambda\left(\ln J\right)^2 - G\ln J + \tfrac{1}{2}G\left(\operatorname{tr}\mathbf{b} - 3\right) = \tfrac{1}{2}\left(\lambda\left(\ln J\right)^2 - \mu\ln J + \mu\operatorname{tr}\underset{2}{\mathbf{E}}\right), \quad \text{where} \tag{1}$$

$$\mu = 2G \quad \text{and} \quad \operatorname{tr}\underset{2}{\mathbf{E}} = \tfrac{1}{2}\left(\operatorname{tr}\mathbf{C} - 3\right) = \tfrac{1}{2}\left(\operatorname{tr}\mathbf{b} - 3\right),$$

the derivative of this gives the Kirchhoff stress $\boldsymbol{\tau}$:

$$\boldsymbol{\tau} = \lambda\ln J\mathbf{I} + \mu\left(\mathbf{b} - \mathbf{I}\right). \tag{2}$$

The strain energy function W uses the Lamé parameters λ and μ in a scalar product with invariant components of the deformation/strain tensor. Here, $\ln J$ is the natural logarithm of J, the determinant of the stretch tensor $\mathbf{U}$ or similarly of the deformation gradient $\mathbf{F}$. Additionally, $\mathbf{b} = \mathbf{F}\mathbf{F}^{\mathrm{T}}$ is the left Cauchy–Green tensor, noting that $\operatorname{tr}\mathbf{b}$ is the trace function of $\mathbf{b}$, which is equal to $\operatorname{tr}\mathbf{C}$, where $\mathbf{C} = \mathbf{F}^{\mathrm{T}}\mathbf{F}$. Two particularly valued properties of the Simo and Pister model are:

  a) The deviatoric component of stress is only a function of μ

  b) The strain energy goes to infinity as either the volume goes to infinity or to zero (singularity)

Surprisingly few models meet criteria a) and b), which can easily be demonstrated. First, the deviatoric part of the stress measure $\mathbf{S}$ is

$$\boldsymbol{\tau}_{dev} = \boldsymbol{\tau} - \boldsymbol{\tau}_{vol}, \quad \text{where} \quad \boldsymbol{\tau}_{vol} = \tfrac{1}{3}\operatorname{tr}\left(\boldsymbol{\tau}\right)\mathbf{I} \tag{3}$$

Substituting Eq. (2), the volumetric part becomes

$$\boldsymbol{\tau}_{vol} = \lambda\ln J\mathbf{I} + \tfrac{1}{3}\mu\left(\operatorname{tr}\mathbf{b} - 3\right)\mathbf{I} \tag{4}$$

and, where $\mathbf{e}$ is the Almansi–Euler strain, the deviatoric part in Eq. (3) becomes

$$\boldsymbol{\tau}_{dev} = \lambda\left(\ln J\mathbf{I} - \ln J\mathbf{I}\right) + \mu\left(\mathbf{e} - \tfrac{1}{3}\operatorname{tr}\mathbf{e}\mathbf{I}\right) = \mu\left(\mathbf{e} - \tfrac{1}{3}\operatorname{tr}\mathbf{e}\mathbf{I}\right) \tag{5}$$

noting that this is independent of the parameter λ. The next property, that of infinite strain energy at zero volume, can simply be seen to follow the logarithm of zero, $\ln 0 = \infty$.

In this paper, we shall propose an orthotropic expansion of Simo and Pister's model that preserves these properties while also remaining a valid orthotropic continuum model that collapses down to the isotropic model by nothing more than material parameters becoming isotropic.

Transformation into standard scalar form using series strain

In order to elevate the form of the strain energy function in Eq. (1) we need to first turn the strain energy in to a standard form that is similar to the St Venant Kirchhoff model.
The first component of the function is simply transformed through the identity

$$\ln J = \ln\left(\det\mathbf{U}\right) = \operatorname{tr}\left(\ln\mathbf{U}\right) = \operatorname{tr}\underset{0}{\mathbf{E}}, \tag{6}$$

where $\underset{0}{\mathbf{E}}$ is the logarithmic strain following the Seth-Hill[4], [5] form of general strain:

$$\underset{n}{\mathbf{E}} = \tfrac{1}{n}\left(\mathbf{U}^n - \mathbf{I}\right) \quad, \quad \begin{cases} n > 0: & \underset{n}{\mathbf{E}} = \dfrac{1}{n}\left(\mathbf{U}^n - \mathbf{I}\right) \\ n = 0: & \underset{0}{\mathbf{E}} = \ln\left(\mathbf{U}\right) \end{cases} \tag{7}$$

This can be used to develop and interesting equality to replace $\underset{2}{\mathbf{E}}$ in Eq. (1). Initially we note

$$\underset{2}{\mathbf{E}} = \tfrac{1}{2}\left(\mathbf{U}^2 - \mathbf{I}\right), \quad \underset{1}{\mathbf{E}} = \mathbf{U} - \mathbf{I} \quad \rightarrow \quad 2\underset{2}{\mathbf{E}} + \mathbf{I} = \left(\underset{1}{\mathbf{E}} + \mathbf{I}\right)^2 \quad \rightarrow \quad \underset{2}{\mathbf{E}} = \tfrac{1}{2}\underset{1}{\mathbf{E}}^2 + \underset{1}{\mathbf{E}}, \tag{8}$$

which represents $\underset{2}{\mathbf{E}}$ in terms of $\underset{1}{\mathbf{E}}$. This process is repeated to the limit as n $\rightarrow \infty$, yielding

$$\underset{2}{\mathbf{E}} = \tfrac{1}{2}\underset{1}{\mathbf{E}}^2 + \tfrac{1}{4}\underset{1/2}{\mathbf{E}}^2 + \tfrac{1}{8}\underset{1/4}{\mathbf{E}}^2 + \tfrac{1}{16}\underset{1/8}{\mathbf{E}}^2 + \ldots + \tfrac{1}{\infty}\underset{1/\infty}{\mathbf{E}}^2 + \underset{1/\infty}{\mathbf{E}} = \tfrac{1}{2^1}\underset{2^0}{\mathbf{E}}^2 + \tfrac{1}{2^2}\underset{2^{-1}}{\mathbf{E}}^2 + \tfrac{1}{2^3}\underset{2^{-2}}{\mathbf{E}}^2 + \ldots + \tfrac{1}{\infty}\underset{0}{\mathbf{E}}^2 + \underset{0}{\mathbf{E}}$$

$$= \sum_{n=0}^{\infty}\left(\tfrac{1}{2^{n+1}}\underset{2^{-n}}{\mathbf{E}}^2\right) + \underset{0}{\mathbf{E}} \tag{9}$$

Now, suppose we define a strain measure called Series Strain $\underset{\Sigma}{\mathbf{E}}$, defined by

$$\underset{\Sigma}{\mathbf{E}}^2 = \sum_{n=0}^{\infty}\left(\tfrac{1}{2^{n+1}}\underset{2^{-n}}{\mathbf{E}}^2\right), \tag{10}$$

then we now have the identity

$$\underset{2}{\mathbf{E}} = \underset{\Sigma}{\mathbf{E}}^2 + \underset{0}{\mathbf{E}}. \tag{11}$$

Substitution of Eq. (6) and (11) into Eq. (1) gives

$$\underset{\text{S\&P}}{W} = \tfrac{1}{2}\left[\lambda\left(\operatorname{tr}\underset{0}{\mathbf{E}}\right)^2 - \mu\operatorname{tr}\underset{0}{\mathbf{E}} + \mu\operatorname{tr}\left(\underset{\Sigma}{\mathbf{E}}^2 + \underset{0}{\mathbf{E}}\right)\right] = \tfrac{1}{2}\left[\lambda\left(\operatorname{tr}\underset{0}{\mathbf{E}}\right)^2 + \mu\operatorname{tr}\underset{\Sigma}{\mathbf{E}}^2\right], \tag{12}$$

which is remarkably similar to the St Venant Kirchhoff (SVK) model:

$$\underset{\text{SVK}}{W} = \tfrac{1}{2}\left[\lambda\left(\operatorname{tr}\underset{2}{\mathbf{E}}\right)^2 + \mu\operatorname{tr}\underset{2}{\mathbf{E}}^2\right] \tag{13}$$

This form provides the basis for representation by the generalised strain energy function present in the section that follows.

Generalised Strain Energy (Gse)

GSE formulation

Given a fourth order tensor possessing major symmetry $\mathbb{B} = \mathbb{B}^{\mathrm{T}}$, we note the identity

$$\mathbf{A} : \mathbb{B} : \mathbf{C} = \mathbb{B} :: \left(\mathbf{A} \otimes \mathbf{C}\right), \tag{14}$$

where the operator $\otimes$ is the tensor product used by Itskov[6], $\left(\mathbf{A} \otimes \mathbf{B}\right)_{ijkl} = A_{ij}B_{kl}$. Using Eq. (14), we can represent the classical linear Hooke's Law

$$\tilde{W} = \tfrac{1}{2}\boldsymbol{\varepsilon} : \overline{\mathbb{C}} : \boldsymbol{\varepsilon} \quad \text{as} \quad \tilde{W} = \tfrac{1}{2}\overline{\mathbb{C}} :: \left(\boldsymbol{\varepsilon} \otimes \boldsymbol{\varepsilon}\right). \tag{15}$$

For a general representation of the model, we allow any order *n* of strain as per the Seth–Hill formula in Eq. (7), and so define a fourth-order tensor

$$\underset{n}{\mathbb{E}} = \underset{n}{\mathbf{E}} \otimes \underset{n}{\mathbf{E}}. \tag{16}$$

This yields a general model than can encompass the sum of any number of strain orders in consideration of the repeated indices on one side of the equation (Einstein summation convention):

$$W = \tfrac{1}{2}\underset{n}{\mathbb{C}} :: \underset{n}{\mathbb{E}}. \tag{17}$$

The capacity of this to represent various models will become apparent in the coming sections. Further to this, we can define a fourth-order *Strain Energy Tensor* (SET), which maintains identity to the components of strain energy:

$$\mathbb{W} = \tfrac{1}{2} \underset{n}{\mathbb{C}} \circ \underset{n}{\mathbb{E}} \tag{18}$$

This can be simply reduced back to the scalar value by summation of all elements of the tensor, as

$$W = \sum_{i,j,k,l} W_{ijkl} \,. \tag{19}$$

The form of Eq. (17) is not limited to a typical Hookean stiffness tensor– as mentioned, the summation index n refers to the order identifier of the general strain, but it also has a corresponding component of $\mathbb{C}$ such that $\mathbb{C} = \sum_{n} \underset{n}{\mathbb{C}}$, where in the isotropic form it is split up into two fourth-order tensors separating the Lamé parameters, i.e.

$$\boxed{W = \tfrac{1}{2}\left(\underset{m}{\mathbb{L}} :: \underset{m}{\mathbb{E}} + \underset{n}{\mathbb{G}} :: \underset{n}{\mathbb{E}} \right)} \,, \tag{20}$$

where

$$\mathbb{C} = \mathbb{L} + \mathbb{G} \,, \quad \underset{iso}{\mathbb{L}} = \lambda . \mathbf{I} \otimes \mathbf{I} \,, \quad \underset{iso}{\mathbb{G}} = \mu . \mathbf{I} \odot \mathbf{I} \tag{21}$$

Eq. (20) has the capacity to encompass a wide range of existing strain energy functions with no approximation. Essentially, it is the transformation of the function of strain energy from the scalar product of *scalar parameters* and *invariants of strain* into the quadruple contractions of *fourth-order material tensors* and *fourth-order strains*.

Simo and Pister in GSE form

With the development on the GSE in Eqs. (17) and (20) we can now easily transform the strain energy function of Simo and Pister, represented in a general scalar form in Eq. (12), simply by specifying the order of the strains:

$$W = \tfrac{1}{2}\left(\mathbb{L} :: \underset{0}{\mathbb{E}} + \mathbb{G} :: \underset{\Sigma}{\mathbb{E}} \right) \tag{22}$$

In the case of the series strain, the corresponding fourth-order tensor is defined as

$$\underset{\Sigma}{\mathbb{E}} = \sum_{n=0}^{\infty} \left(\tfrac{1}{2^{n+1}} \underset{2^{-n}}{\mathbf{E}} \otimes \underset{2^{-n}}{\mathbf{E}} \right). \tag{23}$$

Eq. (22) is an exact representation of Simo and Pister's isotropic hyperelastic model, though it is now in a form that is conducive to the introduction of direction dependence.

Intrinsic-Field Tensors: Strain

Deformation IFTs

Earlier in this we referred to the classical stretch tensor $\mathbf{U}$, we must now differentiate the stretch based on the material property-based domain. We shall introduce E to represent the domain of isotropic materials and the domain Œ to represent orthotropic materials. The property of symmetry is herein only afforded to the stretch tensor existing within the domain of isotropy, hence $\mathbf{U}^{\text{E}}$.

The well-known polar decomposition of the deformation gradient $\mathbf{F}$ into stretch and rotation $\mathbf{R}$ can be represented in isotropic parts

$$\mathbf{F} = \mathbf{R}^{\text{E}} \cdot \mathbf{U}^{\text{E}} \quad , \quad \mathbf{F}^{\text{T}} \cdot \mathbf{F} = \left(\mathbf{U}^{\text{E}} \right)^{\text{T}} \cdot \mathbf{U}^{\text{E}} \,, \tag{24}$$

where the former equation indicates the multiplicative decomposition, and the latter implies the unitary and orthogonal nature of $\mathbf{R}$. These equations are indeterminate – they have infinite solutions – and so in mechanics we impose a symmetry condition onto the stretch tensor as follows:

$$\left(\mathbf{U}^{\mathrm{E}}\right)^{\mathrm{T}} = \mathbf{U}^{\mathrm{E}} \quad \therefore \quad \mathbf{F}^{\mathrm{T}} \cdot \mathbf{F} = \left(\mathbf{U}^{\mathrm{E}}\right)^{2} \tag{25}$$

In the present method, we require a stretch whereby the condition of symmetry is removed, which it turns out is only necessary for orthotropic and anisotropic domains. Thus the stretch tensor $\mathbf{U}^{\text{Œ}}$ is potentially asymmetric. This has been published in the thesis by Kellermann[2]. Hence Eq. (24) remains similar

$$\mathbf{F} = \mathbf{R}^{\text{Œ}} \cdot \mathbf{U}^{\text{Œ}} \quad , \quad \mathbf{F}^{\mathrm{T}} \cdot \mathbf{F} = \left(\mathbf{U}^{\text{Œ}}\right)^{\mathrm{T}} \cdot \mathbf{U}^{\text{Œ}} , \tag{26}$$

while the enforced symmetry is replaced by dependence on the $\mathbf{R}^{\text{Œ}}$, the IFT rotation as a function of the Rodrigues Rotation Vector $\boldsymbol{\Omega}$, hence

$$\mathbf{U}^{\text{Œ}} = \left[\mathbf{R}^{\text{Œ}}(\boldsymbol{\Omega})\right]^{\mathrm{T}} \cdot \mathbf{F} . \tag{27}$$

We do not go into detail of the physical implications of IFTs here, though it should be noted that Eq. (27)is solved simply by minimisation of the strain energy function, the variables being the components of the Rodrigues vector. The result is an asymmetric stretch tensor such that $U^{\text{Œ}}{}_{ij} \neq U^{\text{Œ}}{}_{ji}$.

Generalised strain as and IFT

It follows from the Seth–Hill strain in Eq. (7) and the redefinition of stretch in Eq. (27) that we can define a new IFT form of generalised strain:

$$\underset{\mathrm{n}}{\text{Œ}} = \tfrac{1}{n} \left\lfloor \left(\mathbf{U}^{\text{Œ}}\right)^{n} - \mathbf{I} \right\rfloor \tag{28}$$

This measure is for the domain of orthotropic continua, and is not limited to positive integers, indeed negative values yield Eulerian measures; and, fractions to the limit of zero (the logarithmic strain as an IFT) are similarly useful.

Material Tensors For IFTs

Orthotropic Hookean tensors for IFTs

Since IFT theory differentiates between in-plane shear components[7], we require additional shear parameters in the sense that *xy* and *yx* properties become unique. This is quite a natural extension, as it simply means using the 9×9 stiffness matrix that follows from a 3×3×3×3 material tensor. The most compact form of such properties uses indicial notation, where the compliance material tensor $\mathbb{S}$ in the orthotropic orientation denoted by ${}^{\{\mathcal{M}\}}(\cdot)$ is expressed as

$$ {}^{\{\mathcal{M}\}}\mathbb{S}_{ijkl} = \left(\delta_{ik}\delta_{jl}\left(1 + v_{ji}\right) - \delta_{ij}\delta_{kl}v_{lj}\right)\Big/\underline{E}_{l} , \tag{29}$$

where δ_{ij} is the Kronecker delta, $\underline{E}_{i}$ are the components of the Young's Modulus vector and v_{ij} are the components of the Poisson Ratio matrix (see Reference [7]).
Representing the compliance tensor in flattened matrix form $\left[\mathbb{S}\right]$, it then is inverted as shown

$$\left[\mathbb{C}\right] = \left[\mathbb{S}\right]^{-1} \tag{30}$$

to yield the orthotropic Hookean material tensor $\underset{\text{orth}}{\mathbb{C}}$ for use with IFTs.

Orthotropic Lamé tensors for IFTs

Various previous efforts have proposed a set of "orthotropic Lamé parameters", though none meet a very simple requirement set out here:

a) Reduces to the two isotropic Lamé material tensors in Eq. (21) when properties are isotropic

b) Addition of each yields the Hookean orthotropic material tensor of Eqs. (29) and (30), ensuring consistent tangent stiffness

The resulting proposed orthotropic Lamé tensors are $\underset{\text{orth}}{\mathbb{L}}$ and $\underset{\text{orth}}{\mathbb{G}}$ corresponding to λ and μ in isotropy. These are given for both compliance and stiffness in two-dimension as follows.

$$
\left[\underset{\text{orth}}{\mathbb{L}}\right] = \begin{bmatrix} \frac{1}{2}\left(\nu_{23} + \nu_{32} + \nu_{23}{}^2 + \nu_{32}{}^2\right)\tilde{E}_1/\bar{\nu} & \left(\nu_{21} + \nu_{23}\nu_{31}\right)\tilde{E}_1/\bar{\nu} & 0 & 0 \\ \left(\nu_{12} + \nu_{13}\nu_{32}\right)\tilde{E}_2/\bar{\nu} & \frac{1}{2}\left(\nu_{31} + \nu_{13} + \nu_{31}{}^2 + \nu_{13}{}^2\right)\tilde{E}_2/\bar{\nu} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}
$$

$$
\left[\underset{\text{orth}}{\mathbb{G}}\right] = \begin{bmatrix} \left(2+\nu_{23}+\nu_{32}\right)\left(1-\nu_{23}-\nu_{32}\right)\tilde{E}_1/2\bar{\nu} & 0 & 0 & 0 \\ 0 & \left(2+\nu_{31}+\nu_{13}\right)\left(1-\nu_{31}-\nu_{13}\right)\tilde{E}_2/2\bar{\nu} & 0 & 0 \\ 0 & 0 & \tilde{E}_2/\left(1+\nu_{21}\right) & 0 \\ 0 & 0 & 0 & \tilde{E}_1/\left(1+\nu_{12}\right) \end{bmatrix} \quad (31)
$$

$$
\left[\underset{\text{orth}}{\mathbb{C}}\right] = \left[\underset{\text{orth}}{\mathbb{L}}\right] + \left[\underset{\text{orth}}{\mathbb{G}}\right] = \begin{bmatrix} \left(1-\nu_{23}\nu_{32}\right)\tilde{E}_1/\bar{\nu} & \left(\nu_{21}+\nu_{23}\nu_{31}\right)\tilde{E}_1/\bar{\nu} & 0 & 0 \\ \left(\nu_{12}+\nu_{13}\nu_{32}\right)\tilde{E}_2/\bar{\nu} & \left(1-\nu_{31}\nu_{13}\right)\tilde{E}_2/\bar{\nu} & 0 & 0 \\ 0 & 0 & \tilde{E}_2/\left(1+\nu_{21}\right) & 0 \\ 0 & 0 & 0 & \tilde{E}_1/\left(1+\nu_{12}\right) \end{bmatrix}
$$

where $\bar{\nu} = 1 - \nu_{12}\nu_{21} - \nu_{23}\nu_{32} - \nu_{31}\nu_{13} - \nu_{12}\nu_{23}\nu_{31} - \nu_{21}\nu_{32}\nu_{13}$.

Orthotropic Simo And Pister Model

Orthotropic Simo and Pister using GSE

Having shown the Simo and Pister model in the form of the GSE in Eq. (22), having presented the equivalent intrinsic-field tensors for the strains in orthotropy in Eq. (28) and having given the orthotropic equivalent of the fourth-order Lamé tensors in Box 1 we are able to convert Simo and Pister isotropic hyperelasticity into a fully logically-compliant hyperelastic model. This is achieved through the trivial step of replacing $\underset{\text{iso}}{\mathbb{L}}$ with $\underset{\text{orth}}{\mathbb{L}}$, $\underset{\text{iso}}{\mathbb{G}}$ with $\underset{\text{orth}}{\mathbb{G}}$ and $\underset{n}{\mathbb{E}} = f\left(\underset{n}{\mathbf{E}}\right)$ with $\underset{n}{\mathbb{E}} = f\left(\underset{n}{\mathbf{Œ}}\right)$.

The resulting formula, expressed entirely as fourth-order tensors, is

$$
\underset{\text{orth}}{W} = \tfrac{1}{2}\left(\mathbb{L} :: \underset{0}{\mathbb{E}} + \mathbb{G} :: \underset{\Sigma}{\mathbb{E}}\right) \quad , \quad \underset{n}{\mathbb{E}} = \underset{n}{\mathbf{Œ}} \otimes \underset{n}{\mathbf{Œ}} \quad (32)
$$

with the more familiar form as follows. Orthotropic Simo and Pister Hyperelasticity:

$$
\boxed{\underset{\text{orth}}{W} = \tfrac{1}{2}\left(\underset{0}{\mathbf{Œ}} : \mathbb{L} : \underset{0}{\mathbf{Œ}} + \underset{\Sigma}{\mathbf{Œ}} : \mathbb{G} : \underset{\Sigma}{\mathbf{Œ}}\right)} \quad (33)
$$

In the next section we will complete the development of the equation by demonstrating that it has the correct tangent stiffness.

Linearisation back to Hooke's law

Finally we can demonstrate that the proposed model reduces back to orthotropic Hooke's law for IFTs, which has been shown to have identical strain energy to classical orthotropic Hooke's law. Physically, this also shows that the tangent stiffness of the proposed orthotropic hyperelastic model is consistent with classical elasticity. As deformation gradient gets very close to the identity tensor, all strain measures linearise to the infinitesimal strain measure of Cauchy, though in the case of IFTs it is asymmetric:

$$\text{as } \mathbf{F} \to \mathbf{I}, \quad \underset{n}{\mathbf{E}} \to \underset{n}{\tilde{\mathbf{E}}} = \tilde{\mathbf{E}} = \boldsymbol{\varepsilon}, \quad \underset{n}{\textrm{Œ}} \to \underset{n}{\tilde{\textrm{Œ}}} = \tilde{\textrm{Œ}} \tag{34}$$

Thus equation (33) can be factored by the identical linear strain measures as

$$\underset{\text{orth}}{\tilde{W}} = \tfrac{1}{2} \tilde{\textrm{Œ}} : \left(\mathbb{L} + \mathbb{G} \right) : \tilde{\textrm{Œ}} \tag{35}$$

and from Box 1 we know that the two orthotropic Lamé tensor combine to give the extended orthotropic Hookean material tensor $\underset{\text{orth}}{\mathbb{C}} = \underset{\text{orth}}{\mathbb{L}} + \underset{\text{orth}}{\mathbb{G}}$. Hence Eq. (35) returns to the familiar form

$$\begin{aligned} \underset{\text{orth}}{\tilde{W}} &= \tfrac{1}{2} \tilde{\textrm{Œ}} : \underset{\text{orth}}{\mathbb{C}} : \tilde{\textrm{Œ}} \\ &= \tfrac{1}{2} \boldsymbol{\varepsilon} : \underset{\text{orth}}{\overline{\mathbb{C}}} : \boldsymbol{\varepsilon} \quad \text{Orthotropic Hooke's Law} \end{aligned}, \tag{36}$$

where $\underset{\text{orth}}{\overline{\mathbb{C}}}$ is the classical orthotropic Hookean material tensor for stiffness. The proof of the equality between lines in Eq. (36) is obtained by using a mixing equation to generate the classical 'combined' in-plane shear moduli and then finding that the energies are always identical since

$$\tfrac{1}{2} \left(\tilde{Œ}_{ij} + \tilde{Œ}_{ji} \right) = \varepsilon_{ij}. \tag{37}$$

Thus classical tangent stiffness is guaranteed in the proposed model, and for that matter, any orthotropic hyperelastic model of the form of Eq. (20).

Conclusion

In this paper a new class of hyperelastic, orthotropic strain energy functions is introduced by way of demonstrating the conversion of the well-known Simo and Pister model. This is done by first elevating the model from being *isotropic & hyperelastic* to being *orthotropic & hyperelastic*, and then reducing the *orthotropic & hyperelastic* model to being *orthotropic & infinitesimal*. Both the start point (Simo and Pister's model) and the end points (Hookean infinitesimal orthotropy) are widely accepted models, and no approximations are made from the transition from one to the other. The resulting midpoint, the hyperelastic, orthotropic Simo and Pister model, maintains all the desirable qualities of its isotropic counterpart and of Hookean orthotropy. This alone should serve as a compelling argument for the introduction of intrinsic-field tensors and the greater proposed theory of *Orthotropic Continuum Mechanics* into the domain of contemporary continuum mechanics at large. This is by no means a specialised theory – its ability to encompass and adapt to a wide range of applications should be evident through the mathematics alone.

References

[1] J. C. Simo and K. S. Pister, "Remarks on rate constitutive equations for finite deformation problems: computational implications," *Computer Methods in Applied Mechanics and Engineering*, vol. 46, no. 2, pp. 201–215, 1984.
[2] D. C. Kellermann, "A theory of Strongly orthotropic continuum mechanics," The University of New South Wales, 2008.
[3] S. Klinkel, C. Sansour, and W. Wagner, "An anisotropic fibre-matrix material model at finite elastic-plastic strains," *Computational Mechanics*, vol. 35, no. 6, pp. 409–417, May 2005.
[4] R. Hill, "On constitutive inequalities for simple materials--II," *Journal of the Mechanics and Physics of Solids*, vol. 16, no. 5, pp. 315–322, Aug. 1968.
[5] B. R. Seth, "Generalized strain measure with applications to physical problems," *Second-order effects in elasticity, plasticity and fluid dynamics*, pp. 162–172, 1964.
[6] M. Itskov and A. E. Ehret, "A Universal Model for the Elastic, Inelastic and Active Behaviour of Soft Biological Tissues," *GAMM-Mitteilungen*, vol. 32, no. 2, pp. 221–236, 2009.
[7] D. C. Kellermann, T. Furukawa, and D. W. Kelly, "Strongly orthotropic continuum mechanics and finite element treatment," *International Journal for Numerical Methods in Engineering*, vol. 76, no. 12, pp. 1840–1868, 2008.

Convolution Quadrature BEM for Wave Analysis in General Anisotropic

Fluid-Saturated Porous Solid and its GPU Acceleration

*A. Furukawa[1], T. Saitoh[2] and S. Hirose[1]

[1]Department of Mechanical and Environmental Informatics, Tokyo Institute of Technology,
2-12-1 Ookayama, Meguro-ku, Tokyo 152-8552, Japan.
[2]Division of Environmental Engineering Science, Gunma University, Japan.

*Corresponding author: furukawa.a.aa@m.titech.ac.jp

Abstract

Rocks under the ground include pores and cracks which produce strong anisotropy. In addition, the pores and cracks are saturated with pore fluid. Therefore, the rocks can be considered as a general anisotropic fluid-saturated porous solid. Boundary element method (BEM) is well known as an effective numerical approach for wave scattering analysis because BEM can treat infinite or semi-infinite domains without any modification. Convolution quadrature boundary element method (CQ-BEM), which uses convolution quadrature method (CQM) for time-discretization, is known as a new time-domain BEM. The CQ-BEM requires Laplace-domain fundamental solutions and can produce stable numerical solutions, even for small time increments. In this research, a CQ-BEM for wave analysis in general anisotropic fluid-saturated porous solids is developed and accelerated by GPU parallelization. The proposed method and the effect of GPU acceleration are verified by solving wave scattering problems by a cavity.

Keywords: Boundary element method, Convolution quadrature method, Anisotropy, Pore fluid, GPU acceleration

Introduction

Wave analysis in rocks has been addressed in earthquake engineering and exploration geophysics. Waves propagating in rocks, especially near the fault fracture zones and reservoirs, have some important properties: anisotropy, dispersion, and attenuation. Anisotropy is generated by crystal preferred orientation of the rocks and aligned cracks in the rocks. On the other hand, dispersion and attenuation result from porous structure of the rocks and existence of pore fluid. Therefore, the rocks can be considered as a general anisotropic fluid-saturated porous solid. A theory of fluid-saturated porous solids has been proposed by [Biot (1956a; 1956b)], and the theory including consideration of general anisotropy was subsequently presented by [Biot (1962)]. The mechanical model proposed by Biot is called Biot's model, and there are various expanded models which have been derived from Biot's original one.

Boundary element method (BEM) is an effective numerical approach for wave analysis. This is because BEM requires boundary discretization only and can treat infinite and semi-infinite domain without any modification. Therefore, BEM can provide accurate numerical solutions compared with other computational methods such as finite element method (FEM), and finite difference method (FDM), which require spatial discretization of analysis zone. In recent years, a novel time-domain BEM, convolution quadrature boundary element method (CQ-BEM), has been proposed by [Schanz, M. and Antes, H. (1997)]. CQ-BEM is the BEM that a convolution quadrature method (CQM) is used for time-discretization. CQM evaluates the Riemann convolution $f(t) * g(t)$ numerically, and requires Laplace transform of the function $f(t)$. Therefore, in CQ-BEM formulation, time-domain boundary integral equations (BIEs) are solved using Laplace-domain fundamental solutions. Main advantages of CQ-BEM are to produce stable numerical solutions with small time increments, and to deal with waves affected by dispersion and dissipation.

Boundary element method for dynamic poroelastic problems using Biot's model has been presented by a number of researchers. A frequency-domain BEM was proposed by [Domínguez (1992)]. In
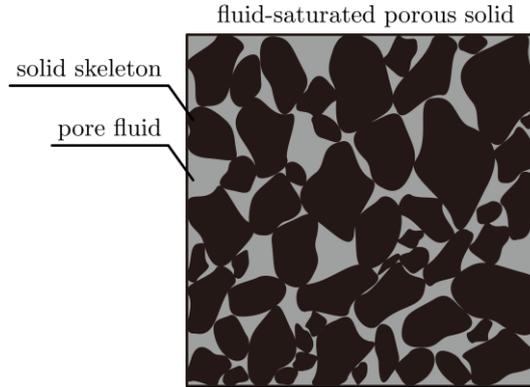
Figure 1. Fluid-saturated porous solid.

this research, the frequency-domain BIEs were derived from the analogy between dynamic poroelasticity and thermoelasticity and, described by two kinds of boundary values, i.e., solid displacements, and fluid pressure. On the other hand, time-domain BEMs for this kind of problem have been developed with the advent of the CQ-BEM [Schanz (2001); Saitoh et al. (2012)]. These formulations involved time-domain BIEs with respect to solid displacements and fluid pressure as boundary values. However, these previous BEM formulations can be applied to isotropic fluid-saturated porous solids only, and the research on BEM for wave analysis in general anisotropic fluid-saturated porous solids has not been carried out, as long as the authors know.

In this paper, a CQ-BEM for two-dimensional wave scattering problems in general anisotropic fluid-saturated porous solids is developed, and the validity of our presented method is confirmed. In general, BEM considering general anisotropy requires much computational time because the computation of fundamental solutions involves integration over the unit circle in two-dimensional problems. Therefore, GPU parallelization is applied to the proposed CQ-BEM and effect of GPU parallelization is confirmed. In the following sections, the theory of Biot's model is summarized and CQ-BEM formulation for Biot's model is subsequently presented. After these expressions, numerical examples and effect of GPU parallelization are illustrated.

Biot's Model for General Anisotropic Fluid-Saturated Porous Solids

In this section, Biot's model, a mechanical model for general anisotropic fluid-saturated porous solids proposed by [Biot (1962)], is described. The presented model shown here involves effects of anisotropy. The Biot's model is based on the following assumptions:
1. Fluid-saturated porous solid consists of solid skeleton and pore fluid as shown in Fig. 1.
2. Infinitesimal transformations occur between the reference and current states of deformation.
3. The wavelength is large compared with the dimensions of macroscopic values.
4. The conditions are isothermal.
5. The fluid is viscous.
6. The fluid flows through the porous skeleton according to Darcy's law.

In this model, anisotropy is due to a preferential alignment of the pores or cracks. Therefore, anisotropic effects of both elasticity of the solid skeleton and permeability of pore fluid can be considered.

The linear stress-strain relations of the Biot's model are given by the following equations:

$$\sigma_{ij} = A_{ijkl}e_{kl} + \alpha_{ij}M\zeta, \tag{1}$$

$$p = \alpha_{kl}Me_{kl} + M\zeta \tag{2}$$

where σ_{ij} and p represents the total stress of the porous solid and the fluid pressure, respectively. e_{ij} is the strain components of the porous solid, and ζ is the increment of the fluid content. In addition, A_{ijkl} represents the elastic tensor of undrained porous solid. M is Biot's elastic modulus, and α_{ij} is

Biot's effective-stress coefficients in general anisotropic case. In Eqs. (1) and (2), the strain component e_{ij} is expressed, using the displacement components of the porous solid u_i, as follows:

$$e_{ij} = \frac{1}{2}\left(u_{i,j} + u_{j,i}\right) \tag{3}$$

where $( )_{,i} = \partial/\partial x_i$. Moreover, the elastic tensor of undrained porous solid A_{ijkl} is expressed as follows:

$$A_{ijkl} = C_{ijkl} + \alpha_{ij}\alpha_{kl}M \tag{4}$$

where C_{ijkl} represents the elastic tensor of drained porous solid (solid skeleton).

Equations of motion of the Biot's model are expressed as follows:

$$\sigma_{ij,j} + \rho b_i = \rho \ddot{u}_i + \rho_f \ddot{w}_i, \tag{5}$$

$$p_{,i} + \rho_f c_i = -\rho_f \ddot{u}_i + m_{ij}\ddot{w}_j - \eta r_{ij}\dot{w}_j \tag{6}$$

where $(\dot{}) = \partial/\partial t$, and w_i represents the flow of the fluid relative to the solid in the unit section. ρ and ρ_f are the densities of the porous solid and pore fluid. b_i and c_i represent body force components of the solid and the pore fluid, respectively. In addition, η is the viscosity of the fluid. m_{ij} is the mass matrix which depends on the pore geometry. Moreover, r_{ij} represents the flow resistivity matrix which is the inverse matrix of the permeability matrix based on Darcy's law. Note that the third term of the right hand side in Eq. (6) is dissipation term. Therefore, dissipation depends on the viscosity of the pore fluid and the relative motion between the solid and the fluid.

Characteristics of wave propagation in general anisotropic fluid-saturated porous solids have been studied by several researchers [Carcione (1996); Sharma (2005)]. According to these previous studies, anisotropic fluid-saturated porous solids generate four body waves; i.e., quasi-fast longitudinal wave (qP1), quasi-slow longitudinal wave (qP2), and quasi-transverse waves (qS1 and qS2). In addition, phase velocities of the waves depend on the propagation direction and the frequency. The viscosity η effects great changes on the behavior of qP2 wave.

Convolution Quadrature Boundary Element Method for Biot's Model

In this section, a formulation of convolution quadrature boundary element method (CQ-BEM) for wave scattering problems in general anisotropic fluid-saturated porous solids is presented. Our formulation is based on that proposed by [Saitoh et al. (2012)]. Firstly, the BEM formulation based on time-domain boundary integral equations (BIEs) is described. Secondly, time- and spatial discretization of the BIEs are expressed. Finally, Laplace-domain fundamental solutions for the target problems are illustrated.

CQ-BEM Formulation for Biot's Model

Formulation based on time-domain BIEs is described here. Considering infinite domain D and its boundary S, time-domain BIEs for Biot's model are expressed as follows:

$$C(\mathbf{x})q_I(\mathbf{x},t) = q_I^{\text{in}}(\mathbf{x},t) + \int_S U_{IK}(\mathbf{x},\mathbf{y},t) * s_K(\mathbf{y},t)dS(\mathbf{y})$$
$$- \int_S W_{IK}(\mathbf{x},\mathbf{y},t) * q_K(\mathbf{y},t)dS(\mathbf{y}) \tag{7}$$

where $*$ denotes the Riemann convolution. For two-dimensional case, the subscript written by capital letter ranges from 1 to 3. In addition, $q_I(\mathbf{x}, t)$ and $s_I(\mathbf{x}, t)$ represent generalized displacement and traction at point $\mathbf{x}$ at time t, respectively, shown as follows:

$$q_I = \{u_1, u_2, p\}^{\mathrm{T}}, \quad s_I = \{t_1, t_2, p_n\}^{\mathrm{T}} \tag{8}$$

where t_i $(i = 1, 2)$ is the traction component of the solid, and p_n is normal derivative of the pressure given by $p_n = \partial p / \partial n$. In addition, $q_I^{\mathrm{in}}(\mathbf{x}, t)$ represents generalized displacement for the incident wave, and the free term $C(\mathbf{x})$ is given by

$$C(\mathbf{x}) = \begin{cases} 1 & : \mathbf{x} \in D \\ 1/2 & : \mathbf{x} \in S \\ 0 & : \mathbf{x} \in D_c \end{cases} \tag{9}$$

where D_c denotes complementary domain of the analytical domain D. In Eq. (7), $U_{IK}(\mathbf{x}, \mathbf{y}, t)$ and $W_{IK}(\mathbf{x}, \mathbf{y}, t)$ represent time-domain fundamental solutions and its double layer kernels, respectively.

Time- and Spatial Discretization of BIEs

Algebraic equation which can be solved computationally is derived via time- and spatial discretization of the time-domain BIEs shown in Eq. (7). Time- and spatial discretization is applied by the following manners: Convolution quadrature method (CQM) based on the backward difference formula is used for time-discretization. Collocation method with constant shape function is used for spatial discretization. Approximation formula of the CQM is given by the following equations:

$$f(n\Delta t) * g(n\Delta t) = \sum_{k=0}^{n} \omega_{n-k}(\Delta t)\, g(n\Delta t) \tag{10}$$

where Δt represents time increment, and ω_m is the weight function as follow:

$$\omega_m(\Delta t) = \frac{1}{2\pi \mathrm{i}} \int_{|z|=R} \hat{f}\left(\frac{\gamma(z)}{\Delta t}\right) z^{-n-1} dz \tag{11}$$

where i is the imaginary unit, and $\hat{f}$ denotes the Laplace transform of f. In the numerical computation, the integration in Eq. (11) is evaluated by L-point trapezoidal rule. Therefore, the weight function ω_m is expressed as follows:

$$\omega_m(\Delta t) \approx \frac{R^{-m}}{L} \sum_{l=0}^{L-1} \hat{f}\left(\frac{\gamma(z_l)}{\Delta t}\right) \mathrm{e}^{-\frac{2\pi \mathrm{i} m l}{L}}. \tag{12}$$

The CQM parameters $\gamma(z_l)$ and R, with an error magnitude $O(\epsilon)$, are written as follows:

$$\gamma(z_l) = \sum_{i=1}^{k} \frac{1}{i}(1 - z_l)^i, \quad z_l = R\mathrm{e}^{-2\pi \mathrm{i}\frac{l}{L}}, \quad R = \epsilon^{\frac{1}{2L}}. \tag{13}$$

Note that Eq. (13) corresponds to the backward differentiation formulas of order k. Using the CQM formula and taking the limit process $\mathbf{x} \in D \to \mathbf{x} \in S$, the following discretized BIEs are obtained:

$$\left[\frac{1}{2}\delta_{MN}\delta_{IK} + B_{MN;IK}^{(0)}\right] q_{N;K}^{(n)} - A_{MN;IK}^{(0)} s_{N;K}^{(n)} = q_{M;I}^{\mathrm{in}(n)}$$

$$+ \sum_{k=1}^{n-1} \left[A_{MN;IK}^{(n-k)} s_{N;K}^{(k)} - B_{MN;IK}^{(n-k)} q_{N;K}^{(k)}\right], (n = 1, 2, \cdots, N_t). \tag{14}$$

In Eq. (14), $M, N = 1, 2, \cdots, N_e$ where N_e represents the number of boundary elements. Equation (14) describes the BIEs at n-th time step, and N_t is the total number of time step. Moreover, $A_{MN;IK}^{(k)}$ and $B_{MN;IK}^{(k)}$ are the influence functions as follows:

$$A_{MN;IK}^{(m)} = \frac{R^{-m}}{L} \sum_{l=0}^{L-1} \left[\int_{S^N} \widehat{U}_{IK}(\mathbf{x}^M, \mathbf{y}^N, s_l) dS(\mathbf{y}) \right] e^{-2\pi i m \frac{l}{L}} \tag{15}$$

$$B_{MN;IK}^{(m)} = \frac{R^{-m}}{L} \sum_{l=0}^{L-1} \left[\iint_{S^N} \widehat{W}_{IK}(\mathbf{x}^M, \mathbf{y}^N, s_l) dS(\mathbf{y}) \right] e^{-2\pi i m \frac{l}{L}} \tag{16}$$

where $s_l = \gamma(z_l)/\Delta t$, and $\widehat{U}_{IK}(\mathbf{x}, \mathbf{y}, s)$ and $\widehat{W}_{IK}(\mathbf{x}, \mathbf{y}, s)$ are the Laplace-domain fundamental solutions and its double layer kernels, respectively. Substituting prescribed boundary conditions into Eq. (14) and solving the resulting algebraic equation, unknown boundary values can be obtained.

Laplace-Domain Fundamental Solutions

Laplace-domain fundamental solutions are required for computation of the influence functions, and given by the following equation:

$$\hat{L}_{PQ}\widehat{U}_{QK}(\mathbf{x}, \mathbf{y}, s) = -\delta(\mathbf{x} - \mathbf{y})\delta_{PK} \tag{17}$$

where $\delta(\cdot)$ is Dirac delta function, δ_{IK} is Kronecker delta, and

$$\begin{aligned}
\hat{L}_{ik} &= C_{ijkl}\frac{\partial}{\partial x_j}\frac{\partial}{\partial x_l} - \tilde{\rho}_{ik}s^2, \quad \hat{L}_{i4} = -\tilde{\alpha}_{ij}\frac{\partial}{\partial x_j}, \\
\hat{L}_{4k} &= \tilde{\alpha}_{kl}\frac{\partial}{\partial x_l}, \quad \hat{L}_{44} = -\frac{1}{s^2}Y_{jl}^{-1}\frac{\partial}{\partial x_j}\frac{\partial}{\partial x_l} - \frac{1}{M}, \\
\tilde{\alpha}_{ik} &= \alpha_{ik} - \rho_f Y_{ik}^{-1}, \quad \tilde{\rho}_{ik} = \rho\delta_{ik} - \rho_f^2 Y_{ik}^{-1}.
\end{aligned} \tag{18}$$

Solving Eq. (17) using Radon and Fourier transform, the following fundamental solutions are obtained:

$$\widehat{U}_{ik}(\mathbf{x}, \mathbf{y}, s) = \frac{1}{4\pi^2}\int_{|\mathbf{p}|=1}\sum_{\alpha=1}^{4}\Lambda_{ik}^{\alpha}(\mathbf{p})\Phi^{\alpha}dL(\mathbf{p}),$$

$$\widehat{U}_{i3}(\mathbf{x}, \mathbf{y}, s) = \frac{1}{4\pi^2}\int_{|\mathbf{p}|=1}\sum_{\alpha=1}^{4}\Lambda_{i4}^{\alpha}(\mathbf{p})\Psi^{\alpha}dL(\mathbf{p}),$$

$$\widehat{U}_{3k}(\mathbf{x}, \mathbf{y}, s) = \frac{1}{4\pi^2}\int_{|\mathbf{p}|=1}\sum_{\alpha=1}^{4}\Lambda_{4k}^{\alpha}(\mathbf{p})\Psi^{\alpha}dL(\mathbf{p}),$$

$$\widehat{U}_{33}(\mathbf{x}, \mathbf{y}, s) = \frac{1}{4\pi^2}\int_{|\mathbf{p}|=1}\sum_{\alpha=1}^{4}\Lambda_{44}^{\alpha}(\mathbf{p})\Phi^{\alpha}dL(\mathbf{p})$$

$$\tag{19}$$

where
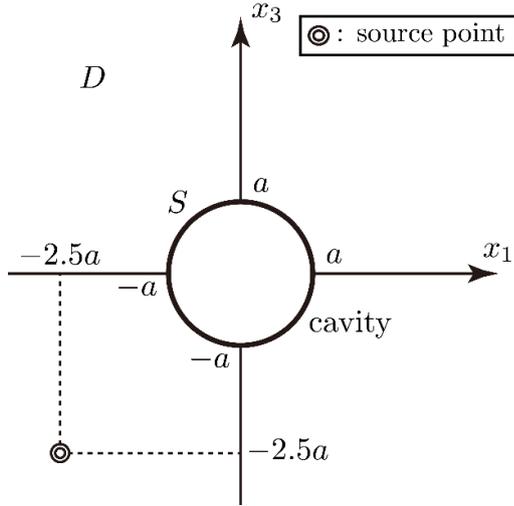
Figure 2. Analytical model for wave scattering analysis by a cavity.
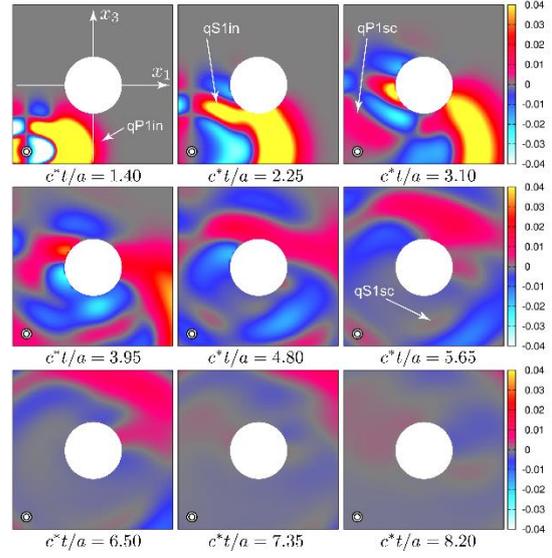
Figure 3. Time histories of the displacement u_1 around a cavity at several times.

$$\Lambda_{IK}^{\alpha}(\mathbf{p}) = k_{\alpha} \frac{\left(\text{adj}\left[\tilde{\check{\mathbf{L}}}\right]\Big|_{k=k_{\alpha}}\right)_{IK}}{\frac{\partial}{\partial k} \det\left[\tilde{\check{\mathbf{L}}}\right]\Big|_{k=k_{\alpha}}}, \tag{20}$$

$$\Phi^{\alpha} = e^{-ik_{\alpha}|\mathbf{p}\cdot\mathbf{r}|} E_1(-ik_{\alpha}|\mathbf{p}\cdot\mathbf{r}|) + e^{ik_{\alpha}|\mathbf{p}\cdot\mathbf{r}|}\{E_1(ik_{\alpha}|\mathbf{p}\cdot\mathbf{r}|) + i\pi\}, \tag{21}$$

$$\Psi^{\alpha} = \text{sgn}(\mathbf{p}\cdot\mathbf{r})[-e^{-ik_{\alpha}|\mathbf{p}\cdot\mathbf{r}|} E_1(-ik_{\alpha}|\mathbf{p}\cdot\mathbf{r}|) + e^{ik_{\alpha}|\mathbf{p}\cdot\mathbf{r}|}\{E_1(ik_{\alpha}|\mathbf{p}\cdot\mathbf{r}|) + i\pi\}]. \tag{22}$$

In Eq. (20), $\widetilde{(\ )}$ and $\check{(\ )}$ denotes Radon and Fourier transform. Moreover, in Eqs. (21) and (22), $E_1(\cdot)$ is the exponential integral, and $\mathbf{r} = \mathbf{x} - \mathbf{y}$. k_{α} is obtained by solving the following equation:

$$\det\left[\tilde{\check{\mathbf{L}}}\right] = C_1 k_{\alpha}^8 + C_2 k_{\alpha}^6 + C_3 k_{\alpha}^4 + C_4 k_{\alpha}^2 + C_5 = 0, \quad \text{Im}[k_{\alpha}] > 0. \tag{23}$$

where $C_1, C_2, \cdots, C_5$ are coefficients.

In Eqs. (19), the subscript written by small letter ranges from 1 to 2, and the integrals over the unit circle ($|\mathbf{p}| = 1$) are implemented as a consideration of the interference of the body waves propagating in every direction. Moreover, the summation with respect to α denotes the superposition of the four body waves. In numerical computation, integrals over the unit circle are evaluated using double exponential formula [Takahasi and Mori (1974)]. The number of sampling points of these integrals is set to 1,400 in this research. Therefore, the computational cost is quite high. On the other hand, the double layer kernels $\widehat{W}_{IK}(\mathbf{x}, \mathbf{y}, s)$ are given by the following equations:

$$\widehat{W}_{IK}(\mathbf{x}, \mathbf{y}, s) = \hat{B}_{KJ}^{\mathbf{y}} \widehat{U}_{IJ}(\mathbf{x}, \mathbf{y}, s) \tag{24}$$

where

$$\hat{B}_{ik}^{\mathbf{y}} = -C_{ijkl} n_j(\mathbf{y}) \frac{\partial}{\partial x_l}, \quad \hat{B}_{i3}^{\mathbf{y}} = \frac{M_{ij}}{M} n_j(\mathbf{y}),$$

$$\hat{B}_{3k}^{\mathbf{y}} = -\rho_f Y_{kl}^{-1} n_l(\mathbf{y}), \quad \hat{B}_{33}^{\mathbf{y}} = \frac{1}{s^2} Y_{jl}^{-1} n_j(\mathbf{y}) \frac{\partial}{\partial x_l}, \tag{25}$$

$$Y_{ik} = m_{ik} + \frac{\eta}{s} r_{ik}.$$
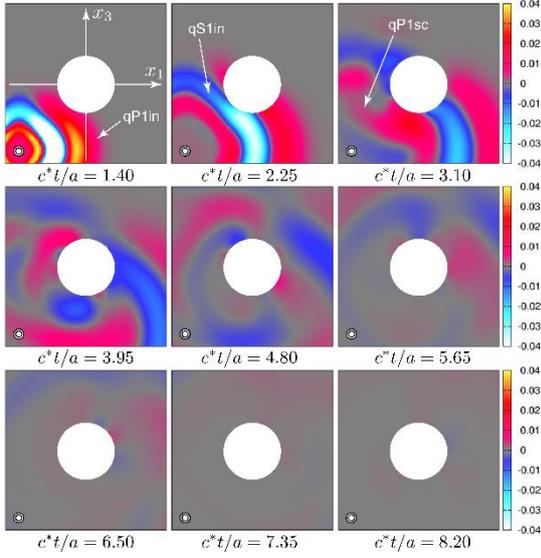
Figure 4. Time histories of the fluid pressure p around a cavity at several times.

Table 1. Effect of GPU parallelization.

	CPU times (s)
CQ-BEM without paral-lelization	23847.59
CQ-BEM with OpenMP using 24 threads	2001.67
CQ-BEM with GPU par-allelization	940.36

In Eqs. (25), $n_i(\mathbf{y})$ represents the unit normal vector at the point $\mathbf{y}$ on the boundary S.

Numerical Examples

Wave Scattering Analysis by a Cavity

Wave scattering problem by a cavity is solved by the proposed CQ-BEM. Analytical model is illustrated in Fig. 2, and sandstone [Carcione (1996)] which is known as a transversely isotropic fluid-saturated porous solid is used. The boundary S is divided into 32 boundary elements, and time increment is given by $c^* \Delta t / a = 0.05$. Note that $c^* = \sqrt{C_{66}/\rho_s}$, C_{66} is the component of elastic tensor, and ρ_s is density of solid skeleton. The total number of time step N_t is set to 512. In addition, incident wave is given by the following equations:

$$q_I^{\text{in}}(\mathbf{x}, t) = -W_{IK}(\mathbf{x}, \mathbf{y}^{\text{src}}, t) * p_K^{\text{src}}(t), \tag{26}$$

$$p_K^{\text{src}}(t) = \frac{\delta_{1K}}{2}\left\{1 - \cos\left(\frac{2\pi t}{T}\right)\right\}\{H(t) - H(t - T)\} \tag{27}$$

where $H(\cdot)$ is Heaviside function. Moreover, the source point is set to $\mathbf{y}^{\text{src}} = \{-2.5a, -2.5a\}^{\text{T}}$, and the period of source function $p_K^{\text{src}}(t)$ is set to $T = 16\Delta t$. The CQM parameters described in Eq. (13) are given as follows: $\epsilon = 1.0 \times 10^{-12}$,and $k = 1$.

Figure 3 and 4 are shown time histories of the displacement components in x_1-direction and fluid pressure, respectively. In these figures, incident waves generated at the source point are propagating with complex wave surfaces. When the incident waves arrived at the boundary of the cavity, scattered waves are generated.

Performance of GPU parallelization

In recent years, the use of GPU for general purpose computing, which is often called GPGPU, has received considerable attentions. GPGPU provides a fine-grained parallelization, and the structure of this parallelization is different from a structure using MPI which is coarse-grained parallelization. As mentioned in the previous section, the influence functions shown in Eqs. (15) and (16) require much computational time because these functions include two integrations: integrations over the boundary

element S^N and the unit circle ($|\mathbf{p}| = 1$). Therefore, in this study, GPU parallelization is applied to the numerical computations of the influence functions. For the numerical computations, TSUBAME 2.5 which is the supercomputer of Tokyo Institute of Technology is used. CPU and GPU of TSUBAME 2.5 are Intel Xeon 2.93 GHz, NVIDIA Tesla K20X, respectively.

Wave scattering problems are solved for evaluating the effect of GPU parallelization. The total number of time step N_t is set to 128, and other analytical parameters are the same as the previous analysis. Table 1 shows the computational time of three cases: CQ-BEM without parallelization, CQ-BEM with OpenMP parallelization using 24 threads, and CQ-BEM with GPU parallelization. From these results, GPU parallelization can reduce computational time of the proposed CQ-BEM dramatically.

Conclusions

In this paper, a convolution quadrature boundary element method for general anisotropic fluid-saturated porous solids and its GPU acceleration are presented. The presented Formulation is based on the formulation proposed by [Saitoh et al. (2012)], and CQM and collocation method are used for time- and spatial discretization. Wave scattering problem by a cavity in sandstone is solved using the proposed CQ-BEM. GPU parallelization is implemented to reduce the computational time of evaluations of the influenced functions given by Eqs. (15) and (16). The computational time is reduced compared with both conventional CQ-BEM and that with OpenMP. In near future, CQ-BEM for three-dimensional wave scattering problems in general anisotropic fluid-saturated porous solid will be developed.

References

Biot, M.A. (1956a) Theory of propagation of elastic waves in a fluid-saturated porous solid. I. Low-frequency range, *The Journal of the Acoustical Society of America*, **28**(2), 168–178.

Biot, M.A. (1956b) Theory of propagation of elastic waves in a fluid-saturated porous solid. II. Higher frequency range, *The Journal of the Acoustical Society of America*, **28**(2), 179–191.

Biot, M.A. (1962) Mechanics of deformation and acoustic propagation in porous media, *Journal of Applied Physics*, **33**(4), 1482–1498.

Carcione, J. M. (1996) Wave propagation in anisotropic saturated porous media: Plane-wave theory and numerical simulation, *Journal of the Acoustic Society of America*, **99**(5), 2655–2666.

Domínguez, J. (1992) Boundary element approach for dynamic poroelastic problems, *International Journal for Numerical Methods in Engineering*, 35, 307–324.

Saitoh, T., Chikazawa, F. and Hirose, S. (2012) Large-scale transient poroelastodynamic analysis using convolution quadrature time-domain boundary element method, *4th International Conference on Computational Methods (ICCM2012)*, CD-ROM.

Schanz, M. (2001) Application of 3D time domain boundary element formulation to wave propagation in poroelastic solids, *Engineering Analysis with Boundary Elements*, **25**, 363–376.

Schanz, M. and Antes, H. (1997) Application of 'Operational quadrature methods' in time domain boundary element method, *Meccanica*, **32**, 179–186.

Sharma, M. D. (2005) Polarisations of quasi-waves in a general anisotropic porous solid saturated with viscous fluid, *Journal of Earth System Science*, **114**(4), 411–419.

Takahasi, H. and Mori, M. (1974) Double exponential formulas for numerical integration, *Publications of the Research Institute for Mathematical Sciences, Kyoto University,* **9**(3), 721–741.

Advances in Computational Hydrodynamics Applied to Wave-in-Deck Loading

*Y.L. Wu[1], G. Stewart[2], Y. Chen[1], J. Gullman-Strand[1], P. Kumar[3], X. Lv[3]

[1]Lloyd's Register Global Technology Centre,
1 Fusionopolis Way, Connexis (North Tower) #17-14, Singapore 138632.
[2]Lloyd's Register EMEA, Aberdeen AB10 1NN, UK.
[3]Institute of High Performance Computing, Agency for Science, Technology and Research (A*STAR)
1 Fusionopolis Way, Connexis #16-16, Singapore 138632

*Corresponding author: yanling.wu@lr.org

Abstract

Owing to increases in predicted wave crest heights, platform settlement and changes to sea-water levels, wave-in-deck loading on offshore structures has increasingly become a concern to the offshore oil and gas industry. In this paper, a numerical approach for simulations of extreme ocean waves interacting with fixed offshore structures is presented in the framework of an open source library, OpenFOAM. A wave generation model based on the "NewWave" focused wave group approach (Tromans *et al.* 2001) has been developed to represent the extreme wave conditions. To validate the simulation, the results from the current approach have been compared to wave profiles obtained by Ning *et al.* (2009) and also with those of Iwanowski *et al.* (2002) for the wave-in-deck loads for a simple box representing the Ekofisk platform deck in the North Sea. The dynamic response of a typical supporting jacket structure when subjected to these loads is also assessed.

Keywords: CFD, NewWave, wave-in-deck, offshore structure, OpenFOAM

1. Introduction

Wave-in-deck loading arises when the total surface elevation exceeds the air gap for which an offshore platform has been designed. This results in very large step changes in the load on the structure and is a major concern for oil and gas operators. This situation is becoming increasingly common due to: changes in the wave crest statistical models that lead to higher crest predictions; seabed subsidence due to oil & gas extraction; and sea-level increases due to climate change.

Historically, the estimation of wave-in-deck load has mainly been conducted using semi-analytical formulations supported by laboratory experiments on scaled-down models (van de Graaf *et al.*, 1995). With the advent of high powered computer clusters, numerical simulation on full size models using state-of-art CFD (Computational Fluid Dynamics) techniques now offers an alternative for determining these complicated hydrodynamics forces. The advantages of numerical simulation through CFD are: 1) full scale simulation of nonlinear phenomena; 2) potentially more accurate (and less costly) prediction compared with model testing as viscous and inertia forces are included whereas a model based on Froude scaling can only capture the inertia forces; and 3) detailed insight into the flow and resulting loads.

To evaluate wave-in-deck impact loads, an extreme wave generator needs to be employed. The extreme wave occurs as a highly transient event within a multi frequency sea state. Regular waves, such as Stokes, do not represent these extreme waves accurately and random wave generation is an extremely time consuming process, as these extreme events occur rarely in random time series. An efficient method is to use a "NewWave" focused wave group that describes the average shape of an extreme wave profile consistent with a random process and a specified energy spectrum (Tromans *et al.*, 1991). NewWave theory combines random wave theory and conditional probability theory to obtain the frequency components and relative amplitudes of the wavelets of the target extreme

waves. These wave components interact and constructively interfere to build up a localized extreme wave, focused at a specified position in the domain. The representation has been studied theoretically by Boccotti (1983) and Tromans *et al.* (1991) and experimentally and numerically by several investigators, such as Taylor and Haagsma (1994), Baldock *et al.* (1996) and Borthwick *et al.* (2006). In this study, we adopt the numerical setup for the generation of NewWave as published in Ning *et al.* (2009).

The NewWave model has been developed by leveraging on the open source CFD tool OpenFOAM in this study. This has an extensive range of features to solve various fluid flow problems. In the current OpenFOAM platform, the Navier-Stokes (N-S) equations are used to describe the fluid flow while the Volume of Fluid method (VOF) is used to capture movement of the water free surface.

2. Methodology and simulation

2.1 NewWave Theory

The concept of the NewWave formulation is to generate the extreme waves from a specified frequency spectrum by superimposing several relatively small waves to form one focused extreme wave at a specified location and specified time. For the linear NewWave, each wave component i, of frequency f_i, the amplitude a_i is defined (see for example Ning *et al.*, 2009) as

$$a_i = A \frac{S(f_i)\Delta f}{\sum\limits_{i}^{N} S(f_i)\Delta f} \tag{1}$$

where $S(f)$ is the spectral density and Δf is the frequency step depending on the number of wave components N and bandwidth. A is the target theoretical linear wave amplitude of the focused wave. The extreme wave represented by linear NewWave theory is simply the scaled auto-correlation function corresponding to a specified spectrum.

The free surface elevation and velocity components are obtained by superposition as:

$$\eta = \eta^{(1)} = \sum_{i=1}^{N} a_i \cos[k_i(x - x_0) - \omega_i(t - t_0)] \tag{2}$$

$$u = u^{(1)} = \sum_{i=1}^{N} \frac{a_i g k_i}{\omega_i} \frac{\cosh k_i(z + h)}{\cosh(k_i h)} \cos[k_i(x - x_0) - \omega_i(t - t_0)]$$

$$w = w^{(1)} = \sum_{i=1}^{N} \frac{a_i g k_i}{\omega_i} \frac{\sinh k_i(z + h)}{\cosh(k_i h)} \sin[k_i(x - x_0) - \omega_i(t - t_0)]$$

where z is the vertical coordinate measured upwards from the Mean Water Level (MWL), η is the instantaneous free surface elevation, x_0, t_0 are the predefined focal location and focal time, respectively, g is the gravitational acceleration, h is the water depth, $k_i = \omega_i^2 / g \tanh(k_i h)$ is the wave number and $\omega_i = 2\pi f_i$ is the frequency. The superscript $^{(1)}$ denotes linear contributions.

For the second order NewWave [Ning *et al.* (2009), Hu *et al.* (2011) and Westphalen *et al.* (2012)], the corresponding wave elevation and velocity components u and w can be expressed as:

$$\eta = \eta^{(1)} + \eta^{(2)}$$
$$u = u^{(1)} + u^{(2)} \tag{3}$$
$$w = w^{(1)} + w^{(2)}$$

where $\eta^{(1)}$, $u^{(1)}$ and $w^{(1)}$ are the linear wave elevation and velocities, respectively and $\eta^{(2)}$, $u^{(2)}$ and $w^{(2)}$ correspond to the second order wave elevation and velocities, respectively. Some conflicts exist amongst different papers on 2nd order terms of Eq.(3). The details of the 2nd order terms in Eq.(3) that we derived and used in this study are given in the Appendix.

Various idealized spectra may be used to represent the sea states. The JONSWAP frequency spectrum $S(f)$ is frequently employed (e.g. Gao *et al*. 2012) and is used herein:

$$S(f) = \beta_J H_s^2 T_p^{-4} f^{-5} \exp[-1.25(T_p f)^{-4}] \gamma_\alpha^{\exp[-(T_p f - 1)^2 / 2\lambda^2]} \tag{4}$$

$$\beta_J \cong \frac{0.06238(1.094 - 0.01915 \ln \gamma_\alpha)}{0.230 + 0.0336 \gamma_\alpha - 0.185(1.9 + \gamma_\alpha)^{-1}} \quad ; \quad \lambda = \begin{cases} 0.07 & f \le f_p \\ 0.09 & f > f_p \end{cases}$$

where H_s is the significant wave height; T_p and f_p are the peak wave period and frequency respectively. The peak enhancement factor γ_α was taken as 3.3. Note that for the NewWave formulation the value of H_s is not relevant since the normalized spectrum is used (see Eq. (1)).

2.2 NewWave boundary conditions for CFD

For the boundary conditions in our CFD simulations we may use either the first order NewWave solution as given by Equation (1) or the second order NewWave solution provided by Equation (3) as the input initial conditions. In this paper, all results are generated using second order NewWave.

In general, there are two different initial conditions that can be used in the CFD simulation.

Type 1 initial condition: surface profile and kinematics prescribed over the entire domain
The surface profile and associate kinematics are imposed over the entire domain at *t*=0. At *t* > 0, the waves and kinematics are input at the boundary *x* = 0m.

Type 2 initial condition: surface profile and kinematics prescribed at inlet boundary, zero conditions over remainder of domain.
The surface profile and associate kinematics are imposed at the inlet boundary only. At t = 0, x > 0 the surface profile and kinematics are zero over the entire domain.

2.3 NewWave generation validation

The OpenFOAM solver with the NewWave generator was validated by comparison with the analytical solution and numerical results from Ning *et al*. (2009). For the numerical simulations, it was assumed that the fluid is incompressible, the surface tension on the free-surface can be ignored and the mean water depth is constant. No turbulence model was applied.

The setup was similar to the one in Ning *et al*. (2009). The computational domain was 13m long, 1m high with a water depth of 0.5m. Between *x* =10m and *x* =13m, a relaxation zone in Wave2Foam was installed to prevent reflections from the right-hand boundary. The wavemaker was located at *x* =0m. We note, in passing, that once the waves leave the input boundary, their

propagation is controlled by fully nonlinear wave-wave interactions. We compare the CFD results with Case 2 in Ning *et al*. (2009) (as shown in Table 1) where the predefined focal point was set at $x_0 = 3m$, $t_0 = 9.2s$ and the linear input amplitude was $A = \eta_{max}^{(1)} = 0.0632$m. The corresponding theoretical second order amplitude at the focal point is $\eta_{max}^{(2)} = 0.0677$m.

Both types of initial conditions mentioned above were tested and the results were almost identical; however, Type 1 can employ a much shorter focal time because the simulation starts from an existing developed wave field and this is much more computationally efficient. Therefore, in the following, we consider only the results for initial condition Type 1.

Ning *et al*. (2009) found that due to nonlinear interaction of the NewWave components, the highest elevation (the real focus point for the wave) occurs at (x_1, t_1), where ($x_1 \geq x_0, t_1 \geq t_0$). In our study, this "focusing delay phenomenon" is investigated. Several "probe points" were set around x_0 and the position of the free surface was extracted at these positions for every time step to identify the maximum surface elevation.

Various mesh sizes were used (Table 2) to find the effect on the amplitude and real focus point location. The models were run on an HP Elitebook 8570W using 4 cores. The run times for calculation of wave propagation for 20s in time domain at coarse ($\Delta x \approx \lambda_P / 67, \Delta y \approx H / 32$, $\Delta t = 0.01$), fine ($\Delta x \approx \lambda_P / 100, \Delta y \approx H / 56$, $\Delta t = 0.001$), and finest mesh ($\Delta x \approx \lambda_P / 200, \Delta y \approx H / 102$, $\Delta t = 0.001$) were about 2 hours, 4.5 hours and 7.5 hours, respectively. It was found that the fine mesh size ($\Delta x \approx \lambda_P / 100, \Delta y \approx H / 56$, $\Delta t = 0.001$) is the optimal choice in term of accuracy and efficiency. Table 2 shows the results of the study: the real focus point occurs at (x_1, t_1) and is delayed in both time and location.

Table 1 Simulated case as per Case 2 in Ning et al., 2009

Frequency band (Hz)	Input Amplitude (m)	No. of wave components	Peak frequency (Hz)	Peak wave period T_p (s)	Characteristic wave length λ_P (m)
0.6 – 1.3	0.0632	16	0.833	1.2	2.0

Table 2 Maximum elevation and focal point comparison for different mesh size
($x_0 = 3m$, $t_0 = 9.2s$, $\eta_{max}^{(1)} = A = 0.0632$m, $\eta_{max}^{(2)} = 0.0677$m)

Case	Maximum elevation η_{max} (m)	Location of x_1 (m)	Time of occurrence t_1 (s)	Max crest/ linear crest, $\eta_{max} / \eta_{max}^{(1)}$	Max crest/ 2nd order crest, $\eta_{max} / \eta_{max}^{(2)}$
Coarse mesh	0.0691	4.0	9.68	1.093	1.021
Fine mesh	0.0750	3.5	9.57	1.187	1.108
Finest mesh	0.0771	3.4	9.37	1.220	1.139
Ning et al. (2009)	0.0704	3.4	9.64	1.114	1.040

Figure 1 shows the time history of surface elevation at the actual focal point obtained by the NewWave CFD simulation based on the fine mesh model (blue dashed line). The second order NewWave analytical solution is also shown for comparison (red line) and the black line is Ning *et al.*'s (2009) results. Ning *et al.* used a Higher-Order Boundary Element Method (HOBEM) with mesh size is $\Delta x \approx \lambda_P / 30$ and time step $\Delta t = T_P / 50 = 0.024$. It is clear that the surface elevations at the real focal point from both OpenFOAM and the HOBEM solver used by Ning *et al.* are higher than the analytical solution. This is because the effect of nonlinear wave-wave interactions beyond second order is not included in the analytical solution. Overall the comparison is good although the surface elevation from the N-S solver (OpenFOAM) is 6% higher than the potential flow solver used by Ning *et al.* This may be due to the different mesh size and time step, or due to the difference between N-S solver and HOBEM solver.
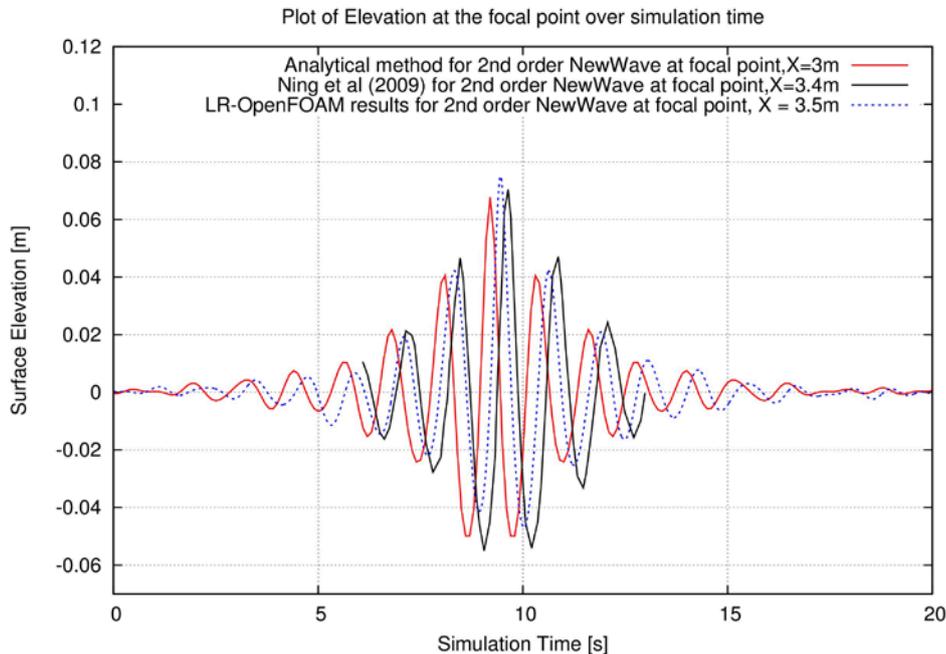
Figure 1 Comparison of wave elevations at focal points

2.4 Numerical results for Wave-in-deck simulations

Iwanowski *et al.* (2002) calculated 100 year wave-in-deck loads for a model representing the Ekofisk platform deck in the North Sea. They presented and compared load time histories calculated by several different approaches, including analytical formulations and CFD simulations. In their work, the incident wave was a regular Stokes 5th order defined by the parameters *H* (wave height), *T* (wave period), and water depth *d*. We compare their results with our NewWave CFD model - the parameters for the target focused wave are summarized in Table 3. A comparison of our CFD results with Iwanowski *et al.*'s for wave-in-deck loads for Stokes 5th order waves (Chen *et al.* 2014) is also presented. The deck was modelled as a simple box being 50m long and 10m high with wave inundation at 4m. Because this study is based on 2D simulation, the actual width of the deck being 30m (normal to the wave propagation direction) is only used in post-processing to calculate the force for comparison with Iwanowski *et al.*

For the NewWave simulation, a conversion factor of 1/0.93 was used to obtain T_p from the Stokes wave period (T). We then matched the amplitude[1] of the Stokes 5th and the simulation based on NewWave. For a Stokes 5th wave of height 24.3m the corresponding wave amplitude is 14.263m. For the Newwave CFD, the input amplitude A was adjusted to 11.91m through trial and error to get the required wave elevation 14.263m. To achieve the required wave impact height $h_{imp} = 4$ the structural model was placed 10.263m above the water free surface.

In the following simulations, the mesh size around the free surface and around the deck is $\Delta x \approx \lambda_P / 150$, $\Delta y \approx H / 50$. Type 1 initial condition was used with second order NewWave.

Table 3 Parameters of waves for the model of Iwanowski et al. (2002)

Parameter	Stokes 5th wave	NewWave
Water depth d (m)	80	80
Wave elevation E (m)	14.263	14.263 ($A = 11.91$)
Wave Height H (m)	24.3	22.13 for the focus wave (calculated from the numerical results)
Wave period T (sec)	14.5	$T_p = (T/0.93) = 15.59$
Peak frequency (Hz)	-	0.06414
Frequency band in JONSWAP spectrum (Hz)	-	0.0237-0.1924 (40 components)
Wave length λ (m)	320	$\lambda_P = 320$
Impact Height (h_{imp}) (m)	4	4
Predefined focal time t_0 for NewWave (s)	-	$t_0 = 1.5 T_p = 23.34$
Predefined focal position x_0 for NewWave (m)	-	$x_0 = 1.5\lambda_P = 480$

Figures 2 and 3 show the time force curves for the horizontal force Fx and vertical force Fz for OpenFOAM NewWave 2D simulation along with the 2D results based on Stokes 5th waves. The Stokes 5th wave 2D results from Iwanowski *et al*. (2002) by FLOW3D are also compared in the figures.

---

1 In practice the crest height would be obtained using Forristall crest statistics (Forristall, 2000) and then NewWave and Stokes 5th waves would be selected to match that crest height.
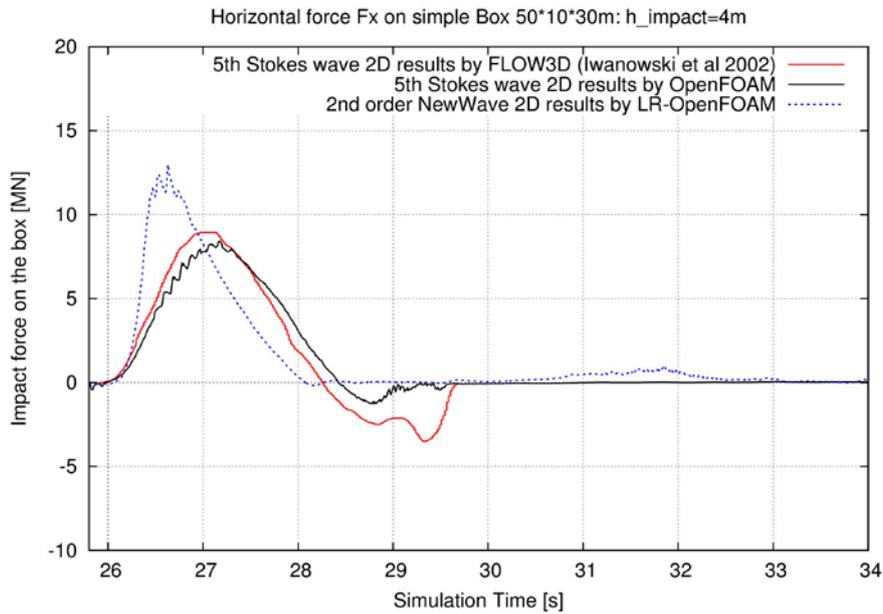
Figure 2   Comparison of horizontal force
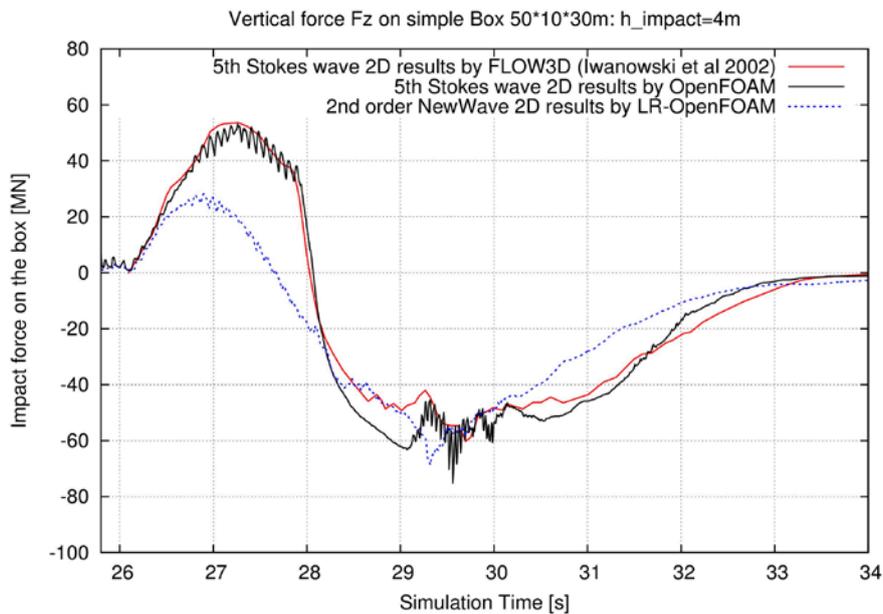
Figure 3   Comparison of vertical force

It is clear that there is good agreement between Stokes 5th wave 2D results for our OpenFOAM simulation and FLOW3D. This demonstrates that the present computational simulation is comparable with other available CFD results in the literature.

For the NewWave simulation with the same elevation as the Stokes 5th wave, it was found that the peak of horizontal force acting on the simple box for NewWave is higher than that of the Stokes 5th wave and its duration is less; however the area under the curve of the NewWave simulation (the impulse) is essentially the same as that for the Stokes 5th wave (NewWave impulse = 11.0 MNs; Stokes impulse = 10.9 MNs). In practice, one would derive the crest elevation based on Forristall crest statistics (Forristall, 2000). This study indicates that if a Stokes wave is matched to the crest amplitude, the corresponding force may be significantly underestimated compared with a NewWave

with the same crest amplitude. It is therefore recommended that focused waves based on NewWave are employed in CFD simulations.

The forces obtained from our simulations are the forces on the deck without consideration of structural dynamic response. The effective force applied to the supporting jacket structure depends on the peak force, the duration of the force and the natural period of the jacket. To determine the effect of dynamics of the structure, the static force-time histories shown in Figure 2 were applied to a single degree of freedom mass-spring-damper system that represents the supporting jacket structure with a natural frequency of 2.5 secs and 3% critical damping. The input force (f_deck) and response curves (force in the jacket) are shown in Figure 4. In this case, although the overall impulse is the same, the resulting force in the jacket is greater for NewWave compared with Stokes 5th. It is this dynamically enhanced force that is used for the assessment of structural integrity of the jacket structure.

From the above we note that the static wave-in-deck force is amplified by the dynamic behavior of the jacket structure resulting in higher loads being transmitted to the jacket. The NewWave *dynamic amplification factor* (DAF) is 1.39 while the Stokes 5th DAF is 1.64. These are fairly typical values for a fixed jacket structure with wave-in-deck loading – the particular value depends on the applied wave-in-deck force-time history and jacket natural frequency as mentioned above.
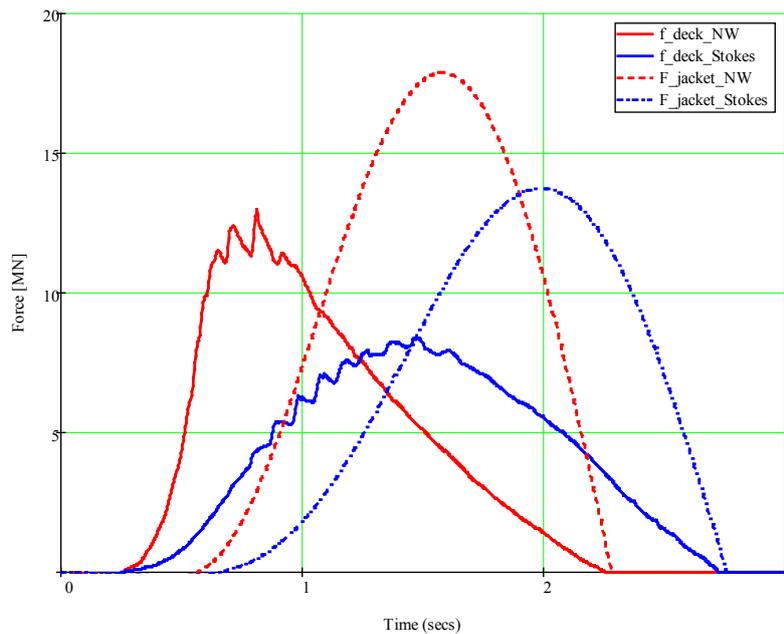
Figure 4   Lateral dynamic response of jacket structure to wave-in-deck loading
(Natural period = 2.5 secs; 3% critical damping)

For the vertical force comparison in Figure 3, there are significant differences between the NewWave and Stokes 5th wave solutions. The upward vertical force Fz for NewWave is about 1/2 that of the Stokes 5th wave. This is most likely due to the different wave shape and crest velocity of the different wave theories.

3. Conclusions

NewWave theory provides an efficient description of the average profile of an extreme event in a random sea. The advantage of NewWave is that the extreme wave can be generated at a predefined

location and time without extensive random time domain simulations. When implemented as an initial condition into a CFD package that solves the Navier-Stokes equations, cost effective wave-in-deck loading simulations can be undertaken that include the full non-linearity of the waves.

This paper describes the development of an extreme wave generator based on second order NewWave theory that was implemented into the OpenFOAM CFD software. A comparison between NewWave and a Stokes 5th wave has been made by calculating the wave-in-deck loading on a simple box and the corresponding response of a supporting jacket structure. This study indicated that if a Stokes 5th wave is matched to the crest amplitude of NewWave, the applied horizontal deck force and the jacket response may be significantly underestimated compared to NewWave. The upward vertical forces from NewWave are substantially lower than the Stokes 5th wave. It is therefore recommended that focused waves based on NewWave are employed in CFD simulations.

Acknowledgments

Disclaimer

The views expressed in this paper are those of the authors and do not necessarily reflect those of their affiliated companies.

References

Baldock T.E., Swan C., Taylor P.H. (1996): A laboratory study of nonlinear surface waves on water, Philosophical Transactions of the Royal Society of London 354A : 649-676.

Boccotti P., (1983): Some New results on Statistical properties of Wind Waves, Applied Ocean Research, Vol. 5, No. 3.

Borthwick A.G.L., Hunt A.C., Feng T., Taylor P.H., Stansby P.K. (2006): Flow kinematics of focused wave groups on a plane beach in the UK coastal research facility, Coastal Engineering 53 (12): 1033-1044.

Chen Y., Wu Y.L., Stewart G., Gullman-Strand J., and Lu X. (2014): Numerical simulation of wave in deck loading on offshore structures, OMAE2014-23847, Proceedings of the ASME 33rd International Conference on Ocean, Offshore and Arctic Engineering, June 8-13, 2014, San Francisco, California, USA.

Dalzell J.F.(1999): A note on finite depth second-order wave-wave interactions, Applied Ocean Research 21:105-111.

Forristall, G.Z. (2000): Wave crest distributions – Observations and second order theory, J. Phys Oceanog. Vol 30, No. 8: 1931-1943.

Gao F., Mingham C., and Causon D.: Simulation of extreme wave interaction with Monopile mounts for offshore wind turbines, Coastal Engineering 2012.

Hu Z.Z., Causon D.M., Mingham C.G., and Qian L.(2011): Numerical simulation of floating bodies in extreme free surface waves, Nat.Hazards Earth Syst. Sci.,11,519-527,2011

ISO 19902+A1, Petroleum and Natural Gas Industries-Fixed Steel Offshore Structures (October 2013)

Iwanowski, B., Grigorian, H., & Scherf, I. (2002). Subsidence of the Ekofisk Platforms: Wave in Deck Impact Study—Various Wave Models and Computational Methods. OMAE2002-28063, Proceedings of OMAE'02, 21st International Conference on Offshore Mechanics and Artic Engineering, June 23-28, 2002, Oslo, Norway.

Jacobsen, N. G., Fuhrman, D. R., & Fredsøe, J. (2012). A wave generation toolbox for the open source CFD library: OpenFoam®. International Journal for Numerical Methods in Fluids, 70(9), 1073-1088.

Ning D.Z., Zang J., Liu S.X., Eatock Taylor R., Teng B., Taylor P.H.(2009): Free surface evolution and wave kinematics for nonlinear uni-directional focusd wave groups, Ocean Engineering, 36:1226-1243.

Taylor P.H., Haagsma I.J.(1994): Focusing of steep wave groups on deep water, Proceedings of the International Symposium: Waves-Physical and Numerical Modelling, Vancouver, Canada, 862-870

Tromans P.S., Anaturk A.R., Hagemeijer A., (1991): A new model for the kinematics of large ocean waves-Application as a design wave, Proceeding of 1st International Offshore and Polar Engineering Conference, vol 3, Edinburgh, UK, 64-71

Van de Graaf, J.W., Tromans, P.S., Vanderschuren, L. (1995): Wave Loads on Decks. Shell Offshore Structures Engineering Newsletter, No. 10. Feb.

Westphalen J, Greaves D.M., Williams C.J.K., Hunt-Raby A.C., and Zang J. (2012): Focused waves and wave-structure interaction in a numerical wave tank, Ocean Engineering 45 9-12, 2012.

Appendix: Second order NewWave theory

For 2nd order NewWave theory (Eq (3)), the underlying equations for 2nd order terms can be derived according to second order Stokes theory (Ning *et al.* 2009), which can be written as

$$\eta^{(2)} = \sum_{i=1}^{N}\sum_{j=i+1}^{N}\{a_i a_j B^+ \cos[(k_i + k_j)(x - x_0) - (\omega_i + \omega_j)(t - t_0)]+$$

$$a_i a_j B^- \cos[(k_i - k_j)(x - x_0) - (\omega_i - \omega_j)(t - t_0)]\}+ \qquad (i)$$

$$\sum_{i=1}^{N}\{\frac{a_i^2 k_i}{4\tanh(k_i h)}(2 + \frac{3}{\sinh^2(k_i h)})\cos[2(k_i(x - x_0) - \omega_i(t - t_0))] - \frac{a_i^2 k_i}{2\sinh(2k_i h)}\}$$

$$u^{(2)} = \sum_{i=1}^{N}\sum_{j=i+1}^{N}\{a_i a_j A^+ (k_i + k_j)\frac{\cosh[(k_i + k_j)(z + h)]}{\cosh[(k_i + k_j)h]}\cos[(k_i + k_j)(x - x_0) - (\omega_i + \omega_j)(t - t_0)]+$$

$$a_i a_j A^- (k_i - k_j)\frac{\cosh(k_i - k_j)(z + h)}{\cosh(k_i - k_j)h}\cos[(k_i - k_j)(x - x_0) - (\omega_i - \omega_j)(t - t_0)]\}+ \qquad (ii)$$

$$\sum_{i=1}^{N}\frac{3a_i^2 k_i \omega_i}{4}\frac{\cosh(2k_i(z + h))}{\sinh^4(k_i h)}\cos[2(k_i(x - x_0) - \omega_i(t - t_0))]$$

$$w^{(2)} = \sum_{i=1}^{N}\sum_{j>i}^{N}\{a_i a_j A^+ (k_i + k_j)\frac{\sinh[(k_i + k_j)(z + h)]}{\cosh[(k_i + k_j)h]}\sin[(k_i + k_j)(x - x_0) - (\omega_i + \omega_j)(t - t_0)]+$$

$$a_i a_j A^- (k_i - k_j)\frac{\sinh[(k_i - k_j)(z + h)]}{\cosh[(k_i - k_j)h]}\sin[(k_i - k_j)(x - x_0) - (\omega_i - \omega_j)(t - t_0)]\}+ \qquad (iii)$$

$$\sum_{i=1}^{N}\frac{3a_i^2 k_i \omega_i}{4}\frac{\sinh[(2k_i(z + h)]}{\sinh^4(k_i h)}\sin[2(k_i(x - x_0) - \omega_i(t - t_0))]$$

$$D^{\pm} = (\omega_i \pm \omega_j)^2 - g(k_i \pm k_j)\tanh[(k_i \pm k_j)h] \qquad (iv)$$

$$A^+ = -\frac{\omega_i \omega_j(\omega_i + \omega_j)}{D^+}\left[1 - \frac{1}{\tanh(k_i h)\tanh(k_j h)}\right] + \frac{1}{2D^+}\left[\frac{\omega_i^3}{\sinh(k_i h)^2} + \frac{\omega_j^3}{\sinh(k_j h)^2}\right] \qquad (v)$$

$$A^- = \frac{\omega_i \omega_j(\omega_i - \omega_j)}{D^-}\left[1 + \frac{1}{\tanh(k_i h)\tanh(k_j h)}\right] + \frac{1}{2D^-}\left[\frac{\omega_i^3}{\sinh(k_i h)^2} - \frac{\omega_j^3}{\sinh(k_j h)^2}\right] \qquad (vi)$$

$$B^+ = \frac{(\omega_i^2 + \omega_j^2)}{2g} - \frac{\omega_i \omega_j}{2g}\left[1 - \frac{1}{\tanh(k_i h)\tanh(k_j h)}\right] \cdot \left[\frac{(\omega_i + \omega_j)^2 + g(k_i + k_j)\tanh((k_i + k_j)h)}{D^+}\right]$$

$$+ \frac{(\omega_i + \omega_j)}{2gD^+}\left[\frac{\omega_i^3}{\sinh(k_i h)^2} + \frac{\omega_j^3}{\sinh(k_j h)^2}\right] \qquad (vii)$$

$$B^- = \frac{(\omega_i^2 + \omega_j^2)}{2g} + \frac{\omega_i \omega_j}{2g}\left[1 + \frac{1}{\tanh(k_i h)\tanh(k_j h)}\right] \cdot \left[\frac{(\omega_i - \omega_j)^2 + g(k_i - k_j)\tanh((k_i - k_j)h)}{D^-}\right]$$

$$+ \frac{(\omega_i - \omega_j)}{2gD^-}\left[\frac{\omega_i^3}{\sinh(k_i h)^2} - \frac{\omega_j^3}{\sinh(k_j h)^2}\right] \qquad (viii)$$

An extra dof-free and well-conditioned XFEM

Longfei Wen[1,2], *Rong Tian[1]

[1]Institute of Computing Technology, Chinese Academy of Sciences, Kexueyuan Nanlu 6, Haidian, Beijing 100190, China
[2]University of Chinese Academy of Sciences, No.19A Yuquan Road, Beijing 100049, China

*Corresponding author: rongtian@NCIC.AC.CN

Abstract

An improved XFEM (in short, *i*XFEM) is introduced. The core of *i*XFEM is an extra-dof free and *interpolating* PU enrichment based on a (moving) least-squares. Compared with the current XFEM, *i*XFEM does not introduce extra dof in PU enrichment and is well-conditioned in geometrical refinement. Selected numerical examples are provided to demonstrate its numerical performance.

Keywords: selectively interpolating moving least squares, partition of unity, XFEM

Introduction

The extended finite element method (XFEM [Belytschko and Black (1999); Moes *et al.* (1999); Belytschko *et al.* (2001)]) is a numerical technique based on the generalized finite element method (GFEM [Duarte *et al.* (2000); Duarte *et al.* (2001); Simone et al. (2006); Duarte and Kim (2008); O'Hara *et al.* (2009); Strouboulis *et al.* (2000a; 2000b; 2001)]) and the partition of unity method (PUM [Babuška and Melenk (1996; 1997)]). Only local parts of the domain are enriched and the mesh generation does not need to consider the internal structure. It adds special enrichment functions to approximate discontinuous fields and has been used for general interface phenomena e.g. in the framework of multi-material problems [Sukumar et al. (2001)], solidification [Chessa *et al.* (2002)], shear bands [Areias and Belytschko (2006)], dislocations [Belytschko and Gracie (2007)], and multi-field problems [Zilian and Legay (2008)].

Recently, a new GFEM without extra dof has been proposed as an improvement on the existing GFEM [Tian (2013)]. Based on the extra-dof free PU approximation, we have developed an improved XFEM with two distinguished features: (1) extra-dof-free; (b) well-conditioned in the so-called geometrical refinement, which means the size of enriched sub-domain is fixed during mesh refinement. This short paper is intended for a brief introduction to the improved version of XFEM and a comparison with the two existing counterparts: the originally standard XFEM and the corrected XFEM [Fries (2008)]. By standard XFEM we mean a XFEM without blending element treatment. The corrected XFEM is an improved version over the standard XFEM by eliminating the blending element issue using a neat ramp formulation.

The extra-dof free and interpolating XFEM

Extra-dof free PU enrichment

Let Ω^e be a sub-domain where all nodes are enriched. Let P_i^r denote nodal patch i composed of nodes from Ω^e, where r is the patch size. The patch size is either the size of nodal support combining $m \geq 1$ layer(s) of elements surrounding node i for a

structured or an non-uniform mesh or simply the radius of an influence circle at node i for a unstructured uniform mesh. Node i is called "patch star" of P_i^r, the topological centre of a patch. $\{x_k | x_k \in P_i^r\}$ is the node set of P_i^r, where index i is solely kept for the patch star and k $(k \neq i)$ for non-patch star nodes on the patch.

On the patch, we construct an interpolant of $u_i(\mathbf{x}), \ \mathbf{x} \in P_i^r$, using the nodal values at $\{x_k | x_k \in P_i^r\}$ as

$$u_i(x) = \sum_{k=1}^{n_i} L_k^{P_i^r}(x) u_k, \quad x, x_k \in P_i^r \tag{0}$$

where $L_k^{P_i^r}(x)$ is the shape function at k-th node of the node set, which also forms a partition of unity, $\sum_{k=1}^{n_i} L_k^{P_i^r}(x) \equiv 1$, n_i is the number of nodes on P_i^r, and u_k is the conventional nodal unknown. Noted is that this approximation is local to patch P_i^r. The approximation is called patch-wise local approximation.

Using the patch-wise local approximation $u_i(\mathbf{x})$ as a local approximation at node i and substitute it for the nodal unknown u_i in the following standard FEM

$$u^h(x) = \sum_{i=1}^{N} N_i(x) u_i \tag{0}$$

we obtain a new PU approximation

$$u^h(x) = \sum_{i=1}^{N} \left(N_i \left(\sum_{k=1}^{n_i} L_k^{P_i^r} u_k \right) \right) \tag{0}$$

where N and u_k are the same as those in the standard FEM (2). The difference is that in the new approximation (3) each u_i now is associated with n_i enrichment functions:

$$\underbrace{N_i L_i^{P_i^r}}_{\text{patch star node}}, \quad \underbrace{N_1 L_i^{P_1^r}, \ N_2 L_i^{P_2^r}, \ \cdots, \ N_{k,k\neq i} L_i^{P_k^r}, \cdots, \ N_{n_i} L_i^{P_{n_i}^r}}_{\text{non patch star nodes}} \tag{0}$$

where $L_i^{P_k^r}$ ($k = 1, 2, \ldots i, \ldots, n_i$) are the local functions constructed on patch k with regard to node i, and index k can be understood as either the k-th node of the patch i or the k-th patch containing node i. Expanding and regrouping all the terms associated with u_i we obtain

$$u^h(x) = \sum_{i=1}^{N} \left(N_i L_i^{P_i^r} + \sum_{k=1,k\neq i}^{n_i} N_k L_i^{P_k^r} \right) u_i = \sum_{i=1}^{N} \tilde{N}_i u_i \tag{0}$$

where n_i, the number of nodes on patch i in (1), now really means the number of patches containing node i, but the two numbers are the same, and $\tilde{N}_i$ denotes the new shape function, which is a summation of n_i functions.

The new PU approximation offers two unique features. One is that there is no extra dof; all nodal dofs are the conventional of the standard FE mesh. The other is that the new PU approximation interpolates as long as the patch-wise local approximation interpolates at its patch star, *no matter it interpolates or not at the rest of nodes on the patch*. Proof is immediate: if and only if $L_i^{P_i^r}(\mathbf{x}_i) = 1$, then $u^h(\mathbf{x}_i) = u_i$ because $N_{k,k\neq i}(\mathbf{x}_i) = 0$. $L_i^{P_i^r}(\mathbf{x}_i) = 1$ is called a *one-point interpolating* condition.

In order to construct the one-point interpolating patch-wise local approximation, a selectively interpolating (moving) least squares approximation is employed.

Selectively interpolating (moving) least squares approximation

An approximation of the field function $u_i(x)$ on the patch i is constructed by,

$$u_i(\mathbf{x}) = \mathbf{p}^{\mathrm{T}}(\mathbf{x})\mathbf{a}(\mathbf{x}), \ \mathbf{x} \in \mathrm{P}_i^r \tag{0}$$

where

$$\mathbf{p}(\mathbf{x}) = \left[p_1(\mathbf{x}), p_2(\mathbf{x}),..., p_m(\mathbf{x}) \right]^{\mathrm{T}}$$

$$= \left[1, \frac{x - x_i}{ch}, \frac{y - y_i}{ch}, f(r,\theta) - f(r_i,\theta_i) \right]^{\mathrm{T}} \tag{0}$$

$$f(r,\theta) = \sqrt{r}\cos\frac{\theta}{2}, \ \sqrt{r}\sin\frac{\theta}{2}, \ \sqrt{r}\sin\frac{\theta}{2}\sin\theta, \ \sqrt{r}\cos\frac{\theta}{2}\sin\theta$$

is the normalized or shifted crack tip basis vector in two dimensions, and $\mathbf{a}$ is the vector of unknown coefficients. In order to construct a moving least-squares approximation passing the patch star, the discrete L_2 error norm is defined by the following *constrained* form,

$$J = \frac{1}{2}\sum_{k=1}^{n_i} w_k(\mathbf{x})\left(\mathbf{p}^{\mathrm{T}}(\mathbf{x}_k)\mathbf{a} - u_k\right)^2 + \lambda\left(\mathbf{p}^{\mathrm{T}}(\mathbf{x}_i)\mathbf{a} - u_i\right) \tag{0}$$

For implemental convenience, the patch star i is also included in the summation term, λ is the Lagrange multiplier that is used to enforce the satisfaction of

$$\mathbf{p}^{\mathrm{T}}(\mathbf{x}_i)\mathbf{a} = u_i \tag{0}$$

Minimizing the L_2 norm with regard to $\mathbf{a}$ and λ ($\partial J/\partial\mathbf{a} = 0, \partial J/\partial\lambda = 0$) leads to a Selectively interpolating moving least squares approximation:

$$u_i(\mathbf{x}) = \sum_{k=1}^{n_i} \phi_k^{\mathrm{P}_i^r}(\mathbf{x})u_k \tag{0}$$

$$\phi_k^{\mathrm{P}_i^r}(\mathbf{x}) = \mathbf{p}^{\mathrm{T}}(\mathbf{x})\left(\mathbf{A}^{-1}\mathbf{p}_k - \frac{1}{\mathbf{A}_{11}^{-1}}\mathbf{A}_{(1)}^{-1}\mathbf{A}_{(1)}^{-\mathrm{T}}\mathbf{p}_k + \frac{1}{\mathbf{A}_{11}^{-1}}\mathbf{A}_{(1)}^{-1}\delta_{ik} \right) \tag{0}$$

$$A = \sum_{k=1}^{n_i} w_k p_k p_k^T \tag{0}$$

where $A_{(1)}^{-1}$ is the first column of $\mathbf{A}^{-1}$, A_{11}^{-1} is the first element of $A_{(1)}^{-1}$, δ is the Kronecker delta. It can be verified that the SIMLS shape function $\phi_k^{\mathrm{P}_i^r}(\mathbf{x})$ interpolates at patch star i

$$\phi_i^{\mathrm{P}_i^r}(\mathbf{x}_i) = 1, \ \phi_{j,j \neq i}^{\mathrm{P}_i^r}(\mathbf{x}_i) = 0 \tag{0}$$

while approximating at the non-patch star nodes.

By letting $w_k \equiv 1$, the local approximation is reduced to a least-squares type which is termed the Selectively Interpolating Least-Squares (SILS) local approximation. *The SILS local approximation simplifies the calculation of derivatives and therefore is much more computationally cheap.*

For both the SIMLS and the SILS, it can be verified that

$$\phi_i^{\mathrm{P}_i^r}\left(\mathbf{x}_i\right)=1,\ \phi_{j\neq i}^{\mathrm{P}_i^r}\left(\mathbf{x}_i\right)=0 \text{ and } u_i\left(\mathbf{x}_i\right)\equiv u_i \tag{0}$$

which means that the local approximation interpolates at the patch star. It can be easily verified that *using this kind of local approximation, iXFEM interpolates at all nodes although the local operation is in nature an approximation.*

Due to the limit of space, the rest details of the method are omitted here and will be presented in the conference.

Selected numerical examples

A crack tip benchmark problem is shown in Fig. 1. A square area with side-length of $2a$ and a crack of a, where $a = 0.5$. The uniform mesh is used. The enriched sub-domain is an area containing the crack tip while the rest of the domain is the standard FEM. The enriched sub-domain is the same in size during mesh refinement, which is respectively the half (Fig. 1(a)) and the one third (Fig. 1(b)) of the domain size. The patch size takes $r=2h$ on Ω_{GFE}. 10×10 Gaussian quadrature is only used on the 2×2 elements directly neighbored to the crack-tip in Fig. 1(a) and subdivided 2×2 cells on the element containing the crack tip in Fig. 1(b). On the rest elements, 2×2 Gaussian quadrature is employed. Young's modulus $E = 1000$ and the Poisson ratio $v = 0.3$ are assumed for plain strain. Along the outer boundary of the area, the displacements are prescribed to the following exact solution:

$$u = 2\left(1+v\right)\frac{K_{\mathrm{I}}}{E}\sqrt{\frac{r}{2\pi}}\cos\frac{\theta}{2}\left[1-2v+\sin^2\frac{\theta}{2}\right]$$

$$v = 2\left(1+v\right)\frac{K_{\mathrm{I}}}{E}\sqrt{\frac{r}{2\pi}}\sin\frac{\theta}{2}\left[2-2v-\cos^2\frac{\theta}{2}\right]$$

where K_{I} is taken as 1.253.

The two patterns of crack layout, mesh aligned and mesh independent, are considered for convergence tests (Fig. 1). The SIMLS and the SILS are tested and compared.
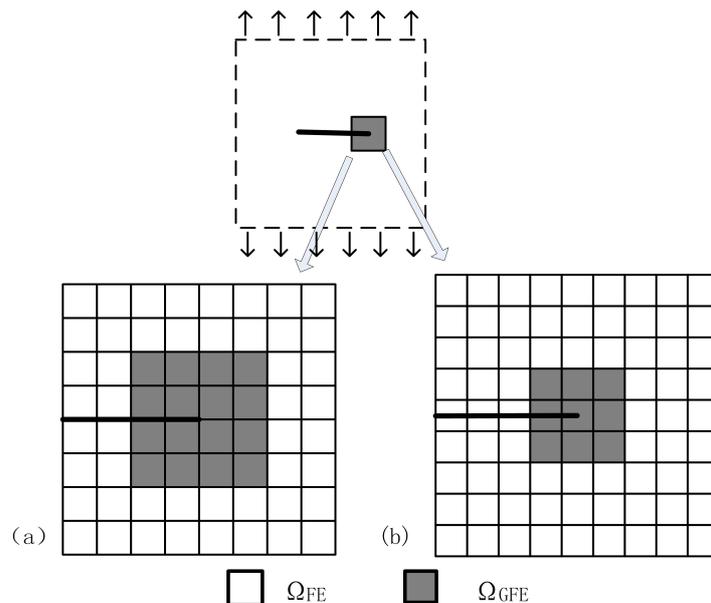
Figure 1. Crack tip benchmark problems. (a) mesh aligned crack; (b) mesh independent crack.

Assessment on convergence in geometrical refinement

Errors are measured by the L_2 and energy norms defined respectively below

$$\|u\| = \left[\int_\Omega \left(\mathbf{u}^h - \mathbf{u}^{\text{exact}}\right)^{\text{T}} \left(\mathbf{u}^h - \mathbf{u}^{\text{exact}}\right) d\Omega\right]^{\frac{1}{2}}, \|e\| = \left[\int_\Omega \left(\boldsymbol{\varepsilon}^h - \boldsymbol{\varepsilon}^{\text{exact}}\right)^{\text{T}} \left(\boldsymbol{\sigma}^h - \boldsymbol{\sigma}^{\text{exact}}\right) d\Omega\right]^{\frac{1}{2}}$$

Convergences are tested in the case of the so-called geometrical refinement. The two kernels, the SIMLS and the SILS, are tested. The results are displayed in Fig. 2. It is shown that *i*XFEM delivers optimal convergence for the crack tip benchmark problem.
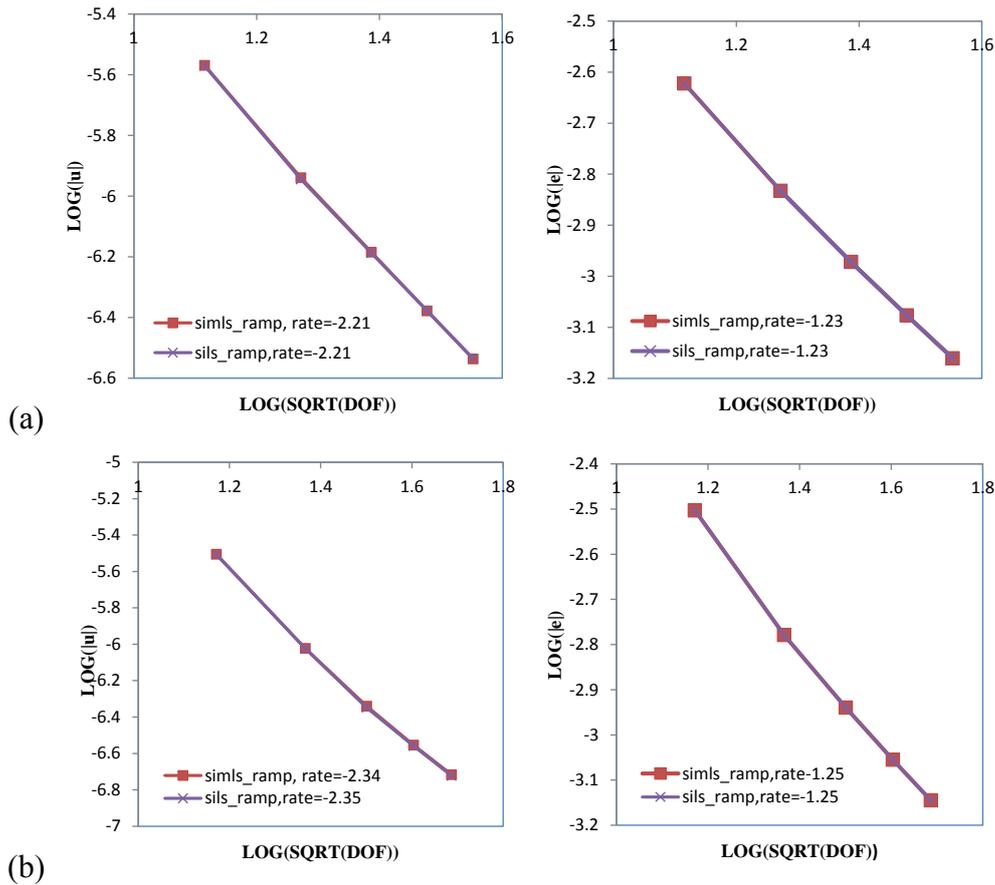
Figure 2. Convergences in geometrical refinement.
(a) mesh aligned crack; (b) mesh independent crack.

Assessment on accuracy of stress intensity factors

Comparisons are based on the crack configuration of Fig. 1(b). The SILS approximation is adopted as it is the same accurate as the SIMLS (refer to Fig. 2) but is more computationally efficient. The patch size takes *r*=2*h*. *i*XFEM and XFEM use the same size of enriched domain. *i*XFEM is compared with the standard XFEM and the corrected XFEM proposed by Fries [Fries (2008)]. 7×7 Gaussian quadrature is used on the enriched elements and 2×2 on the rest. The embedded direct solver of Matlab® is used for linear systems. This circumvents bad conditioning in XFEM, but the solver complains the singularity of the linear system in the corrected XFEM. The normalized K_I values for each method are shown in Table 1.

Table 1. Normalized K_I values

element size (h)	iXFEM		XFEM		corrected XFEM	
	dof	K_I	**dof**	K_I	**dof**	K_I
1/9	220	1.0075	344	1.0068	344	1.0045
1/15	540	1.0022	828	1.0007	828	1.0010
1/21	1004	1.0010	1520	1.0000	1520	1.0002
1/27	1612	1.0006	2420	0.9993	2420	0.9999
1/33	2364	1.0004	3528	0.9990	3528	0.9999
1/39	3260	0.9999	4844	0.9991	4844	0.9998
1/45	4300	0.9996	6368	0.9990	6368	0.9998
1/51	5484	0.9997	8100	0.9990	8100	0.9998

Assessment on conditioning properties

The conditioning is studied for the above same example by computing the maximum and the minimum eigenvalues of the global stiffness matrix before essential boundary treatment (the rigid body modes are included). The results are listed in Table 2. The eigenvalues of the standard FEM (crack and enrichment) are provided for reference. An eigenvalue less than 10^{-15} is taken to be zero. The variation of the condition number versus mesh size h is plotted in Fig. 3 for the standard XFEM, FEM and iXFEM.

Table 2. Conditioning properties
(the eigenvalues of the standard FEM are provided for reference)

	iXFEM			**XFEM**		
element size (h)	max eigenvalue	min eigenvalue	# of zero eigenvalues	max eigenvalue	min eigenvalue	# of zero eigenvalues
1/9	6.61E+03	3.40	3	6.19E+03	1.14E-03	3
1/15	5.77E+03	1.47	3	6.44E+03	3.63E-05	3
1/21	5.60E+03	0.813	3	6.52E+03	2.93E-06	3
1/27	5.63E+03	0.517	3	6.57E+03	4.72E-07	3
1/33	5.66E+03	0.358	3	6.60E+03	1.15E-07	3
1/39	5.67E+03	0.263	3	6.62E+03	3.64E-08	3
1/45	5.68E+03	0.201	3	6.64E+03	1.34E-08	3
1/51	5.69E+03	0.159	3	6.65E+03	5.70E-09	3

	corrected XFEM			**FEM**		
element size (h)	max eigenvalue	min eigenvalue	# of zero eigenvalues	max eigenvalue	min eigenvalue	# of zero eigenvalues
1/9	6.10E+03	2.41E-06	**7**	5.18E+03	51.7	3
1/15	6.36E+03	4.17E-08	**7**	5.30E+03	22.7	3
1/21	6.44E+03	3.77E-09	**7**	5.34E+03	12.1	3
1/27	6.49E+03	6.83E-10	**7**	5.36E+03	7.50	3
1/33	6.53E+03	1.81E-10	**7**	5.37E+03	5.10	3
1/39	6.56E+03	6.11E-11	**7**	5.37E+03	3.69	3
1/45	6.58E+03	2.44E-11	**7**	5.37E+03	2.79	3
1/51	6.60E+03	1.10E-11	**7**	5.38E+03	2.19	3
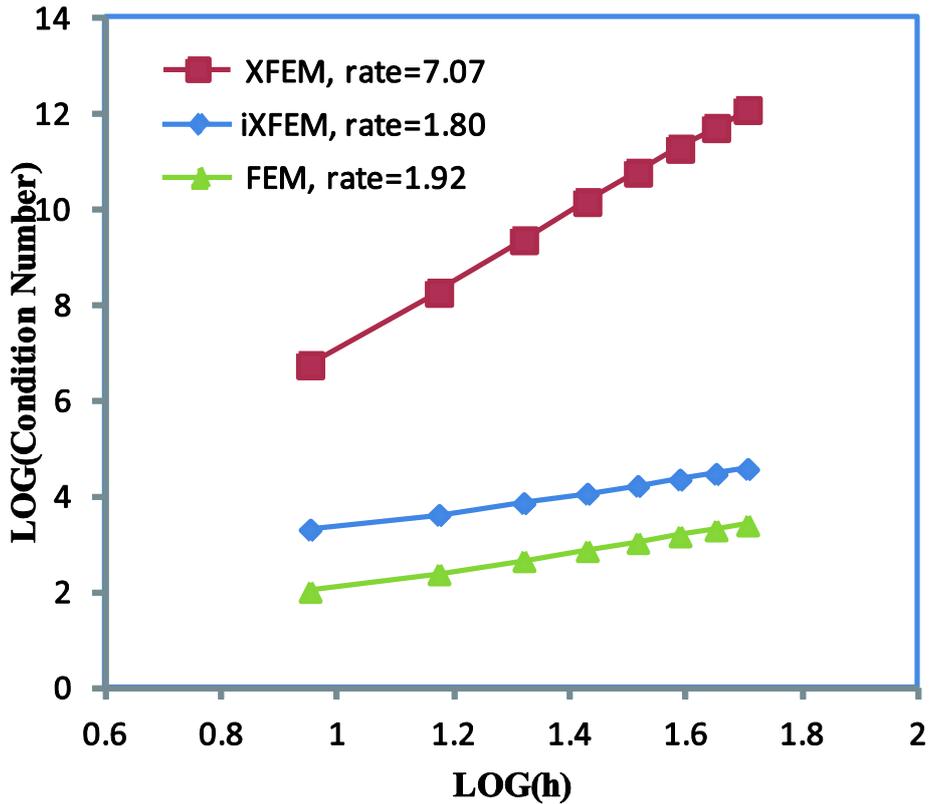
Figure 3. Variation of condition number versus mesh size.

The following conclusions can be drawn: (a) *i*XFEM offers the similar stability as the standard FEM (Fig.3) and it is much better conditioned than the current XFEM. The condition number of the global stiffness matrix in *i*XFEM grows with $h^{-1.8}$, where h is the mesh size, whereas the condition number in XFEM grows with h^{-7}; For reference, the condition number in the standard FEM grows with $h^{-1.9}$ in the same tests. (b) the corrected XFEM contains four spurious zero eigenvalues in stiffness matrices, which signals singularity of the stiffness matrix. The minimum nonzero eigenvalue is also generally smaller than that in XFEM.

Assessment on expenses on equation solving

The computational expense of the three methods is compared in terms of the number of dofs to be solved and the number of convergence iterations in a linear solve. The reason we choose an iterative solver for comparison is that subspace iterative methods are de-facto solvers in large scale problems and a direct solver is hardly scalable for problems at scale.

The conjugate gradient method is used to solve the linear system of the methods. The convergence tolerance is set to 10^{-10} for *i*XFEM and 10^{-8} for the XFEMs. The standard conjugate gradient method *without* preconditioning is used for *i*XFEM and the SSOR preconditioned conjugate gradient method, which is also an embedded solver of Matlab®, is used for XFEM and the corrected XFEM to circumvent bad conditioning. The data are listed in Table 3.

Although a better pre-conditioner or an alternative robust solver definitely can be found for the XFEMs, we still can safely conclude that the standard XFEM and the corrected XFEM obviously suffer from very slow convergence due to bad conditioning; *i*XFEM, in contrast, not only computes a smaller size of linear system but also converges remarkably easily. The latter observation is as expected and should be understandable.

Table 3. The number of dofs to be solved and convergence iterations

h	*i*XFEM (w/o preconditioning) (error tolerance: 10^{-10})		XFEM (preconditioned) (error tolerance: 10^{-8})		corrected XFEM (preconditioned) (error tolerance: 10^{-8})	
	dofs	*iterations*	*dofs*	*iterations*	*dofs*	*iterations*
1/9	220	70	344	276	344	447
1/15	540	93	828	1068	828	2337
1/21	1004	111	1520	2389	1520	5796
1/27	1612	130	2420	4395	2420	10000*
1/33	2364	149	3528	6801	3528	10000*
* the maximum number of iterations is reached but convergence is not yet observed						

Crack growth simulation in a double cantilever beam

In this section, *i*XFEM is compared to the standard and the corrected XFEMs for crack growth simulation. The dimensions of the double cantilever beam (see Fig. 4) are 6cm × 2cm and an initial pre-crack with length of *a* = 2cm is considered. Plane stress conditions are assumed with Young's modulus *E* = 1000 and the Poisson ratio *v* = 0.3. The crack is given a small perturbation at the tip of length Δa=0.1cm in with initial angle dθ=5.71° as shown in Fig. 3. A structured mesh (60 × 180) is used and the crack advances 0.1cm at each step. The stress intensity factors are computed using the interaction integral method. The evolution of the crack paths are shown in Fig. 5. The crack path obtained using *i*XFEM for a fine mesh of 150 × 450 is provided as the reference solution.
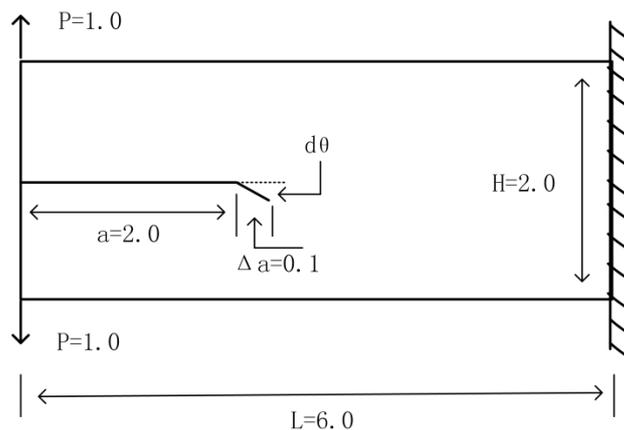
Figure 4. Geometry for the double cantilever beam
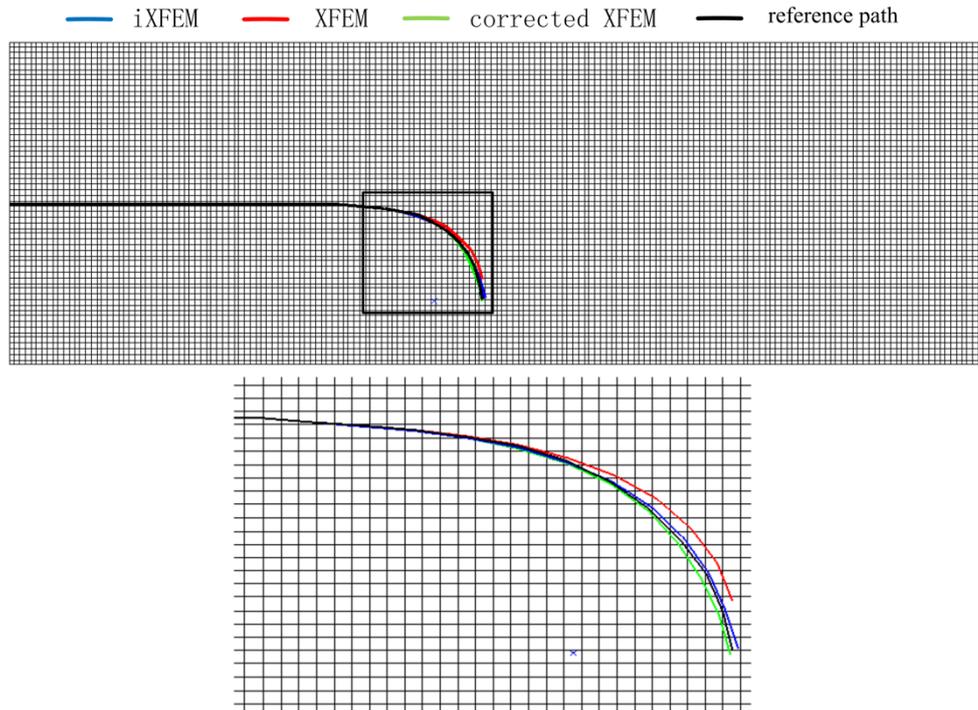
Figure 5. Double cantilever beam: comparison of crack path after 11 steps using the standard XFEM and *i*XFEM for dθ=5.71°.

Due to the limit of space, the rest of numerical examples are to be presented in the conference.

Conclusions

An improved XFEM, in short *i*XFEM, has been briefly introduced. *i*XFEM has been compared with the standard XFEM and the corrected XFEM in terms of convergences, accuracy, conditioning properties, and the expenses on equation solving. The following conclusions are drawn:

(1) The *i*XFEM offers the similarly excellent accuracy as the standard XFEM and the corrected XFEM provided that the blending element issue is dealt with properly and the enriched domain is the same in size.

(2) *i*XFEM shows remarkable improvement on conditioning. The condition number of the global stiffness matrix in *i*XFEM grows with $h^{-1.8}$, where h is the mesh size, while the existing XFEM grows with h^{-7}. In the corrected XFEM, four spurious zero eigenvalues are detected, which means the singularity of global stiffness matrix. For reference, the condition number in the standard FEM grows with $h^{-1.9}$ in the same tests. *i*XFEM shows the similar conditioning as the standard FEM. Due to the good conditioning, *i*XFEM is robust to deliver optimal convergences in geometrical refinement.

(3) *i*XFEM is computationally efficient. The shape function of *i*XFEM involves a matrix inversion operation. This increases computational expenses to a certain extent. On the other hand, for the same size of enriched domain, *i*XFEM uses less dofs and convergence iteration is also much faster compared with the standard and the corrected XFEMs.

Acknowledgements

References

Belytschko T., and Black, T. (1999) Elastic crack growth in finite elements with minimal remeshing. *International Journal for Numerical Methods in Engineering* **45**, 601–620.

Moes, N., Dolbow, J. and Belytschko T. (1999) A finite element method for crack growth without remeshing. *International Journal for Numerical Methods in Engineering* **46**,131–150.

Belytschko, T. Moes, N., Usui, S., and Parimi C.(2001) Arbitrary discontinuities in finite elements. *International Journal for Numerical Methods in Engineering* **50**, 993–1013.

Duarte, C.A., Babuska, I., and Oden, J.T. (2000) Generalized finite element methods for three-dimensional structural mechanics problems. *Computers & Structures* **77**, 215–232.

Duarte, C.A., Hamzeh, O.N., Liszka, T.J., and Tworzydlo W.W. (2001) A generalized finite element method for the simulation of three-dimensional dynamic crack propagation. *Computer Methods in Applied Mechanics and Engineering* **190**, 2227–2262.

Simone, A., Duarte, C.A., and Van der Giessen E. (2006) A generalized finite element method for polycrystals with discontinuous grain boundaries. *International Journal for Numerical Methods in Engineering* **67**, 1122–1145.

Duarte C.A., and Kim D.J. (2008) Analysis and applications of a generalized finite element method with global-local enrichment functions. *Computer Methods in Applied Mechanics and Engineering* **197**(6-8), 487–504.

O'Hara, P., Duarte, C.A., and Eason T. (2009) Generalized finite element analysis for three dimensional problems exhibiting sharp thermal gradients. *Computer Methods in Applied Mechanics and Engineering* **198**, 1857–1871.

Strouboulis, T., Babuška, I, and Copps K. (2000) The design and analysis of the generalized finite element method. *Computer Methods in Applied Mechanics and Engineering* **181**(1–3), 43–69.

Strouboulis, T., Copps, K., and Babuška, I. (2000) The generalized finite element method: an example of its implementation and illustration of its performance. *International Journal for Numerical Methods in Engineering* **47**, 1401–1417.

Strouboulis, T., Copps, K., and Babuška, I. (2001) The generalized finite element method. *Computer Methods in Applied Mechanics and Engineering* **190**(32–33), 4081– 4193.

Strouboulis, T., Zhang, L., and Babuška, I. (2003) Generalized finite element method using mesh-based handbooks: application to problems in domains with many voids. *Computer Methods in Applied Mechanics and Engineering* **192**, 3109–3161.

Babuška, I., Melenk, J.M. (1997) Partition of unity method. *International Journal for Numerical Methods in Engineering* **40**, 727 –758.

Melenk, J.M., and Babuška, I. (1996) The partition of unity finite element method: basic theory and applications. *Computer Methods in Applied Mechanics and Engineering* **139**, 289 –314.

Sukumar, N., Chopp, D.L., Moes, N., and Belytschko, T. (2001) Modeling holes and inclusions by level sets in the extended finite-element method. *Comp. Methods Appl. Mech. Engrg.* **190**, 6183 – 6200.

Chessa, J., Smolinski, P., and Belytschko, T. (2002) The extended finite element method (XFEM) for solidification problems. *International Journal for Numerical Methods in Engineering* **53**, 1959– 1977.

Areias, P.M.A. and Belytschko, T. (2006) Two-scale shear band evolution by local partition of unity. *International Journal for Numerical Methods in Engineering* **66**, 878–910.

Belytschko, T. and Gracie, R. (2007) On XFEM applications to dislocations and interfaces. *International Journal of Plasticity* **23**, 1721 –1738.

Zilian, A. and Legay A. (2008) The enriched space-time finite element method (EST) for simultaneous solution of fluid-structure interaction. *International Journal for Numerical Methods in Engineering* **75**, 305–334.

Tian, Rong (2013) Extra-dof-free and linearly independent enrichments in GFEM. *Computer Methods in Applied Mechanics and Engineering* **266**: 1-22.

Fries T.P.. (2008) A corrected XFEM approximation without problems in blending elements. *International Journal for Numerical Methods in Engineering* **75**(5), 503-532.

Pais, M., MATLAB Extended Finite Element (MXFEM) Code v1.3, www.matthewpais.com, 2011.

Finite integration method for solving multi-dimensional partial differential equations

*P. H. Wen[1,2], C.S. Chen[2,3], Y. C. Hon[2,4], M. Li[2]

[1] School of Engineering and Materials Science, Queen Mary, University of London, London E1 4NS, UK
[2] College of Mathematics, Taiyuan University of Technology, Taiyuan, China
[3] Department of Mathematics, University of Southern Mississippi, Hattiesburg, MS 39406, USA
[4] Department of Mathematics, City University of Hong Kong, Hong Kong SAR, China

*Corresponding author: p.h.wen@qmul.ac.uk

Abstract

Based on the recently developed Finite Integration Method (FIM) for solving one-dimensional ordinary and partial differential equations, this paper extends the technique to higher dimensional partial differential equations. The main idea is to extend the first order finite integration matrices constructed by using either Ordinary Linear Approach (OLA) (uniform distribution of nodes) or Radial Basis Function (RBF) interpolation (uniform/random distributions of nodes) to higher order integration matrices. Illustrative two-dimensional numerical examples are given in two-dimension to compare the FIM (FIM-OLA and FIM-RBF) with the Finite Difference Method and Point Collocation Method to demonstrate its superior accuracy and efficiency.

Keywords: finite integration method, radial basis functions, partial differential equation.

Introduction

Mathematical models in terms of partial differential equations (PDEs) have commonly been used to describe a wide variety of physical phenomena such as sound, heat, electrostatics, electrodynamics, fluid flow, and elasticity. Under various boundary conditions, it is very rare that these models can be solved in closed form solutions. Numerical methods are unavoidable for seeking approximate solutions to simulate the dynamic and characteristics of the models. Due to the advance of computational methods, these kinds of numerical approximation can usually be achieved inexpensively to high accuracy together with a reliable bound on the error between the analytical solution and its numerical approximation. There are many numerical techniques available for solving differential equations [Lambert (1991), Hairer (1993)] including the Finite Element Method (FEM) and Boundary Element Method (BEM). In the last decade, the development of the Radial Basis Functions (RBFs) as a truly meshless method has drawn attention from many researchers. In particular, the use of multiquadric radial basis function (MQ-RBF) [Hardy (1971), Goldberg and Chen (1997), Hon and Mao (1997), Atluri (2002), Liu (2003),] has shown the superior convergence of the method in comparing with FEM and BEM. Numerical results indicated that these meshless methods provide a similar optimal accuracy for solving both elliptic and parabolic equations in 2D. Recently, Wen et al (2013) and Li et al (2013) developed a Finite Integration Method (FIM) for solving differential equation in 1D and demonstrated its applications to nonlocal elasticity problems. It has been shown that the FIM gives higher degree of accuracy than the Finite Difference Method (FDM) and Point Collocation Method (PCM). In this paper, the FIM is further extended to solve multi-dimensional partial differential equations. Two-dimensional partial differential equations are given in illustrative examples. Similar to the FDM and the PCM, a finite number of points, known as field points, are distributed in the computational domain. The field points are generated either uniformly (grid) along the independent coordinate or randomly in the domain. The integration matrix of the first order is obtained by the direct integration with either OLA approximation. Based on these first order integration matrices, any finite integration matrix

with multi-layer integration can easily be obtained. To compare with other numerical methods, the PCM and analytical solution are used.

FIM for one-dimensional problems

Numerical quadrature rule based on Ordinary Linear Approach (OLA) is the simplest computational scheme for integration [see Wen et al (2013)]. Starting from one-dimension problem, an integral of a given function $u(x)$ can be written as

$$U(x) = \int_0^x u(\xi)d\xi \qquad (1)$$

Applying the linear interpolation technique to Eq. (1), we have

$$U(x_k) = \int_0^{x_k} u(\xi)d\xi = \sum_{i=1}^k a_{ki}u(x_i) \qquad (2)$$

where, using trapezoidal rule,

$$a_{1i} = 0 \qquad (3)$$

$$a_{ki} = \begin{cases} \Delta, & i=1, \\ 0.5\Delta, & i=2,3,...,k-1, \\ 0.5\Delta, & i=k, \\ 0, & i>k, \end{cases} \qquad (4)$$

and $x_i = \Delta*(i-1)$, $\Delta = b/(N-1)$, $i=1,2,..,N$ are nodal points in [0, b], and $x_1 = 0, x_N = b$. Note that Eq. (2) can be written in a matrix form as

$$\mathbf{U} = \mathbf{A}\mathbf{u} \qquad (5)$$

where $\mathbf{U} = [U_1, U_2,...,U_N]^T$ $\mathbf{u} = [u_1, u_2,...,u_N]^T$, the first order integration matrix

$$\mathbf{A} = (a_{ki}) = \Delta \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 & 0 & 0 \\ 1/2 & 1 & 1/2 & 0 & 0 & 0 \\ 1/2 & 1 & 1 & 1/2 & 0 & 0 \\ ... & ... & ... & ... & ... & ... \\ 1/2 & 1 & 1 & 1 & 1 & 1/2 \end{pmatrix}_{N\times N}$$

.

and $U_i = U(x_i)$, $u_i = u(x_i)$ are the values of integration and the integral function respectively at each nodes. Thereafter, consider a multi-integral for one-dimensional problem

$$U^{(2)}(x) = \int_0^x \int_0^\zeta u(\xi)d\xi d\zeta, \qquad x \in [0,b]. \qquad (6)$$
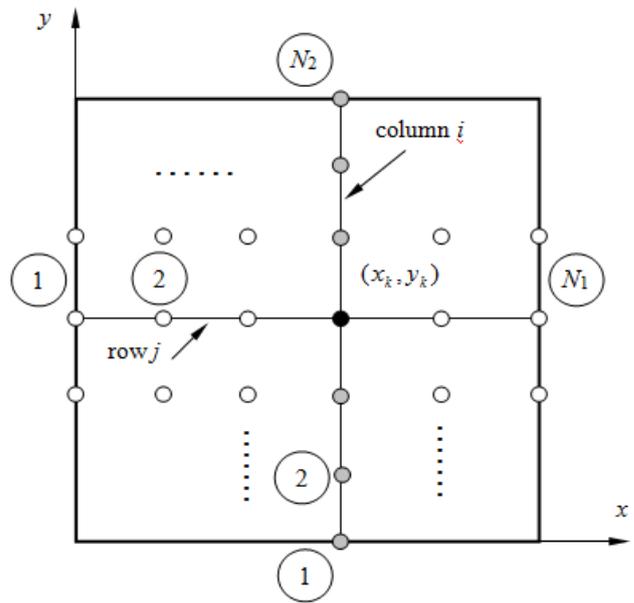
Figure 1. Uniform distribution of collocation points.

$$U^{(2)}(x_k) = \int_0^{x_k}\int_0^\zeta u(\xi)d\xi d\zeta = \sum_{i=0}^k\sum_{j=0}^i a_{ki}a_{ij}u(x_i) = \sum_{i=0}^k a_{ki}^{(2)}u(x_i) \qquad (7)$$

.

The above multi-integral can also be written in a matrix form as

$$\mathbf{U}^{(2)} = \mathbf{A}^{(2)}\mathbf{u} = \mathbf{A}^2\mathbf{u} \qquad (8)$$

where

$$\mathbf{A}^{(2)} = \mathbf{AA} = \Delta^2 \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 1/4 & 1/4 & 0 & 0 & 0 & 0 \\ 3/4 & 1 & 1/4 & 0 & 0 & 0 \\ 5/4 & 2 & 1 & 1/4 & 0 & 0 \\ ... & ... & ... & ... & ... & ... \\ [1+2(N-1)]/4 & N-1 & N-2 & ... & 1 & 1/4 \end{pmatrix}_{N \times N} \qquad (9)$$

and the elements of matrix $\mathbf{A}^{(2)}$ are

$$a_{1i}^{(2)} = 0 \qquad (10)$$

$$a_{ki}^{(2)} = \begin{cases} [1+2(k-2)]\Delta^2/4, & i=1, \\ (k-i)\Delta^2, & i=2,3,...,k-1, \\ \Delta^2/4, & i=k, \\ 0, & i>k. \end{cases}$$

For two-dimensional problems, let us consider a uniform distribution of collocation points as shown in Fig. 1. Similar to Eq. (1), we define

$$U_x(x,y) = \int_0^x u(\xi,y)d\xi, U_x(x_k,y_k) = \int_0^{x_k} u(\xi,y_k)d\xi \qquad (11)$$

and the total number of point is $k = N_1(j-1)+i$, where i and j denote the number of column and the number of row respectively. This numbering system is called the global number system. We can also express each nodal value of integration in Eq. (11) in a matrix form as

$$\mathbf{U}_x = \mathbf{A}_x \mathbf{u} \qquad (12)$$

where integral nodal value $U_x = [U_{x1}, U_{x2}, ..., U_{xM}]^T$, nodal value $\mathbf{u} = [u_1, u_2, ..., u_M]^T$ and M is the total number of collocation points ($M = N_1 \times N_2$ for grid shown in Fig. 1). For a rectangular domain, the first order integration matrix

$$\mathbf{A}_x = \begin{pmatrix} \mathbf{A} & 0 & ... & 0 \\ 0 & \mathbf{A} & 0 & 0 \\ ... & ... & ... & ... \\ 0 & 0 & 0 & \mathbf{A} \end{pmatrix}_{N_2} \qquad (13)$$

in which, $\mathbf{A}$ is integration matrix for one-dimension given in Eq. (5) with dimension $N_1 \times N_1$. Similarly, the integration along y axis is

$$U_y(x,y) = \int_0^y u(x,\eta)d\eta, \quad U_y(x_k,y_k) = \int_0^{y_k} u(x_k,y)dy \qquad (14)$$

which can be written in the matrix form as

$$\mathbf{U}_y = \mathbf{A}'\mathbf{u} \qquad (15)$$

in the local system for the collocation points, where $k = N_2(i-1)+j$. The first order integration matrix in the local system is

$$\mathbf{A}' = \begin{pmatrix} \mathbf{A} & 0 & ... & 0 \\ 0 & \mathbf{A} & 0 & 0 \\ ... & ... & ... & ... \\ 0 & 0 & 0 & \mathbf{A} \end{pmatrix}_{N_2} \qquad (16)$$

in which $\mathbf{A}$ is the integration matrix for one-dimension integral given in Eq. (5) with dimension $N_2 \times N_2$. By a simple re-arrangement of the number of the nodes, Eq. (15) can be rewritten, in the global system, as

$$\mathbf{U}_y = \mathbf{A}_y \mathbf{u}. \qquad (17)$$

For the multi-integration in two-dimensional problem in a rectangular domain, we consider the following integral with respect to coordinate x

$$U_x^{(2)}(x_i,y) = \int_0^{x_i}\int_0^{\zeta} u(\xi,y)d\xi d\zeta, \qquad x_i \in [0,b_1],\ y \in [0,b_2] \qquad (18)$$

and use the same procedure for one-dimension, one has

$$U^{(2)}(x_k,y_k) = \int_0^{x_k}\int_0^{\zeta} u(\xi,y_k)d\xi d\zeta = \sum_{i=0}^{k}\sum_{j=0}^{i}(a_{kj})_x(a_{ji})_x u_i \qquad (19)$$

or in a matrix form

$$\mathbf{U}_x^{(2)} = \mathbf{A}_x^2 \mathbf{u} \qquad (20)$$

where

$$\mathbf{A}_x^2 = \mathbf{A}_x \mathbf{A}_x = \begin{pmatrix} \mathbf{A}^2 & 0 & ... & 0 \\ 0 & \mathbf{A}^2 & 0 & 0 \\ ... & ... & ... & ... \\ 0 & 0 & 0 & \mathbf{A}^2 \end{pmatrix}_{\underbrace{\qquad\qquad}_{N_2}} \tag{21}$$

Similarly, one has multi-integration $U_y^{(2)}(x,y)$ with respect to coordinate y

$$U_y^{(2)}(x_k, y_k) = \int_0^{y_k}\int_0^{\tau} u(x_k, \eta)\, d\eta\, d\tau = \sum_{i=0}^{k}\sum_{j=0}^{i}(a_{kj})_y (a_{ji})_y u_i \tag{22}$$

and

$$\mathbf{U}_y^{(2)} = \mathbf{A}_y^2 \mathbf{u}. \tag{23}$$

This method can be extended to the higher order integrations, i.e.

$$U_x^{(m)}(x_k, y_k) = \underbrace{\int_0^{x_k}...\int_0^{...} u(\xi_1, y_k)\, d\xi_1...d\xi_m}_{m\ \text{layers}}, \quad U_y^{(m)}(x_k, y_k) = \underbrace{\int_0^{y_k}...\int_0^{...} u(x_k, \eta_1)\, d\eta_1...d\eta_m}_{m\ \text{layers}} \tag{24}$$

$$x_k \in [0, b_1], y_k \in [0, b_2]$$

Applying ordinary linear interpolation technique again for integral function $U^{(m)}(x,y)$, we have

$$U_x^{(m)}(x_k, y_k) = \int_0^{x_k}...\int_0^{...} u(\xi_1, y_k)\, d\xi_1...d\xi_1 = \sum_{i=0}^{M}...\sum_{j=0}^{M}(a_{kj})_x...(a_{ni})_x u(x_i, y_i) = \sum_{i=0}^{M}\left(a_{ki}^{(m)}\right)_x u_i \tag{25}$$

$$U_y^{(m)}(x_k, y_k) = \int_0^{y_k}...\int_0^{...} u(x_k, \eta_1)\, d\eta_1...d\eta_m = \sum_{i=0}^{M}...\sum_{j=0}^{M}(a_{kj})_y...(a_{ni})_y u(x_i, y_i) = \sum_{i=0}^{M}\left(a_{ki}^{(m)}\right)_y u_i \tag{26}$$

Again, it can also be written, in a matrix form, as

$$\mathbf{U}_x^{(m)} = \mathbf{A}_x^m \mathbf{u}, \quad \mathbf{U}_y^{(m)} = \mathbf{A}_y^m \mathbf{u} \tag{27}$$

In addition, this method can be extended to multi-layers integration with two coordinates x and y as follow:

$$U^{(mn)}(x_k, y_k) = \int_0^{x_k}...\int_0^{...}\int_0^{y_k}...\int_0^{...} u(\xi, \eta)\underbrace{d\xi_1...d\xi_n}_{n-\text{layer}}\underbrace{d\eta_1...d\eta_m}_{m-\text{layer}} \qquad x_k \in [0, b_1],\ y_k \in [0, b_2] \tag{28}$$

and the nodal values of the above integration are obtained in the matrix form as

$$\mathbf{U}^{(mn)} = \mathbf{A}_x^m \mathbf{A}_y^n \mathbf{u}. \tag{29}$$

FIM with radial basis functions

For uniform distribution of nodes (grid), the multi-layer integrations at each node can be obtained quite easily in a matrix form. However, in general case, if the nodes distribution is random, the algorithm OLA discussed in the Section above is not valid. In this case, interpolation schemes have to be introduced. Recently, the radial basis functions interpolation schemes and moving least square method are very popular meshless methods. For example, the MQ-RBF was introduced by Hardy (1971) for the interpolation of topographical surfaces in the early stage of radial bases function application. Note that $u(\mathbf{x})$ in the domain Ω can be interpolated over a number of randomly distributed nodes $\mathbf{x}_i = (x_i, y_i)$, $i = 1, 2, ..., M$, as

$$u(\mathbf{x}) = \sum_{i=1}^{M} R_i(\mathbf{x}, \mathbf{x}_i)\alpha_i + \sum_{q=1}^{Q} P_q(\mathbf{x})\beta_q = \mathbf{R}(\mathbf{x})\boldsymbol{\alpha} + \mathbf{P}(\mathbf{x})\boldsymbol{\beta} = \sum_{i=1}^{M} \varphi_i(\mathbf{x})u_i \quad \mathbf{x} \in \Omega, \tag{30}$$

where $\mathbf{R}(\mathbf{x}) = \left[R_1(\mathbf{x}, \mathbf{x}_1), R_2(\mathbf{x}, \mathbf{x}_2), ..., R_M(\mathbf{x}, \mathbf{x}_M) \right]$ is a set of radial basis functions centred at $\mathbf{x} = (x, y)$, $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, \cdots, \alpha_M]^T$ and $\boldsymbol{\beta} = [\beta_1, \beta_2, \cdots, \beta_Q]^T$ are the coefficients to be determined, $\varphi_i(\mathbf{x})$ is shape function. Therefore, the integration matrices of the first order are

$$\mathbf{A}_x = \begin{pmatrix} \bar{\phi}_{x1}(\mathbf{x}_1) & \bar{\phi}_{x2}(\mathbf{x}_1) & ... & \bar{\phi}_{xM}(\mathbf{x}_1) \\ \bar{\phi}_{x1}(\mathbf{x}_2) & \bar{\phi}_{x2}(\mathbf{x}_2) & ... & \bar{\phi}_{xM}(\mathbf{x}_2) \\ ... & ... & ... & ... \\ \bar{\phi}_{x1}(\mathbf{x}_M) & \bar{\phi}_{x2}(\mathbf{x}_M) & ... & \bar{\phi}_{xM}(\mathbf{x}_M) \end{pmatrix}_{M \times M} \text{ and } \mathbf{A}_y = \begin{pmatrix} \bar{\phi}_{y1}(\mathbf{x}_1) & \bar{\phi}_{y2}(\mathbf{x}_1) & ... & \bar{\phi}_{yM}(\mathbf{x}_1) \\ \bar{\phi}_{y1}(\mathbf{x}_2) & \bar{\phi}_{y2}(\mathbf{x}_2) & ... & \bar{\phi}_{yM}(\mathbf{x}_2) \\ ... & ... & ... & ... \\ \bar{\phi}_{y1}(\mathbf{x}_M) & \bar{\phi}_{y2}(\mathbf{x}_M) & ... & \bar{\phi}_{yM}(\mathbf{x}_M) \end{pmatrix}_{M \times M}.$$

where $\bar{\phi}_{xi} = \int \phi_i(\mathbf{x})dx$, $\bar{\phi}_{yi} = \int \phi_i(\mathbf{x})dy$.

The FIM for multi-dimensional problems

The FIM is readily extendable to solving higher dimensional problems. For illustration, consider the following two-dimensional partial differential equation

$$\alpha_1(x, y)\frac{\partial^2 u}{\partial x^2} + \alpha_2(x, y)\frac{\partial^2 u}{\partial y^2} + \alpha_3(x, y)u = b(x, y), \qquad \mathbf{x} \in \Omega, \tag{31}$$

$$\Lambda[u(x, y)] = h(x, y), \qquad\qquad\qquad \mathbf{x} \in \partial\Omega,$$

where Λ is a boundary operator, $\alpha_1(x, y), \alpha_2(x, y), \alpha_3(x, y), b(x, y)$ and $h(x, y)$ are given functions. u is generally referred as potential, which represents the transversal displacement of a membrane. Ω and $\partial\Omega$ are simple connected domain and its boundary respectively. Integrating twice in Eq. (31) with respect to coordinates x and y respectively, one has

$$\iint \iint \left[\alpha_1(x, y)\frac{\partial^2 u}{\partial x^2} + \alpha_2(x, y)\frac{\partial^2 u}{\partial y^2} + \alpha_3(x, y)u \right] dxdxdydy = \iint \iint b(x, y)dxdxdydy + \tag{32}$$

$$xf_0(y) + f_1(y) + yg_0(x) + g_1(x)$$

where $f_0(y), f_1(y), g_0(x)$ and $g_1(x)$ are unknown one-dimensional functions. Using the technique of integration by part, we have

$$\iint \left[\alpha_1 u - 2\int u\frac{\partial \alpha_1}{\partial x}dx + \iint u\frac{\partial^2 \alpha_1}{\partial x^2}dxdx \right] dydy + \iint \left[\alpha_2 u - 2\int u\frac{\partial \alpha_2}{\partial y}dx + \iint u\frac{\partial^2 \alpha_2}{\partial y^2}dydy \right] dxdx + \tag{33}$$

$$\iint \iint \alpha_3 u dxdxdydy = \iint \iint b(x, y)dxdxdydy + xf_0(y) + f_1(y) + yg_0(x) + g_1(x).$$

By using integration matrix mentioned in the previous sections, we have

$$\left[\mathbf{A}_y^2 \boldsymbol{\alpha}_1 + \mathbf{A}_x^2 \boldsymbol{\alpha}_2 - 2\mathbf{A}_x \mathbf{A}_y \left(\mathbf{A}_y \boldsymbol{\alpha}_{1,x} + \mathbf{A}_x \boldsymbol{\alpha}_{2,y} \right) + \mathbf{A}_x^2 \mathbf{A}_y^2 \left(\boldsymbol{\alpha}_{1,xx} + \boldsymbol{\alpha}_{2,yy} + \boldsymbol{\alpha}_3 \right) \right] \mathbf{u} = \tag{34}$$

$$\mathbf{A}_y^2 \mathbf{A}_x^2 \mathbf{b} + \mathbf{X}\boldsymbol{\Psi}_y \mathbf{f}_0 + \boldsymbol{\Psi}_y \mathbf{f}_1 + \mathbf{Y}\boldsymbol{\Psi}_x \mathbf{g}_0 + \boldsymbol{\Psi}_x \mathbf{g}_1$$

where $\mathbf{X} = \{x_1, x_2, ..., x_M\}$, $\mathbf{Y} = \{y_1, y_2, ..., y_M\}$, $\mathbf{f}_0 = \left[f_0^1, f_0^2, ..., f_0^r \right]^T$, $\mathbf{f}_1 = \left[f_1^1, f_1^2, ..., f_1^r \right]^T$, $\mathbf{g}_0 = \left[g_0^1, g_0^2, ..., g_0^p \right]^T$, $\mathbf{g}_1 = \left[g_1^1, g_1^2, ..., g_1^p \right]^T$, p and r are numbers of point to be used for interpolation of functions $f(y)$ and $g(x)$ respectively, $\boldsymbol{\Psi}_x$ and $\boldsymbol{\Psi}_y$ are matrices of one-dimensional shape functions with respect to coordinates x and y respectively, and

$$\boldsymbol{\alpha}_l = \begin{pmatrix} \alpha_l(\mathbf{x}_1) & 0 & ... & 0 \\ 0 & \alpha_l(\mathbf{x}_2) & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & \alpha_l(\mathbf{x}_M) \end{pmatrix} \qquad l = 1,2,3 \tag{35}$$

Integral functions $f_0(y), f_1(y), g_0(x)$ and $g_1(x)$ can be interpolated in terms of the nodal values in the following procedure:

(1) Determine the regions of functions $f(y)$ and $g(x)$, i.e. $[\bar{y}_1, \bar{y}_r]$, $[\bar{x}_1, \bar{x}_p]$, and uniformly distributed points in these regions as shown in Fig. 2;

(2) Determine one-dimensional shape function matrices $\mathbf{\Psi}_x$ and $\mathbf{\Psi}_y$

By using linear interpolation, one has

$$f(y) = 0 \quad \text{if } \bar{y}_1 < y < \bar{y}_{m-1} \text{ or } \bar{y}_m < y < \bar{y}_r$$
$$g(x) = 0 \quad \text{if } \bar{x}_1 < x < \bar{x}_{n-1} \text{ or } \bar{x}_n < x < \bar{x}_p$$
(36)

$$f(y) = \frac{\bar{y}_m - y}{\bar{y}_m - \bar{y}_{m-1}} f_{m-1} + \frac{y - \bar{y}_{m-1}}{\bar{y}_m - \bar{y}_{m-1}} f_m \quad \text{if } \bar{y}_{m-1} < y < \bar{y}_m$$

$$g(x) = \frac{\bar{x}_n - x}{\bar{x}_n - \bar{x}_{n-1}} g_{n-1} + \frac{x - \bar{x}_{n-1}}{\bar{x}_n - \bar{x}_{n-1}} g_n \quad \text{if } \bar{x}_{n-1} < x < \bar{x}_n$$
(37)

Therefore, the matrices of shape function are

$$\mathbf{\Psi}_x = \begin{pmatrix} \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \frac{\bar{x}_n - x_i}{\bar{x}_n - \bar{x}_{n-1}} & \frac{x_i - \bar{x}_{n-1}}{\bar{x}_n - \bar{x}_{n-1}} & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \end{pmatrix} \begin{matrix} \text{row 1} \\ \\ i \\ \\ M \end{matrix}$$
$$\begin{matrix} \text{columm 1} & & n\text{-}1 & n & & p \end{matrix}$$
(38)

and

$$\mathbf{\Psi}_y = \begin{pmatrix} \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \frac{\bar{y}_m - y_i}{\bar{y}_m - \bar{y}_{m-1}} & \frac{y_i - \bar{y}_{m-1}}{\bar{y}_m - \bar{y}_{m-1}} & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \end{pmatrix} \begin{matrix} \text{row 1} \\ \\ i \\ \\ M \end{matrix}$$
$$\begin{matrix} \text{columm 1} & & m\text{-}1 & m & & r \end{matrix}$$
(39)

in which $\psi_i(x)$ and $\psi_i(y)$ are shape functions in one dimensional case as shown in Fig. 2. In Eq. (34), we have M nodal unknowns of $\mathbf{u}$, $2q$ unknowns of $\mathbf{f}_0$, $\mathbf{f}_1$ and $2r$ unknowns of $\mathbf{g}_0$, $\mathbf{g}_1$.

For a rectangular plate with uniform distribution of nodes $(N_1 \times N_2)$, obviously one has $2p + 2r$ nodes located on the boundary. By selecting $p = N_1 - 1$ and $r = N_2 - 1$ for uniform distribution of node, there are $N_1 \times N_2 + 2(N_1 + N_2 - 2)$ linear system of equations to determine all unknowns, i.e. $\mathbf{u}$, $\mathbf{f}_0$, $\mathbf{f}_1$, $\mathbf{g}_0$, and $\mathbf{g}_1$. In fact, the number of boundary points to determine four one-dimensional integral functions is arbitrary. The number of points (L) on the boundary should be greater than or equal to $2(p+r)$. If $L = 2(p+r)$, the standard Gaussian solver can be used directly. Otherwise, the Singular Value Decomposition [Press et al (1992)] scheme should be introduced.

Figure 2. Interpolations for one dimensional function $f(y)$ and $g(x)$.

Numerical example

Consider the following partial differential equation

$$x(1-x)\frac{\partial^2 u}{\partial x^2} + y(1-y)\frac{\partial^2 u}{\partial y^2} = -4xy(1-x)(1-y), \quad (x,y) \in \Omega,$$
$$u(x,y) = 0, \quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad (x,y) \in \partial\Omega,$$
(41)

where $\Omega \cup \partial\Omega = [0,1] \times [0,1]$. The analytical solution is given by $u^*(x,y) = xy(1-x)(1-y)$. The average relative error is defined as

$$\varepsilon = \frac{1}{M} \sum_{i=1}^{M} \frac{|u_i - u_i^*|}{|u_{\max}^*|}, \quad u_{\max}^* = u^*(0.5, 0.5) = \frac{1}{16}$$
(42)

Figure 3. Distributions of nodes.

Uniformly distributed nodes is shown in Fig.3. For radial basis functions approach in Eq. (30), three radial basis functions are considered, i.e.

(1) MQ function: $R(r) = \sqrt{c^2 + r^2}$;

(2) Linear function (LF): $R(r) = r$;

(3) Thin-Plate Splines (TPS): $R(r) = r^2 \ln r$.

In this example, we chose $N_1 = N_2 = N$ and $p = r = N$. The shape parameter c of MQ is selected as $c = 1 / N$. For the RBF approach, single integration matrix in Eq. (70) is used. The average errors ε for various number of collocation point are shown in Table 1. Among these algorithms, the accuracy of OLA is the lowest and PSF of radial basis function is the highest.

Table 1. Average errors (ε).

N	OLA	MQF	LF	TPS
10	0.019110	0.013707	0.017554	0.012985
20	0.005689	0.005365	0.006462	0.004496
30	0.013820	0.003083	0.003671	0.002373

Conclusion

In this paper, the Finite Integration Method (FIM) with Ordinary Linear Approach and Radial Basis Functions interpolation was extended to solve multi-dimensional differential equations. Compared with the Point Collocation Method (PCM) and the Finite Difference Method, the proposed FIM performs much superior in accuracy and stability. For the FIM with Radial Basis Functions interpolation, the use of randomly distributed nodes in the domain allows solving problems under irregular domains.

References

Atluri, S.N. (2002) *The meshless local Petrov-Galerkin (MLPG) method*, Tech. Science Press.
Goldberg, M.A., Chen, C.S. and Karur, S.R. (1996) Improved multiquadric approximation for partial differential equations, *Engng Anal with Boundary Elements* **18**, 9-17.
Hairer, E. (1993) *Solving ordinary differential equations, Volumes 1 and 2*, Springer-Verlag, Berlin-New York.
Hardy, R.L. (1971) Multiquadric equations of topography and other irregular surface, *J. Geophys. Res*. **76**, 1905-1915.
Hon Y.C. and Mao, X.Z. (1997) A multiquadric interpolation method for solving initial value problems, *J. Scientific Computing* **12**, 51-55.
Lambert J.D. (1991) *Numerical methods for ordinary differential systems: the initial value problem*, John Wiley & Sons, New York.
Li, M. Hon, Y.C., Korakianitis, T. and Wen, P.H. (2013) Finite integration method for nonlocal elastic bar under static and dynamic loads, *Engng Anal with Boundary Elements*, **37**, (5) 842-849.
Liu, G. R. (2009) *Meshfree Methods: Moving Beyond the Finite Element Method*, 2nd edn, CRCPress, Boca Raton,USA.
Press, W.H., Teukolsky, S.A., Vetterling, W.T. and Flannery, B.P. (1992) *Numerical Recipes in Fortran*, Cambridge University Press: Cambridge, UK.
Wen, P.H., Hon, Y.C. Li, M. and Korakianitis, T. (2013) Finite integration method for partial differential equations, *Applied Mathematic Modelling* **37**, 10092-10106.

Identification of a Position and Time Dependent Heat Flux Using the Kalman Filter and Improved Lumped Analysis in Heat Conduction

*C.C. Pacheco[1], H.R.B. Orlande[1], M.J. Colaço[1], and G.S. Dulikravich[2]

[1]Department of Mechanical Engineering, Federal University of Rio de Janeiro, Brazil
[2]Department of Mechanical and Materials Engineering, FIU-MAIDROC Laboratory, Miami, USA

*Corresponding author: cesar.pacheco@poli.ufrj.br

Abstract

This paper aims to estimate a position and time dependent heat flux with high magnitude in a heat conduction problem. The heat flux is applied on one side of a flat plate, while the inverse problem is solved by using temperature measurements taken on the opposite side. The proposed forward problem is a surrogate model, derived from the simplification of a complete model. The inverse problem is then solved with the Kalman Filter. The temperature at the surface of the plate is approximated by using the improved lumped analysis, where the temperature gradients across the thickness of the plate are accounted for in an approximate manner. The measurements are simulated with the complete model, while the inverse problem is solved with the surrogate model. The temperature estimates show a good agreement with reference values.

Keywords: Inverse Problems, Heat Conduction, Kalman Filter, Improved Lumped Analysis

Introduction

Despite the modern and reliable available techniques for measuring temperature and heat flux, some particular scenarios are still challenging. Situations involving complex geometries or hazardous environments might make direct measurements of these quantities impractical [Dennis and Dulikravich, 2001]. Thus, estimation of these unknowns by using inverse analysis with temperature and/or heat flux measurements taken at other regions of the body of interest should be considered as a possible solution. Situations of this type become increasingly common, for example, due to the recent development of powerful microprocessors, which dissipates high amounts of heat. Techniques for dealing with such thermal loads are available, but new methodologies for proper quantification and more efficient cooling of these thermal loads are desired. Some results on the estimation of a high magnitude boundary heat flux in a heat conduction problem can be found in the literature [Dennis and Dulikravich, 2001; Feng et al., 2011; Dennis and Dulikravich, 2012; Afrin et al., 2013; Orlande et al., 2013]. All of these works emphasize the difficulties of solving the inverse problem with an accurate mathematical model, which would be a three-dimensional nonlinear heat conduction problem, thus resulting in high computational times [Dennis and Dulikravich, 2012]. In this work, the proposed forward problem is obtained by simplification of a more general heat conduction problem which, together with the modeling of uncertainties of observations and unknowns as Gaussians, allows one to use the Kalman filter [Kalman, 1960; Chen, 2003; Kaipio and Somersalo, 2004; Grewal and Andrews, 2008]. The physical problem considered in this paper involves heat conduction in a flat plate, where the temperatures at both surfaces are approximated by the Improved Lumped System Analysis [Cotta and Mikhailov, 1997]. In this formulation, the temperature gradient across the plate is approximated by Hermite's formulae. The use of the Kalman filter requires much less computational effort in comparison with techniques such as particle filters and is more readily adaptable to parallel processing.

Forward Problem

The physical problem considered in this paper involves a high magnitude heat flux applied to the top surface of a flat plate, while temperature measurements are taken at the opposite side, as shown in Fig. 1. The dimensions of the flat plate are given by Tab. 1.
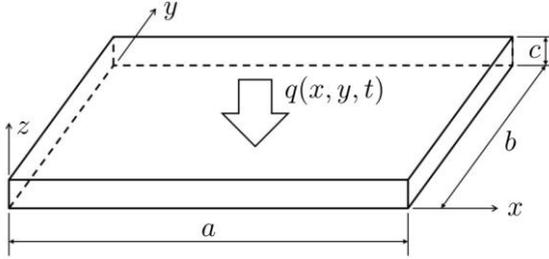
Figure 1: Geometry of the physical problem

Table 1: Dimensions of the flat plate

Dimension	Value [mm]
a	120
b	120
c	3

All other boundaries are thermally insulated. The heat flux is position-and-time dependent and the initial temperature distribution is considered to be uniform. Based on these assumptions, the resulting mathematical model [Ozisik, 1993], named "Complete Model", is given by Eqs. (1.a)-(1.f).

$$C(T_c)\frac{\partial T_c}{\partial t} = \frac{\partial}{\partial x}\left[k_T(T_c)\frac{\partial T_c}{\partial x}\right] + \frac{\partial}{\partial y}\left[k_T(T_c)\frac{\partial T_c}{\partial y}\right] + \frac{\partial}{\partial z}\left[k_T(T_c)\frac{\partial T_c}{\partial z}\right] \text{ in } \begin{matrix} 0 \le x \le a, & 0 \le y \le b \\ 0 \le z \le c, & t \ge 0 \end{matrix} \tag{1.a}$$

$$\frac{\partial T_c}{\partial x} = 0 \quad \text{at} \quad x = 0 \quad \text{and} \quad x = a \tag{1.b}$$

$$\frac{\partial T_c}{\partial y} = 0 \quad \text{at} \quad y = 0 \quad \text{and} \quad y = b \tag{1.c}$$

$$\frac{\partial T_c}{\partial z} = 0 \quad \text{at} \quad z = 0 \tag{1.d}$$

$$k_T(T_c)\frac{\partial T_c}{\partial z} = q(x,y,t) \quad \text{at} \quad z = c \tag{1.e}$$

$$T_c = T_0 \quad \text{at} \quad t = 0 \tag{1.f}$$

Since high temperature variations are expected, the volumetric heat capacity and the thermal conductivity are supposed to vary with respect to the temperature according to [Orlande et al., 2013]:

$$C(T) = 1324.75T + 3557900 \text{ [J/m}^3\text{]} \tag{2.a}$$

$$k_T(T) = 12.45 + 0.014T + 2.517 \times 10^{-6}T^2 \text{ [W/mK]} \tag{2.b}$$

The Kalman filter cannot be used to solve the inverse problem related to the estimation of the applied heat flux using this mathematical model, since it is non-linear. Regarding other techniques, a similar inverse problem solved with the complete model, using the Metropolis-Hastings algorithm in a time range of 2.0 seconds with $\Delta x = \Delta y = 5$mm, $\Delta z = 0.5$mm, $\Delta t = 0.01$s and 10^5 states of the Markov Chain, led to 8 days of computational time [Orlande et al., 2013]. In order to reduce this extremely high computational cost, a surrogate model is proposed in this paper as described below.

Surrogate Model

The first step to obtain the surrogate model is to linearize the thermal properties, evaluating Eqs. (2.a) and (2.b) at a reference temperature $T^* = 600K$. This gives rise to the constant thermal properties presented on Eq. (3).

$$C^* = C(T^*) \quad \text{and} \quad k_T^* = k_T(T^*) \tag{3}$$

The next step aims to reduce the number of dimensions of the model. This is achieved by calculating the mean temperature in the z direction, using the operator described in Eq. (4).

$$\overline{T}(x, y, t) = \frac{1}{c} \int_0^c T(x, y, z, t) dz \tag{4}$$

Application of this operator in Eq. (1.a) is straightforward, except for the diffusion term in the z-direction, where the result is the heat flux at $z = 0$ and $z = c$ surfaces of the plate. This result can be combined with the linearized versions of Eq. (1.d) and (1.e), as shown in Eq. (5).

$$\frac{1}{c} \int_0^c \frac{\partial}{\partial z} \left[k_T^* \frac{\partial T}{\partial z} \right] dz = \frac{1}{c} k_T^* \frac{\partial T}{\partial z} \Big|_0^c = \frac{q(x, y, t)}{c} \tag{5}$$

Operation of the linearized versions of Eqs. (1.b), (1.c) and (1.f), is also straightforward. The final result is the following surrogate model, which is a linear two dimensional problem:

$$C^* \frac{\partial \overline{T}}{\partial t} = k_T^* \frac{\partial^2 \overline{T}}{\partial x^2} + k_T^* \frac{\partial^2 \overline{T}}{\partial y^2} + \frac{q(x, y, t)}{c} \quad \text{in} \quad \begin{array}{l} 0 \le x \le a, \quad t \ge 0 \\ 0 \le y \le b \end{array} \tag{6.a}$$

$$\frac{\partial \overline{T}}{\partial x} = 0 \quad \text{in} \quad x = 0 \text{ and } x = a \tag{6.b}$$

$$\frac{\partial \overline{T}}{\partial y} = 0 \quad \text{in} \quad y = 0 \text{ and } y = b \tag{6.c}$$

$$\overline{T} = T_0 \quad \text{in} \quad t = 0 \tag{6.d}$$

This model is much simpler to solve than the complete model. However, its solution leads to the mean temperature in the z-direction, but the desired quantity is the temperature at the z = 0 surface. The Improved Lumped Analysis [Cotta and Mikhailov, 1997] allows one to approximate this quantity by using the Hermite's formulas for integrals given by:

$$\int_0^h y(x) dx = \frac{h}{2} [y(0) + y(h)] + O(h^3) \tag{7.a}$$

$$\int_0^h y(x) dx = \frac{h}{2} [y(0) + y(h)] + \frac{h^2}{12} \left[\frac{dy}{dx} \Big|_{x=0} - \frac{dy}{dx} \Big|_{x=h} \right] + O(h^5) \tag{7.b}$$

These formulas are used to approximate the mean temperature in the z-direction and the integral of the temperature gradient in the z-direction, that is,

$$\overline{T}(x, y, t) \approx \frac{1}{2} [T(x, y, 0, t) + T(x, y, c, t)] + \frac{c}{12} \left[\frac{\partial T}{\partial z} \Big|_{z=0} - \frac{\partial T}{\partial z} \Big|_{z=c} \right] \tag{8.a}$$

$$\int_0^c \frac{\partial T(x,y,z,t)}{\partial z} dz = T(x,y,c,t) - T(x,y,0,t) \approx \frac{c}{2}\left[\left. \frac{\partial T}{\partial z}\right|_{z=0} + \left. \frac{\partial T}{\partial z}\right|_{z=c} \right] \tag{8.b}$$

The final result is an approximation of the temperature at the $z = 0$ surface given by:

$$T(x,y,0,t) \approx \overline{T}(x,y,t) - \frac{c}{6k_T^*} q(x,y,t) \tag{9}$$

Inverse Problem

The inverse problem related to the estimation of the applied high intensity heat flux is solved in this paper by using a Bayesian approach. A probability density function (pdf) of the unknown state variables $\mathbf{x}_n$ given the set of observations $\mathbf{y}_{0:n}$ is built with Bayes' Theorem. Statistical inference techniques can be applied to this pdf, called "posterior", to extract information about the unknowns [Chen, 2003]. This work uses the Kalman filter, which requires the forward problem to be cast in the form of the Evolution-Observed Model given by Eqs. (10.a) and (10.b), where $\mathbf{w}_n$ and $\mathbf{v}_n$ are zero mean Gaussian noise vectors, with covariance matrices $\mathbf{Q}_n$ and $\mathbf{R}_n$, respectively.

$$\mathbf{x}_n = \mathbf{F}_n \mathbf{x}_{n-1} + \mathbf{w}_{n-1} \tag{10.a}$$

$$\mathbf{y}_n = \mathbf{H}_n \mathbf{x}_n + \mathbf{v}_n \tag{10.b}$$

The state vector for this problem, given by Eq. (11), is composed by the mean temperature and heat flux values, represented by the vectors $\overline{\mathbf{T}}_n$ and $\mathbf{q}_n$, at each control volume of the discretization grid. Thus, considering a grid with I volumes in the x direction and J volumes in the y direction, the number of unknowns is $2IJ$.

$$\mathbf{x}_n = \begin{bmatrix} \overline{\mathbf{T}}_n \\ \mathbf{q}_n \end{bmatrix} \tag{11}$$

The $\mathbf{F}_n$ matrix for the evolution model, with size $2IJ \times 2IJ$, is built with four matrices of size $IJ \times IJ$, as:

$$\mathbf{F}_n = \begin{bmatrix} \mathbf{A}_n & \mathbf{B}_n \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \tag{12}$$

The $\mathbf{A}_n$ and $\mathbf{B}_n$ matrices result from the discretization of the forward surrogate problem with the explicit scheme of the finite volume method [Patankar, 1980; Versteeg and Malalasekera, 1995; Ferziger and Peric, 2002]. $\mathbf{A}_n$ accounts for the heat diffusion in the domain, and $\mathbf{B}_n$ is given by Eq. (13).

$$\mathbf{B}_n = \frac{\Delta t}{cC^*} \mathbf{I} \tag{13}$$

The $\mathbf{0}$ and $\mathbf{I}$ terms are the zero and identity matrices, where a *random walk model* was used for the evolution of the unknown local heat fluxes. Considering the state noise as uncorrelated and with a standard deviation σ_q, this model can be described by Eq. (14), where $\boldsymbol{\omega}$ is a standard Gaussian vector.

$$\mathbf{q}_n = \mathbf{q}_{n-1} + \sigma_q \boldsymbol{\omega} \tag{14}$$

For the observation model, the matrix $\mathbf{H}_n$ is described by Eq. (15), where the diagonal matrix results from the Improved Lumped formulation:

$$\mathbf{H}_n = \left[\mathbf{0} \quad -\frac{c}{6k_T^*} \mathbf{I} \right] \tag{15}$$

The Kalman filter is applied to the solution of the present state estimation problem, (see Eqs. (16.a)-(16.e)). If the hypotheses of linear problem and Gaussian noise are respected, this set of equations produces an unbiased and minimal variance recursive estimator [Chen, 2003; Grewal and Andrews, 2008; Orlande et al., 2012]. Also, the covariance matrix of the estimates error $\mathbf{P}_n$ allows the construction of confidence interval for better analysis of the obtained results.

$$\mathbf{x}_{n|n-1} = \mathbf{F}_n \mathbf{x}_{n-1} \tag{16.a}$$

$$\mathbf{P}_{n|n-1} = \mathbf{F}_n \mathbf{P}_{n-1} \mathbf{F}_n^T + \mathbf{Q}_n \tag{16.b}$$

$$\mathbf{K}_n = \mathbf{P}_{n|n-1} \mathbf{H}_n^T (\mathbf{H}_n \mathbf{P}_{n|n-1} \mathbf{H}_n^T + \mathbf{R}_n)^{-1} \tag{16.c}$$

$$\mathbf{x}_n = \mathbf{x}_{n|n-1} + \mathbf{K}_n (\mathbf{y}_n - \mathbf{H}_n \mathbf{x}_{n|n-1}) \tag{16.d}$$

$$\mathbf{P}_n = (\mathbf{I} - \mathbf{K}_n \mathbf{H}_n) \mathbf{P}_{n|n-1} \tag{16.e}$$

Results

In this work, the experimental measurements were simulated with the complete model, using a fine grid with $768 \times 768 \times 64$ volumes and time step $\Delta t = 10^{-4} s$ to ensure numerical convergence. The inverse problem was solved with 24×24 volumes and a time step $\Delta t = 0.01s$. This was done so that the simulated measurements are free of inverse crime [Kaipio and Somersalo, 2004]. The initial temperature was considered as $300K$. The observation noise was assumed as Gaussian, uncorrelated, with zero mean and constant standard deviation, σ_y. In a real situation, these measurements could be obtained with modern infrared cameras, which presents standard deviations of the order of $0.01°C$ [Orlande et al., 2013]. For testing the performance of the Kalman filter, a relatively high value ($\sigma_y = 1°C$) was selected. The proposed heat flux is described by Eq. (17) and Tab. 2. The size of the region of application of the heat flux is selected so that it does not necessarily coincide with the control volume size.

$$q(x, y, t) = \begin{cases} q_0 & \text{if} \quad x_1 \leq x \leq x_2, \quad y_1 \leq y \leq y_2 \quad \text{and} \quad t \geq t_1 \\ 0 & \text{otherwise} \end{cases} \tag{17}$$

Table 2: Parameters of the proposed heat flux

Quantity	Value	Quantity	Value	Quantity	Value
x_1	60mm	y_1	60mm	q_0	$10^7 \, \text{W/m}^2$
x_2	72mm	y_2	72mm	t_1	0.4s

The comparison between the projection of the exact temperature field on the coarse grid and the estimated values at time $t = 2.0s$ is presented in Figs. 2.a-2.b. The agreement between these values

is excellent, once the region where the heating occurs is adequately identified and the estimated temperatures are very close to the exact values. No signs of correlation were detected in the obtained residuals. The largest residual was in the heated region and its vicinity, but its value was approximately 0.45°C. Thus, both the largest residual and the standard deviation of the experimental measurements have the same order of magnitude ($1°C$), and the temperature estimates can be considered as good [Ozisik and Orlande, 2000].
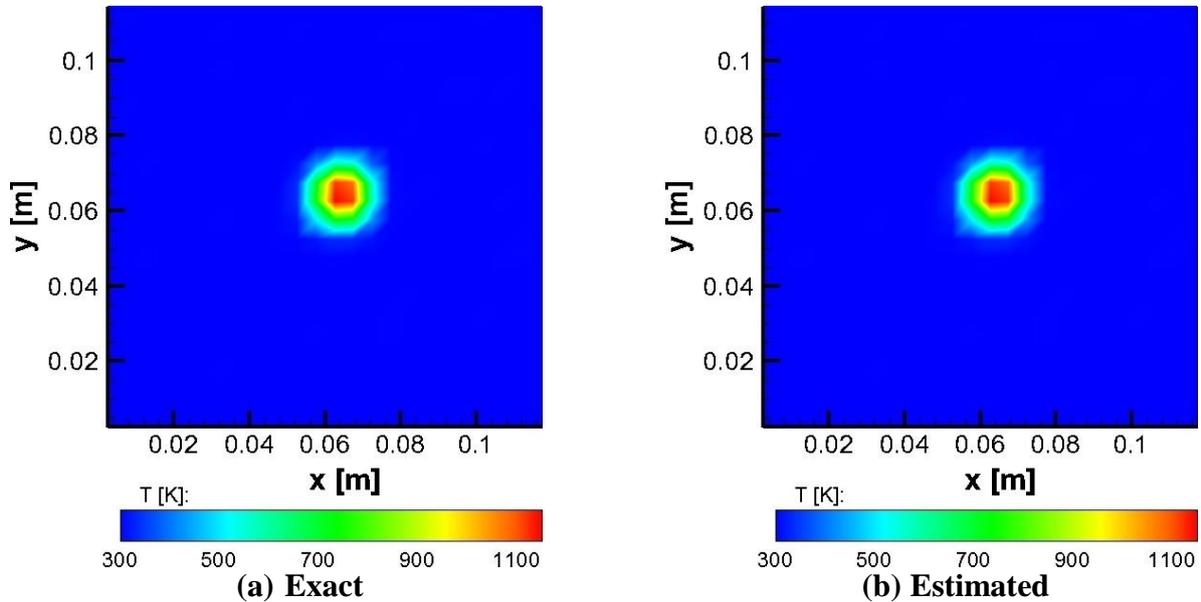
Figure 2: Comparison of the exact and estimated temperature field at time $t = 2.0s$.

The same comparison made for the temperature is presented for the heat flux in Figs. 4.a-4.d, where the projections of the exact heat flux values in the coarse grid is presented in Fig. 4.a, while the estimated values at times t = 1.0s, 1.5s and 2.0s are presented in Figs. 4.b, 4.c and 4.d, respectively.

(c) t=1.5s

(d) t=2.0s
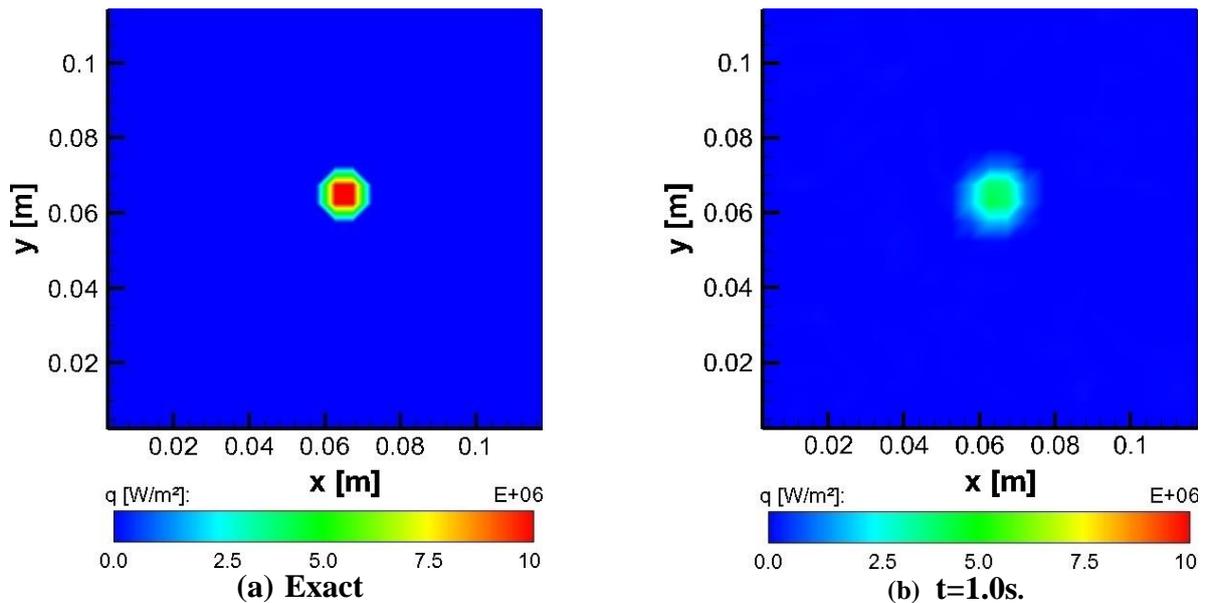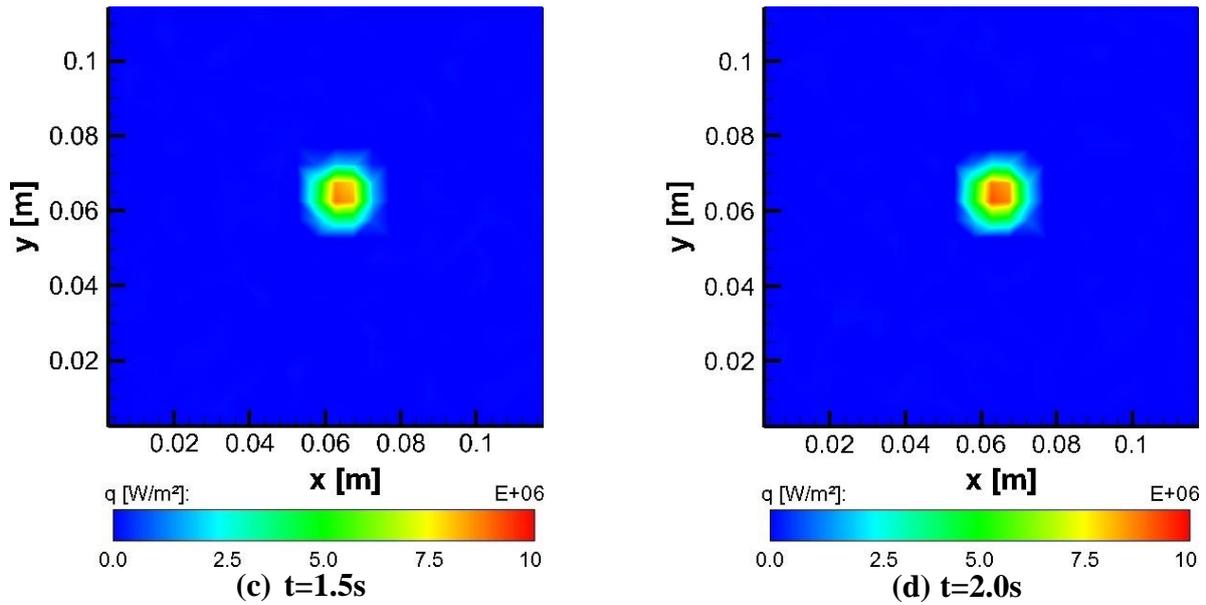
Figure 3: Comparison between the exact heat flux and the estimates at different times.

The results show that the region where the heating occurs is very well identified. However, some quantitative differences can be observed between the reference values and the estimates. For better understanding of these differences, the evolutions in time of the exact and estimated values are presented in Figs. 5.a-5.b for the point $(x, y) = (62.5; 62.5)mm$, located inside the heated region.

(a) Temperature

(b) Heat Flux
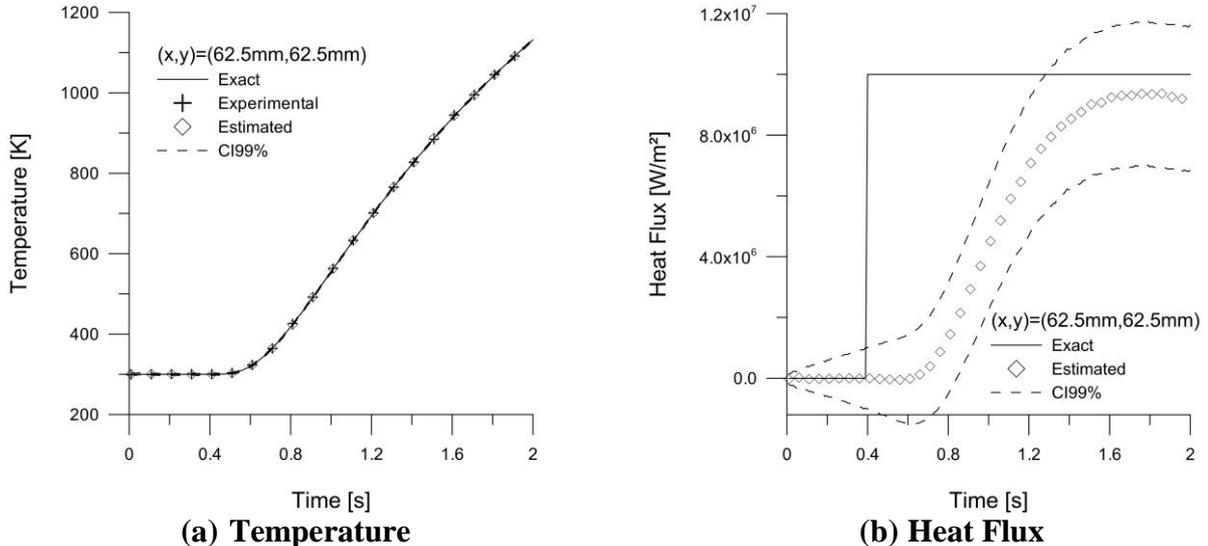
Figure 4: Evolution of the reference and estimated values with time at $(x, y) = (62.5; 62.5)mm$.

For these results, the 99% confidence intervals show again the good quality of the temperature estimates. However, since the heating occurs on the opposite side from where the measurements are obtained, a time of approximately $0.2s$ is elapsed before the filter shows any change in the heat flux estimates resulting from the applied heat flux. This is due to the time required for the diffusion of heat through the thickness of the plate. Also, in the vicinity of $t = 2.0s$, the estimates show a decreasing behavior as a result of the modeling errors of the surrogate model at high temperatures.

On the other hand, it is possible to extract information about the order of magnitude of the heat flux and its region of application.

Conclusions

The proposed inverse problem, for which nonlinear and three dimensional models would be needed, could be reasonably well solved with simplified models, allowing for the use of fast and computationally efficient algorithms, such as the Kalman Filter. The temperature estimates present very good agreement with reference values. For the estimation of the heat flux, although the effect of the modeling errors of the surrogate model is noticeable, the heating location is adequately identified and the obtained estimates have the same order of magnitude as the exact values. Improvement of these results relies in accounting for modeling errors in the solution of the inverse problem.

Acknowledgements

References

Afrin, N., Feng, Z. C., Zhang, Y. and Chen, J. K. (2013). Inverse estimation of front surface temperature of a locally heated plate with temperature-dependent conductivity via Kirchhoff transformation. *International Journal of Thermal Sciences,* **69,** 53-60.

Chen, Z. (2003). Bayesian filtering: From Kalman filters to particle filters, and beyond. *Statistics,* **182,** 1-69.

Cotta, R. M. and Mikhailov, M. D. (1997). *Heat Conduction: Lumped Analysis, Integral Transforms, Symbolic Computation*. Wiley Ed.

Dennis, B. H. and Dulikravich, G. S. (2001). Simultaneous determination of steady temperatures and heat fluxes on surfaces of three dimensional objects using FEM. *IMECE*, New York, NY.

Dennis, B. H. and Dulikravich, G. S. (2012). Inverse determination of unsteady temperatures and heat fluxes on inaccessible boundaries. *Inverse and Ill-Posed Problems,* **20,** 791-803.

Feng, Z. C., Chen, J. K., Zhang, Y. and Griggs Jr., J. L. (2011). Estimation of front surface temperature and heat flux of a locally heated plate from distributed sensor data on the back surface. *International Journal of Heat and Mass Transfer,* **54,** 3431-3439.

Ferziger, J. H. and Peric, M. (2002). *Computational Methods in Fluid Flow*. Springer-Verlag.

Grewal, M. S. and Andrews, A. P. (2008). *Kalman Filtering: Theory and Practice Using MATLAB*. Wiley.

Kaipio, J. P. and Somersalo, E. (2004). *Statistical and Computational Inverse Problems*. Springer Science+Business Media, Inc.

Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Transactions of the ASME-Journal of Basic Engineering,* **82,** 35-45.

Orlande, H. R. B., Colaço, M. J., Dulikravich, G. S., Vianna, F. L. V., da Silva, W. B., Fonseca, H. M. and Fudym, O. (2012). State Estimation Problems in Heat Transfer. *International Journal for Uncertainty Quantification,* **2,** 239-258.

Orlande, H. R. B., Dulikravich, G. S., Neumayer, M., Watzenig, D. and Colaço, M. J. (2013). Accelerated bayesian inference for the estimation of spatially varying heat flux in a heat conduction problem. *Numerical Heat Transfer, Part A,* **65**.

Ozisik, M. N. (1993). *Heat Conduction*. John Wiley & Sons, Inc.

Ozisik, M. N. and Orlande, H. R. B. (2000). *Inverse Heat Transfer: Fundamentals and Applications*. Taylor & Francis.

Patankar, S. V. (1980). *Numerical Heat Transfer and Fluid Flow*. Hemisphere Publishing Company.

Versteeg, H. K. and Malalasekera, W. (1995). *An Introduction to Computational Fluid Dynamics: The Finite Volume Method*. Longman Group Ltd.

Simulation of liquid-solid interaction using a phase-field-lattice Boltzmann method

*R. Rojas[1], and T. Takaki[2]

[1]Mechanical and System Engineering, Graduate School of Science and Technology, Kyoto Institute of Technology, Matsugasaki, Sakyo, Kyoto 606-8585, Japan
[2]Mechanical and System Engineering, Graduate School of Science and Technology, Kyoto Institute of Technology, Matsugasaki, Sakyo, Kyoto 606-8585, Japan

*Corresponding author: rojas@kit.ac.jp

Abstract

In this study, a combination of the lattice Boltzmann method (LBM) and the phase-field method (PFM) is used for the calculation of liquid-solid two-phase flows with solidification. PFM is used as a numerical tool to capture interface topology changes of solid and the flow of liquid is solved by using LBM. The no-slip boundary condition at the liquid-solid interface is satisfied by adding a diffusive-forcing term in LBM formulation. Calculations of Poiseuille flows and flows past a circular cylinder at different Reynolds numbers confirm that the no-slip boundary condition is effectively satisfied at the diffuse liquid-solid interface. Then, the present method is applied to the calculation of two-dimensional anisotropic dendritic growth of a binary alloy under melt convection. Two cases have been studied. Initially, the solid is stationary, and then the solid is free to move under the influence of the flow. The equations of motion are solved to track the translation and rotation of the solid. Qualitative comparisons of the solidification patterns reveal that the microstructure dendritic growth is mainly affected in the direction of the fluid flow. The results obtained with the present method agree well with those obtained with other available numerical techniques.

Keywords: Liquid-solid two-phase flow, Phase-field method, Lattice Boltzmann method, Solidification

Introduction

Dendrites are common structures in the solidification of metals and alloys. They reproduce due to unstable movement of a liquid-solid interface. Dendritic solidification is governed by complex processes such as surface tension, heat transport (for pure materials) and/or concentration gradient (for alloys). Due to the importance of this phenomenon, some numerical techniques using the phase-field method (PFM) have been proposed to effectively capture the interface morphology [Karma&Rappel (1998); Ohno&Matsuura (2009)]. Other important aspect in solidification is convection. Convection plays a crucial role on crystal growth, and its effects have been studied by taking into account the effect of melt convection [Beckermann et al. (1999), Lu et al. (2003), Miller et al. (2001), Medvened et al (2006)]. In order to simulate solidification under melt convection, the aforementioned methods combine PFM to solve the phase transformation process and an appropriate method to solve the fluid flow, i.e. the solution of the Navier-Stokes equations or the lattice Boltzmann method (LBM) [Chen&Doolen (1998)]. In this studies, the solid is assumed to be stationary, so that the momentum equation for the solid is not solved.

In this study a combination of PFM proposed by Ohno&Matsuura [Ohno&Matsuura (2009)] and LBM has been used for simulation of solidification. The main advantage of this PF-LBM over others is the possibility of choosing different diffusivities in the solid and liquid parts, and therefore allowing solidification involving diffusion in solid of a binary alloy. Most importantly, the solid is allowed to move freely, and the equations of motion are solved to calculate the translational and rotational velocities. Additionally, LBM has been adopted for modeling the fluid flow due to its computational advantages related to easy programming and suitability for parallel computing.

First, the present method is applied to benchmark simulations with stationary diffuse interfaces. Poiseuille flows and flows past a circular cylinder are carried out to validate the no-slip boundary condition at the liquid-solid interface. Then, calculations of isothermal dendritic solidification under melt convection of an Al-Cu alloy are carried out. Two cases are considered to study the effect of melt convection. Initially, the solid part is considered to be stationary, and then the solid part is allowed to move with the fluid flow. The analysis of the results is based on qualitative comparisons of the differences in the solidification patterns.

Numerical Method

The present method uses the quantitative phase-field method for dilute alloy solidification [Ohno&Matsuura (2009)] and the lattice Boltzmann method with discrete forcing term [Chen&Doolen (1998)]. The interaction between solid and liquid parts is modeled by a diffuse force proposed by Beckermann et al. [Beckermann et al. (1999)]. A detailed explanation of the numerical methods is given in the following.

Phase-field method

The time evolution of the phase field, ϕ, is given by [Ohno&Matsuura (2009)]

$$a_s(\theta)^2 \partial_t \phi = \nabla \cdot \left(a_s(\theta)^2 \nabla \phi\right) - \partial_x \left[a_s(\theta)a_s{}'(\theta)\partial_y \phi\right] + \partial_y \left[a_s(\theta)a_s{}'(\theta)\partial_x \phi\right] - f'(\phi) - \lambda^* g'(\phi)u \quad (1)$$

where $a_s(\theta)$ is a function to represent crystalline anisotropy, $\phi=-1$ for liquid and $\phi=+1$ for solid, $a_s{}'(\theta) = da_s/d\theta$, $f(\phi)$ and $g(\phi)$ are interpolating functions associated with the double-well and chemical potential, respectively, λ^* is a dimensionless parameter that controls the coupling between the phase field and the concentration field represented by the dimensionless supersaturation, u.

The time evolution of the supersaturation, u, is given by

$$\left[1 + k - (1-k)h(\phi)\right](\partial_t u + \boldsymbol{U} \cdot \nabla u)/2 = \nabla\left[D_l q(\phi)\nabla u - j_{AT}\right] + \left[1 + (1-k)u\right]\partial_t h(\phi)/2 - \nabla \cdot \boldsymbol{J} \quad (2)$$

where k is the partition coefficient, $h(\phi)$ and $q(\phi)$ are interpolating functions, $\boldsymbol{U}$ is the fluid velocity, D_l is the liquid diffusivity, j_{AT} is the antitrapping current term, and $\boldsymbol{J}$ is a function to include noise fluctuations [Echebarria (2010)].

Lattice Boltzmann method

The lattice Boltzmann equation with single relaxation time, τ, and discrete forcing term, G_i, is given by

$$f_i(\boldsymbol{x} + \boldsymbol{c}_i \delta t, t + \delta t) = f_i(\boldsymbol{x},t) - \frac{1}{\tau}[f_i(\boldsymbol{x},t) - f_i^{eq}(\boldsymbol{x},t)] + G_i(\boldsymbol{x},t)\delta t \quad (3)$$

where f_i is the particle velocity distribution in the ith direction, $\boldsymbol{x}$ is the position vector, $\boldsymbol{c}_i$ the discrete particle velocity, t the time, δt the time step size, f_i^{eq} is the equilibrium distribution function. The fluid density, ρ, and velocity $\boldsymbol{U}$ are given by

$$\rho = \sum_{i=0}^{Q-1} f_i \quad (4)$$

$$\rho \boldsymbol{U} = \sum_{i=0}^{Q-1} \boldsymbol{c}_i f_i \quad (5)$$

where Q is the number of discrete velocities. In this study, the two-dimensional nine-velocity (D2Q9) model is used in the calculations of the discrete velocity. The discrete forcing term, G_i, with second-order accuracy is given by

$$G_i = \rho W_i \left[3 \frac{c_i - u}{c^2} + 9 \frac{(c_i \cdot u)c_i}{c^4} \right] \cdot G \quad (6)$$

where W_i is the weighting function related to f_i^{eq} and c the lattice velocity ($c=\delta x/\delta t=1$). G is an external force.

In the present study, effects of the gravity and other external forces are neglected. Therefore, the only force acting on the fluid flow is the one related to drag. A dissipative drag force is used to impose the no-slip boundary condition at the diffuse liquid-solid interface. It is given by [Beckermann et al. (1999)]

$$G(x,t) = 2\rho \nu g \left(\frac{1+\phi}{2} \right)^2 (U_S - U) \quad (7)$$

where ν is the kinematic viscosity, g is a dimensionless constant, i.e. $g = 2.757$ and U_S is the solid velocity. The motion of the solid part is calculated by solving the following equations [Glowinski et al. (2001)]:

$$M_S \frac{dU_T}{dt} = G_S \quad (8)$$

$$I_S \frac{d\omega_S}{dt} = T_S \quad (9)$$

where M_S is the mass of the solid, U_T is the translational velocity of the solid, I_S is the tensor for the moment of inertia, ω_S is the angular velocity, G_S and T_S are the total force and torque acting on the solid, respectively. They are given by

$$G_S = -\sum_x G(x,t)\Delta V \quad (10)$$

$$T_S = -\sum_x (x - X_S) \times G(x,t)\Delta V \quad (11)$$

where ΔV is the volume of the computational cell and X_S is the center of mass of the solid. The velocity U_S is given by $U_S = U_T + \omega_S \times (x-X_S)$. The location of the phase field is updated with an advection equation. The advection term is discretized with the WENO fifth-order scheme.

The basic solution of the present method is illustrated as follows:
1. Solve Eqs. (1) - (2) with the explicit Euler method.
2. Solve the lattice Boltzmann equation, Eq. (3), with discrete forcing term, Eqs. (6) and (7), and obtain U.
3. If necessary, calculate the motion of the solid with Eqs. (8), (9), (10) and (11) and update the location of the phase field with an advection equation.

Validation

Initially, calculations of Poiseuille flow and flows past a circular cylinder at different Reynolds numbers are carried out to validate the no-slip boundary conditions. The phase field does not
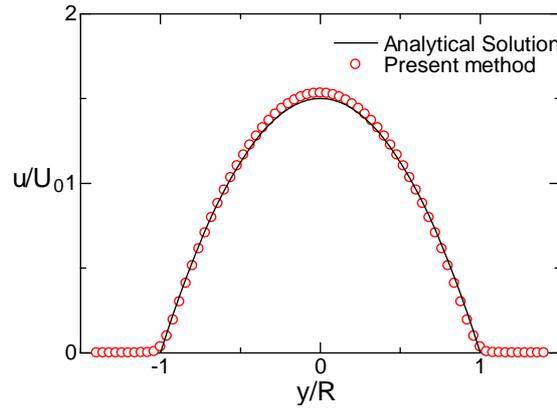
Figure 1. Velocity profile of Poiseuille flow

Table 1 Comparison of drag coefficient for steady flows past a circular cylinder

Re	10	20	40
IB-FDLBM [Rojas et al. (2011)]	2.95	2.135	1.58
Present method	3.16	2.31	1.75

Table 2 Properties of Al-Cu alloy

Diffusivity in liquid	D_l [m^2/s]	3.0×10^{-9}
Diffusivity in solid	D_s [m^2/s]	3.0×10^{-13}
Partition coefficient	k	0.14

change in these calculations and it is only used to distingue the fluid and the solid, e.g. walls or boundaries. For Poiseuille flows, the computational domain is long enough so that the predicted velocity profile corresponds to the fully developed flow. The phase field is set at $\phi = -1$ at the flow region and $\phi = 1$ at the top and bottom walls. Figure 1 shows a comparison of the theoretical velocity profile and the one predicted with the present method. As can be seen, both velocity profiles agree well. Moreover, the no-slip boundary condition at the top and bottom walls is effectively satisfied.

Flows past a circular cylinder are calculated using a square computational domain. The dimensions are 1600 and 1600 lattice points in the x and y directions, respectively. The left boundary condition is inlet flow at U_0, and the right, top and bottom walls are outflow boundary conditions. A circular cylinder is located at the center of the domain and the phase field is $\phi = 1$, in the fluid flow $\phi = -1$. Table 1 shows a comparison of the drag coefficient with other numerical results in literature [Rojas et al. (2011)]. The results obtained with the present method slightly differ from the results obtained by using an immersed boundary method. This is because in the present calculation the interface has a finite width and better agreements are expected as the interface width approaches cero.

Simulation of dendrites

In this section, numerical simulations are carried out to analyze the effect of melt convection on isothermal dendritic growth of an undercooled Al-Cu alloy. Some physical properties of an Al-Cu alloy are given in Table 2. Two cases are analyzed. In Case 1, the solid part of the alloy is stationary, i.e. $U_S = 0$. In case 2, the solid part is allowed to move within the fluid flow, i.e. free motion. The computational domain is shown in Fig. 2. The computational domain sizes, $W \times H$, are 1001×1001 and 2001×1001 lattice points for Cases 1 and 2, respectively. The computational domain of Case 2 is larger in the x-axis because the solid part is moving along this direction. A seed is located at (501,501) and its diameter occupies six lattice spacings. A uniform inflow at U_0 comes

from the left boundary. The right boundary is continuous outflow condition and the top and bottom walls are slip boundaries. In this study, $U_0 = 0.01$, $\tau = 1$, and $u = 0.4$.

The phase field distributions at two different instants, t_1 and t_2 are shown in Figs. 3 and 4. Figure 3 illustrates the case when the solid part is stationary. In the absence of melt convection, dendritic growth is symmetric in all directions [Ohno&Matsuura (2009)]. On the contrary, the effect of melt convection breaks the symmetrical growth morphology. As can be seen, the upstream facing parts of the alloy grow faster than those downstream orientated. These patterns agree well with other qualitative results obtained by using other numerical methods [Beckermann et al. (1999), Lu et al. (2003), Miller et al. (2001), Medvened et al (2004)].

Figure 3 shows the phase field distribution of Case 2 at t_1 and t_2. The solid part has been displaced from its initial position with U_S. Despite dendritic growth seems to be symmetrical, a detailed examination revels that the rear tip grows slightly faster than the front and normal tips. This behavior is because the fluid flow enhances dendritic growth along its direction. As can be seen, the motion of the solid part is mainly related to the translational velocity of the solid. The effect of the rotational velocity is very small in this case. Therefore, calculations of dendritic growth in a shear flow are carried out to evaluate the rotation of the solid. The computational domain size is $W \times H = 1001 \times 1001$. The dimensionless velocities at top and bottom walls are -0.02 and +0.02, respectively. Periodic boundary conditions are set along the x-axis. Figures 5 (a) and (b) show the distribution of ϕ at t_1 and t_2. As can be seen, the simultaneous growth and rotation of the solid is effectively predicted. Dendritic growth is favorable along the preferable directions, and the fluid flow slightly modifies the growth in the direction perpendicular to the preferable directions.
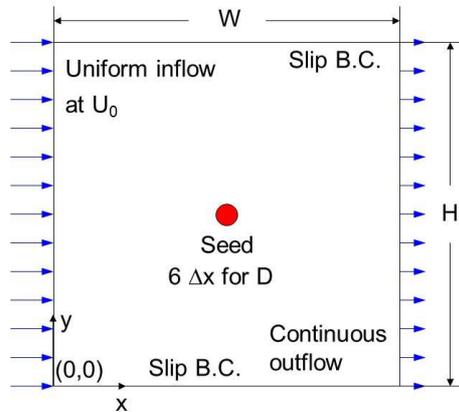
Figure 2. Computational domain for simulation of dendrites

(a) t_1 (b) t_2
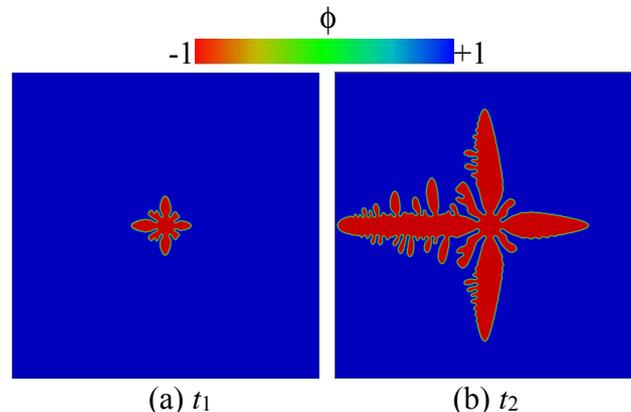
Figure 3. Phase field distribution, ϕ, at $U_S = 0$

(a) t_1 (b) t_2
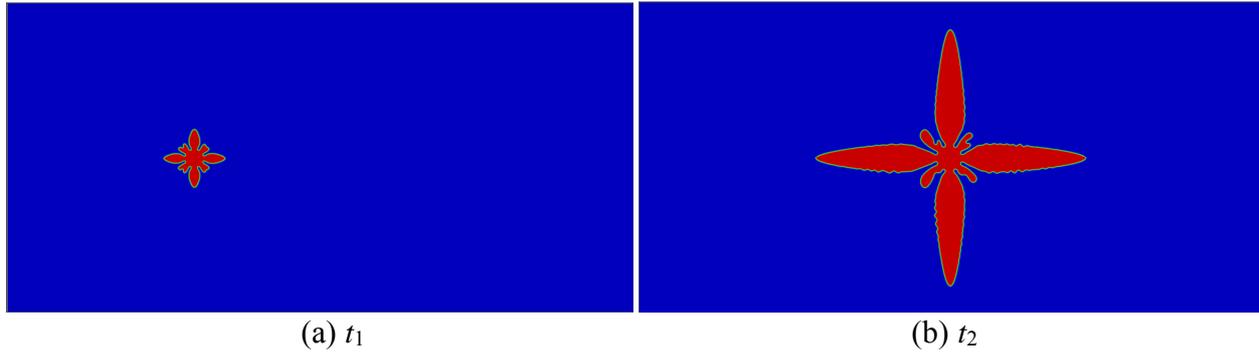
Figure 4. Phase field distribution, φ, at free motion

(a) t_1 (b) t_2
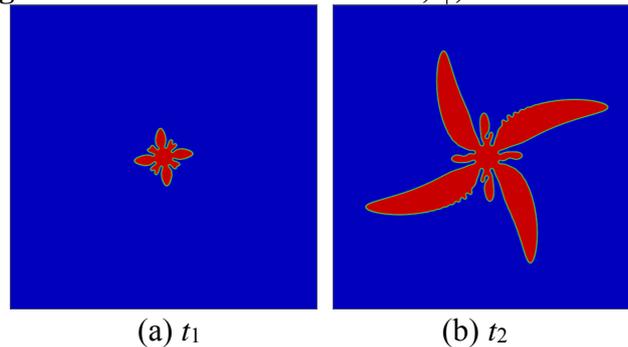
Figure 5. Phase field distribution, φ, of dendritic growth in a shear flow

Conclusions

A combination of the phase-field method and the lattice Boltzmann method was used for simulations of dendrites. Two cases were studied to analyze the effect of melt convection on dendritic growth. First, the solid part of the alloy was stationary and then it was allowed to move freely within the fluid flow. As a result, the following conclusions were obtained: (1) melt convection strongly affects the morphology of dendritic growth; it accelerates growth along its direction and (2) the motion of the solids affects dendritic growth in all directions, specially the direction of the fluid flow

References

Beckermann, C., Diepers, H.-J., Steinbach, I., Karma, A., and Tong, X. (1999) Modeling melt convection in phase-field simulations of solidification, *Journal of Computational Physics* **154**, 468-496.

Chen, S. and Doolen G.D. (1998) Lattice Boltzmann method for fluid flows, *Annual Review of Fluid Mechanics* **30**, 329-364.

Echebarria, B. (2010) Onset of sidebranching in directional solidification, *Physical Review E* **81**, 021608.

Glowinski, R., Pan, T.W., Hesla, T.I., Joseph, D.D., and Periaux, J. (2001) A fictitious domain approach to the direct simulation of incompressible viscous flow past moving rigid bodies: application to particulate flow, *Journal of Computational Physics* **169**, 363-426.

Karma, A. and Rappel W.J. (1998) Quantitative phase-field modeling of dendritic growth in two and three dimensions, *Physical Review E* **57 (4)**, 4323-4349.

Lu, Y., Beckermann, C., and Ramirez, J.C. (2005) Three-dimensional phase-field simulations of the effect of convection on free dendritic growth, *Journal of Crystal Growth* **280**, 320-334.

Miller, W., Succi, S., and Mansutti, D. (2001) Lattice Boltzmann method for anisotropic liquid-solid phase transition, *Physical Review Letters* **86 (16)**, 3578-3581.

Medvened, D., Fischaleck, T., and Kassner, K. (2006) Influence of external flows on crystal growth: Numerical investigation, *Physical Review E* **74**, 031606.

Ohno, M. and Matsumura K. (2009) Quantitative phase-field modeling for dilute alloy solidification involving diffusion in the solid, *Physical Review E* **79**, 031603.

Rojas, R., Seta, T., Hayashi, K., and Tomiyama, A. (2011) Immersed boundary-finite difference lattice Boltzmann method for liquid-solid two-phase flows, *Journal of Fluid Science and Technology* **6(6)**, 1051-1064.

ICCM2014, 28th-30th July 2014, Cambridge, England

Implicit computational method for compressible flows with high and low Mach numbers

K. Aoki[1], S. Ushijima[2], D. Toriu[1], and H.Itada[1]

[1] Kyoto University, Civil and Earth Resources Engineering, Kyoto, 615-8540, Japan
[2] Prof., Kyoto University, ACCMS, Kyoto, 606-8501, Japan

*Corresponding author : ushijima@media.kyoto-u.ac.jp

Abstract

This paper presents a new computational method for compressible flows with both high and low Mach numbers. In this method, the principal governing equations, such as the equation of mass, momentum and internal energy, are described with conservation forms, which are discretized with finite volume method (FVM) on a collocated grid system. In particular, since the special implicit method (C-ISMAC method) is applied to these governing equations, their convection terms can be estimated with higher-order TVD schemes in addition to the larger time increment Δt compared with the usual explicit methods that decreases computational time of our method.

The proposed computational method was applied to the shock wave problems with high Mach numbers as well as the natural convection flows driven by non-uniform temperatures with low Mach numbers. It was shown that the conservation of mass and other physical properties are sufficiently satisfied and that the predicted results are in good agreement with the theoretical values and calculated results reported by other researchers.

Keywords : compressible fluid, FVM, implicit method, Mach number, natural convection

Introduction

In this paper, a numerical method is investigated to predict compressible flows for high Mach numbers in addition to low Mach numbers. In order to construct such prediction methods, it is necessary to capture accurately shock discontinuity due to shock waves with high Mach numbers as well as to predict the compressible flows with low Mach numbers affected by large pressure or temperature differences.

To propose such computational method, the basic ideas used in the computations for incompressible fluids are employed: the numerical procedures in collocated grid system, special implicit method applicable to higher-order schemes in convection terms. Thus, the main numerical procedures are described as follows:

1). The governing equations are described in the conservative forms and the convection and diffusion terms are discretized with finite volume method (FVM).

2). As a result, the mass conservation law is satisfied accurately in the present computational method compared with the usual finite difference methods (FDM) based on non-conservative forms.

3). The employed implicit method (C-ISMAC method [Ushijima and Nezu (2002)]) enables us to utilize the higher-order TVD schemes and to use larger time increment Δt than explicit methods. Thus, accurate numerical results can be obtained with short elapse time.

Applying the present computational method to multiple problems, it will be confirmed that the present computational method allows us to predict the one- to three-dimensional shock-wave problems with high Mach numbers as well as the natural convection flows with low Mach numbers.

Numerical procedures

Governing Equations

The governing equations are the conservation for mass, momentum and internal energy. In Cartesian coordinate system they are given by

$$\frac{\partial \rho}{\partial t} + \frac{\partial (\rho u_j)}{\partial x_j} = 0 \tag{1}$$

$$\frac{\partial (\rho u_i)}{\partial t} + \frac{\partial (\rho u_i u_j)}{\partial x_j} = -\frac{\partial p}{\partial x_j} + \frac{\partial \tau_{ij}}{\partial x_j} - \rho g \delta_{i3} \tag{2}$$

$$\frac{\partial (\rho e)}{\partial t} + \frac{\partial (\rho e u_j)}{\partial x_j} = -p\frac{\partial u_i}{\partial x_i} + \tau_{ij}\frac{\partial u_i}{\partial x_j} - \frac{\partial q_j}{\partial x_j} \tag{3}$$

where ρ, u_i, p, τ_{ij}, g, δ_{ij}, q_i are the density, the velocity component in the x_i direction, pressure, viscous stress tensor, gravity, Kronecker's delta and the heat flux in the x_j direction, respectively. The x_3 coordinate towards the vertically upward direction.

In the above equations, the relationship between the internal energy e and the temperature T is given by

$$e = C_v T \tag{4}$$

where C_v is the specific heat at constant volume. The equation of state for ideal gasses is given by

$$p = \rho e(\gamma - 1) \tag{5}$$

where the γ is the ratio of specific heat. The viscous stress τ_{ij} and the heat flux q_j are defined as

$$\tau_{ij} = \mu\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right) - \lambda\frac{\partial u_k}{\partial x_k}\delta_{ij} \tag{6}$$

$$q_j = -\kappa\frac{\partial T}{\partial x_j} \tag{7}$$

where $\lambda = -(2/3)\mu$ and κ is the coefficient of thermal conductivity.

Computational Method for Governing Equations

As shown in Fig.1, the governing equations are discretized on the collocated grid points in the cell where Q denotes the scalar variables. The numerical procedure is given as follows:

1). The tentative velocity components u_i^* are calculated at the cell-center points without pressure-gradient terms.

2). The pressure-gradient terms are added to the velocity components after the interpolation of $u_{b,i}$ on cell boundaries. The cell-boundary velocity components, including the pressure gradient terms, are denoted by $u_{b,i}^*$.
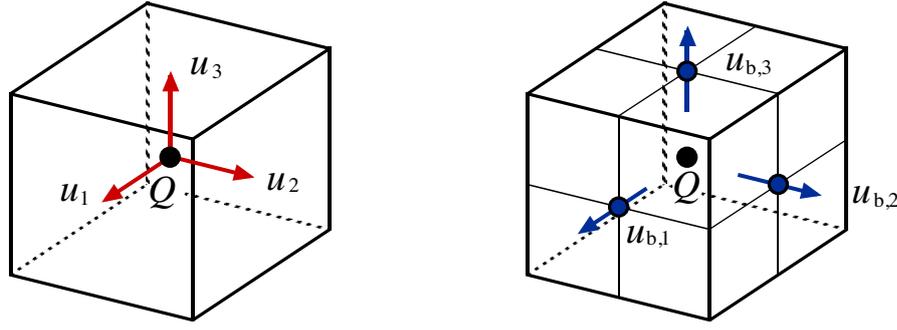
Figure 1. Collocated grid system

3). The density ρ^{n+1} at $n+1$ time step is calculated from Eq.(1) with the implicit method as detailed later.

4). $(\rho T)^{n+1}$ is calculated from Eqs.(3) and (4) with the implicit method as well. The temperature T^{n+1} is determined with the ρ^{n+1}.

5). The pressure p^{n+1} is obtained by Eqs.(4) and (5) with ρ^{n+1} and T^{n+1}. Thus, the equation of state is satisfied by all variables at $n+1$ time step.

6). The momentum $(\rho u_i)^{n+1}$ is calculated from Eq.(2) with the implicit method using the variables at $n+1$ time step. The velocity u_i^{n+1} is determined with ρ^{n+1}.

The details of the above numerical procedures are as follows: the tentative velocity component u_i^* is calculated at the center of the cells with FVM:

$$u_i^* = \frac{1}{\rho^n}\left[(\rho u_i)^n + \Delta t \left\{ -\left(\frac{\partial(\rho u_i u_j)}{\partial x_j}\right)^n + \left(\frac{\partial \tau_{ij}}{\partial x_j}\right)^n \right\} \right] \tag{8}$$

where superscripts denote the time step. In order to estimate pressure-gradient terms accurately, these terms are not included in the estimation of u_i^* at the cell-center points.

After this procedure, the pressure-gradient terms estimated on the cell boundaries are added to the interpolated velocity. Thus, we obtain the cell-boundary velocity component $u_{b,i}^*$ as follows:

$$u_{b,i}^* = f_b(u_i^*) - \frac{1}{\rho^n}\frac{\partial p^n}{\partial x_i}\bigg|_b \Delta t \tag{9}$$

where $f_b()$ denotes the function to interpolate the variable on a cell boundary, which is a simple linear average between the cell-center variables in the present paper.

Equation (1) is discretized with the implicit method called C-ISMAC method [Ushijima and Nezu (2002)] proposed for the collocated grid system, which is based on the implicit SMAC method [Shin et al. (1993)] in the staggered grid system. The C-ISMAC method allows us to decrease computational time without decreasing numerical accuracy.

The equation discretized with respect to time by the C-ISMAC method is given by

$$\frac{\rho^{n+1} - \rho^n}{\Delta t} + \alpha_\rho \frac{\partial(\rho^{n+1} u^*)}{\partial x_j} + (1 - \alpha_\rho)\frac{\partial(\rho^n u_j^*)}{\partial x_j} = 0 \tag{10}$$

where α_ρ is a parameter whose range is $0 \le \alpha_\rho \le 1$. With the following definition,

$$\rho^{n+1} = \rho^n + \tilde{\rho} \tag{11}$$

Eq.(10) is transformed to the following equation:

$$\frac{\tilde{\rho}}{\Delta t} + \alpha_\rho \frac{\partial(\tilde{\rho} u_j^*)}{\partial x_j} = -\frac{\partial(\rho^n u_j^*)}{\partial x_j} \tag{12}$$

where $\tilde{\rho}$ becomes nearly zero when the flow field is almost steady or the time-scale of the flow field is sufficiently larger than the time increment Δt. Thus, we can apply a simple first-order spatial discretization method to the left-hand side of Eq.(10), while higher-order scheme to the right-hand side. After solving the simultaneous linear equations of $\tilde{\rho}$, which is derived from the discretized equation of Eq.(12) with respect to space, ρ^{n+1} can be obtained from Eq.(11).

With Eq.(4) in which C_v is assumed to be constant, Eq.(3) is rewritten as

$$\frac{\partial(\rho T)}{\partial t} + \frac{\partial\{(\rho T) u_j\}}{\partial u_j} = -(\rho T)(\gamma - 1)\frac{\partial u_i}{\partial x_i} + \frac{1}{C_v}\left\{ \tau_{ij}\frac{\partial u_i}{\partial x_j} + \frac{\partial}{\partial x_j}\left(\kappa \frac{\partial T}{\partial x_j} \right) \right\} \tag{13}$$

Similarly, the equation discretized by the C-ISMAC method is given by

$$\frac{(\rho T)^{n+1} - (\rho T)^n}{\Delta t} + \alpha_{\rho T}\frac{\partial\{(\rho T)^{n+1} u_j^*\}}{\partial x_j} + (1 - \alpha_{\rho T})\frac{\partial\{(\rho T)^n u_j^*\}}{\partial x_j}$$
$$= (\gamma - 1)\left[\beta_{\rho T}\left\{ -(\rho T)^{n+1}\frac{\partial u_j^*}{\partial x_j} \right\} + (1 - \beta_{\rho T})\left\{ -(\rho T)^n\frac{\partial u_j^*}{\partial x_j} \right\} \right]$$
$$+ \frac{1}{C_v}\left\{ \tau_{ij}\frac{\partial u^*}{\partial x_j} + \frac{\partial}{\partial x_j}\left(\kappa \frac{\partial T^n}{\partial x_j} \right) \right\} \tag{14}$$

where $\alpha_{\rho T}, \beta_{\rho T}$ are parameters whose ranges are $0 \le \alpha_{\rho T}, \beta_{\rho T} \le 1$. With the following definition,

$$(\rho T)^{n+1} = (\rho T)^n + (\tilde{\rho T}) \tag{15}$$

Eq.(14) is transformed to the following equation:

$$\frac{(\tilde{\rho T})}{\Delta t} + \alpha_{\rho T}\frac{\partial\{(\tilde{\rho T}) u_j^*\}}{\partial x_j} + \beta_{\rho T}(\gamma - 1)\left\{ (\tilde{\rho T})\frac{\partial u_i^*}{\partial x_i} \right\}$$
$$= -\frac{\partial\{(\rho T)^n u_j^*\}}{\partial x_j} + (\gamma - 1)\left\{ -(\rho T)^n\frac{\partial u_j^*}{\partial x_j} \right\} + \frac{1}{C_v}\left\{ \tau_{ij}\frac{\partial u_j^*}{\partial x_j} + \frac{\partial}{\partial x_j}\left(\kappa \frac{\partial T^n}{\partial x_j} \right) \right\} \tag{16}$$

After solving the simultaneous linear equations of $(\tilde{\rho T})$, which is derived from the discretized equation of Eq.(16) with respect to space, we obtain $(\rho T)^{n+1}$ with Eq.(15).

With the similar procedures, Eq.(2) discretized with respect to time is given by

$$\frac{(\rho u_i)^{n+1} - (\rho u_i)^n}{\Delta t} + \alpha_{\rho u_i}\frac{\partial\{(\rho u_i)^{n+1} u_j\}}{\partial x_j} + (1 - \alpha_{\rho u_i})\frac{\partial\{(\rho u_i)^n u_j\}}{\partial x_j} = -\frac{\partial p^{n+1}}{\partial x_i} + \frac{\partial \tau_{ij}}{\partial x_j} - \rho^{n+1} g \delta_{i3} \tag{17}$$

where $\alpha_{\rho u_i}$ is a parameter whose range is $0 \leq \alpha_{\rho u_i} \leq 1$. The component $(\rho u_i)^{n+1}$ is defined by

$$(\rho u_i)^{n+1} = (\rho u_i)^n + (\tilde{\rho u_i}) \tag{18}$$

Substituting Eq.(18) into Eq.(17), we have

$$\frac{(\tilde{\rho u_i})}{\Delta t} + \alpha_{\rho u_i} \frac{\partial\{(\tilde{\rho u_i})u_j^*\}}{\partial x_j} = -\frac{\partial\{(\rho u_i)^n u_j^*\}}{\partial x_j} - \frac{\partial p^{n+1}}{\partial x_i} + \frac{\partial \tau_{ij}}{\partial x_j} - \rho^{n+1}g\delta_{i3} \tag{19}$$

After solving the simultaneous linear equations of $(\tilde{\rho u_i})$, which is derived from the discretized equation of Eq.(19) with respect to space, we obtain $(\rho u_i)^{n+1}$ with Eq.(18).

Applicability of the Numerical Method

Sod's Shock Tube Problem

Firstly, the present computational method was applied to one dimensional Sod's shock tube problem [Sod (1978)], in order to confirm that the method is able to capture discontinuity of the variables. Figure 2 shows the initial conditions on the rectangular domain with $l_1 = 0.3$ [m] and $l_2 = 1.0$ [m]. A diaphragm at $x_2 = 0.5$ [m] separates two regions which have different densities and pressures.

The two regions are initially in a static state. The variables in the initial conditions are given as follows:

$$\left.\begin{array}{ll} \rho_L = 1.0\,[\text{kg/m}^3] & , \quad \rho_R = 0.125\,[\text{kg/m}^3], \\ u_L = 0.0\,[\text{m/s}] & , \quad u_R = 0.0\,[\text{m/s}], \\ p_L = 1.0\,[\text{Pa}] & , \quad p_R = 0.1\,[\text{Pa}], \end{array}\right\} \tag{20}$$

where subscripts L and R denote the values on the left and right sides of the diaphragm respectively. T_L and T_R are determined by the equation of state. The ratio of specific heats γ is chosen to be 1.4 assuming that γ of the gas is similar to that of air. The specific heat at constant volume $C_v = 7.17 \times 10^2$ [J/(kg·K)]. The coefficient of thermal conductivity $\kappa = 0$ [W/(m·K)]. The coefficient of viscosity $\mu = 0$ [Pa·s] and the effect of gravity is negligible.

The boundary conditions are given as follows: A free-slip boundary condition is imposed on the top and bottom walls and a non-slip boundary condition is imposed on the left and right walls. On all walls, Neumann boundary conditions are employed: $\partial T/\partial n = 0$, $\partial P/\partial n = 0$ and $\partial \rho/\partial n = 0$.
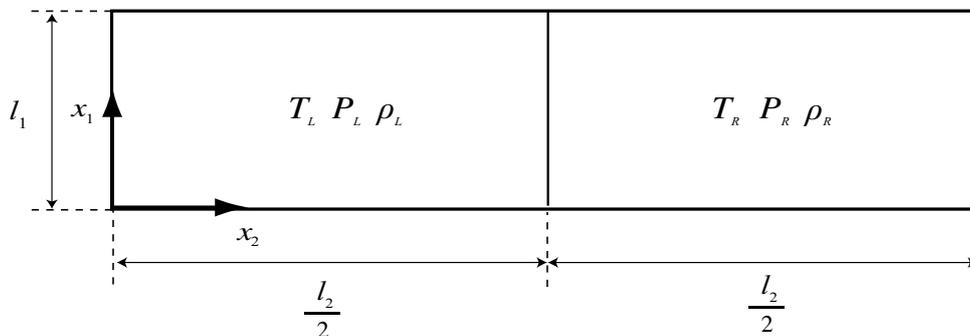
Figure 2. Initial conditions for Sod's shock tube problem

In the computations, the time step Δt is 4.0×10^{-4} [s] and the number of computational cells is 102×1502. All convection terms on the right-hand side in Eqs.(12), (16) and (19) are evaluated with the third-order MUSCL TVD scheme [Yamamoto and Daiguji(1993)].

Figure 3 shows the dimensionless mass error $err = (M_0 - M')/M_0$ at $t = t'$ [s]. Here M_0 and M' denote the total mass at $t = 0$ [s] and at $t = t'$ [s], respectively. It was shown that the mass conservation law is satisfied accurately in the present computational method.

The numerical results at $t = 0.2$ [s] are shown in Fig.4, in which internal energy e is given by Eq.(4). While the calculated internal energy in the range of $0.68 \leq x_2 \leq 0.85$ is somewhat smaller than the theoretical values, the other predicted results reasonably agree with the exact solutions. Thus, it was shown that the numerical method is applicable to the one-dimensional shock wave problem and that it is able to capture discontinuity without artificial viscosity.
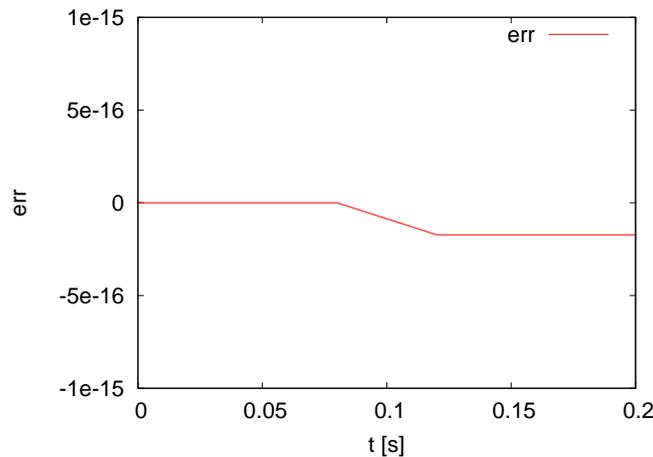
Figure 3. Dimensionless mass error
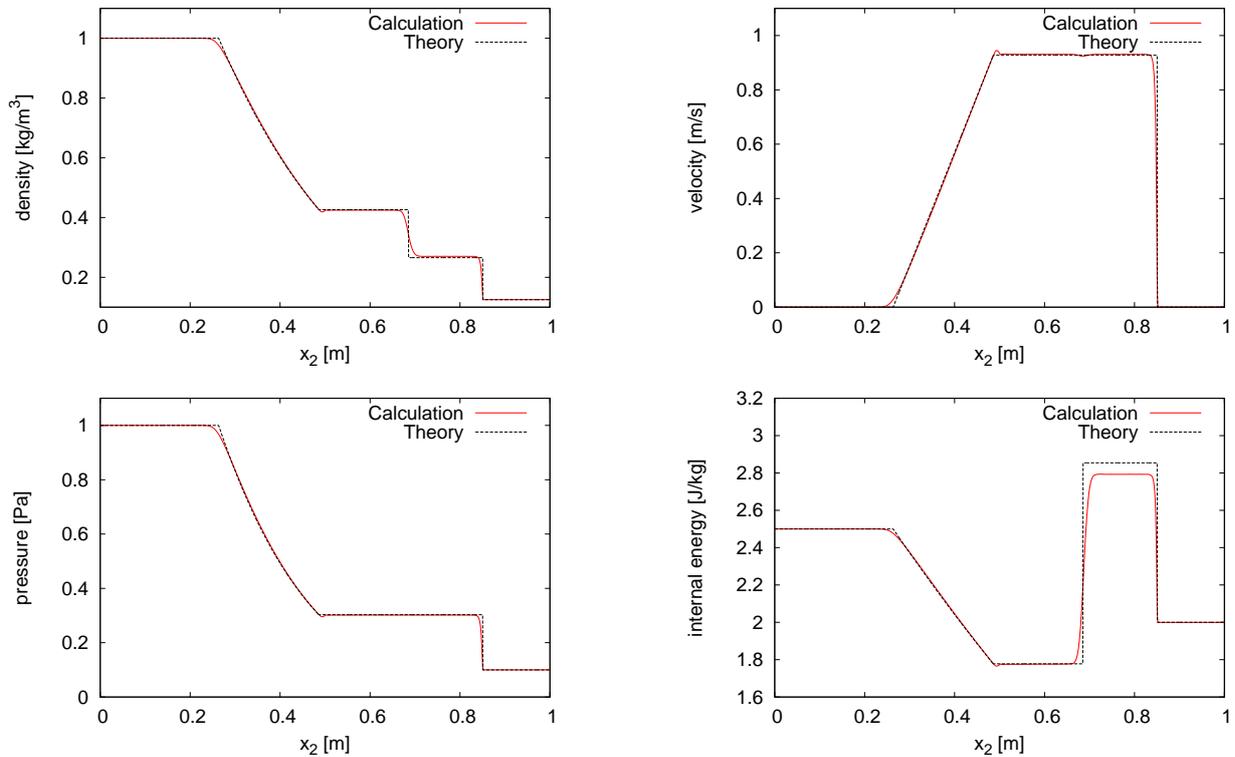
Figure 4. Numerical results ($x_1 = 0.5l_1$ and $t = 0.2$ [s])

820

Explosion Test in Two-Space Dimensions

Secondary, the present computational method was applied to the two-dimensional problem in the square domain of $l_1 = l_2 = 2.0$ [m], including a circular area at its center, as shown in Fig.5. The initial conditions are different between the circular region, whose radius r is 0.4 [m], and the region outside the circle.

The flow is in a static condition in all regions at $t = 0$ [s]. The initial conditions are given as follows:

$$\left.\begin{array}{lll}
\rho_{in} = 1.0\,[\text{kg/m}^3] & , & \rho_{out} = 0.125\,[\text{kg/m}^3], \\
u_{1in} = 0.0\,[\text{m/s}] & , & u_{1out} = 0.0\,[\text{m/s}], \\
u_{2in} = 0.0\,[\text{m/s}] & , & u_{2out} = 0.0\,[\text{m/s}], \\
p_{in} = 1.0\,[\text{Pa}] & , & p_{out} = 0.1\,[\text{Pa}],
\end{array}\right\}
\tag{21}$$

where subscripts *in* and *out* denote the values inside and outside the circle respectively. T_{in} and T_{out} are given by the equation of state. The ratio of specific heats γ was chosen to be 1.4 assuming that air behaves as an ideal gas. The coefficient of thermal conductivity $\kappa = 0$ [W/(m·K)]. The coefficient of viscosity $\mu = 0$ [Pa·s]. The time step Δt is 5.0×10^{-4} [s] and the effects of gravity is also negligible.

The boundary conditions are given as follows: A free-slip boundary condition is imposed on all walls. On all walls, $\partial T/\partial n = 0$, $\partial P/\partial n = 0$ and $\partial \rho/\partial n = 0$.

All convection terms on the right-hand side in Eqs.(12), (16), and (19) are evaluated with the third-order MUSCL TVD scheme [Yamamoto and Daiguji(1993)].The number of computational cells is 1002×1002.

The predicted results at $t = 0.25$ [s] are shown in Figs. 6, 7 and 8. As shown in Figs. 6 and 7, the density and pressure distributions qualitatively agree with the predicted values of [Toro(1997)]. Figure 8 shows a comparison between the present numerical results and the analytical solutions of [Toro(1997)]. While only the internal energy e in the range of $1.62 \leq x_1 \leq 1.81$ is smaller than the analytical solution, the other numerical results reasonably agree with the reference values. From the above results, it can be seen that the present numerical method is also applicable to two-dimensional problems.
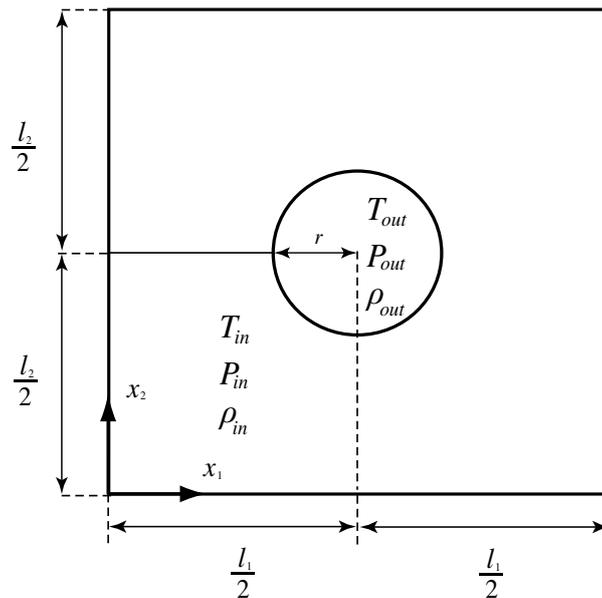
Figure 5. Initial conditions for cylindrical explosion
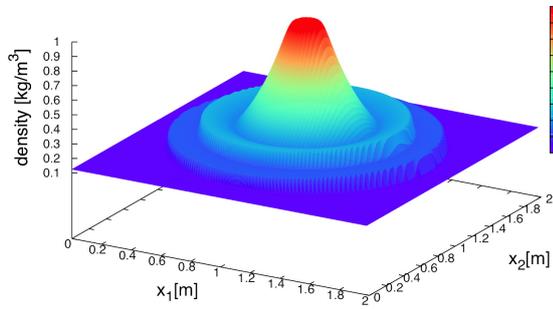
Figure 6. Density distribution at $t = 0.25$ **[s]**
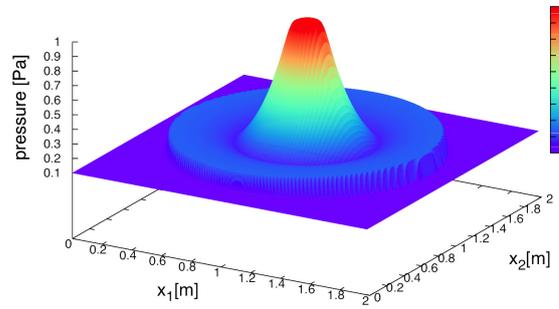
Figure 7. Pressure distribution at $t = 0.25$ **[s]**
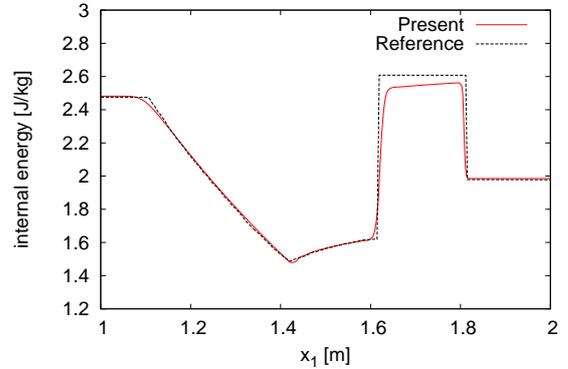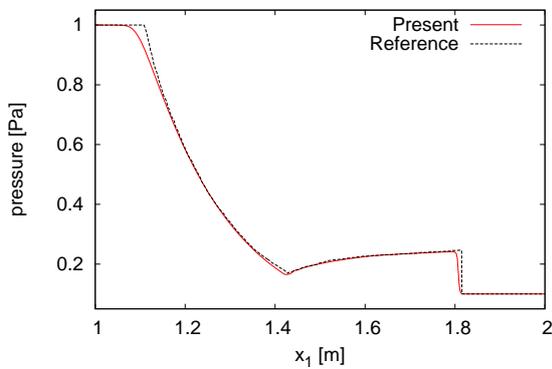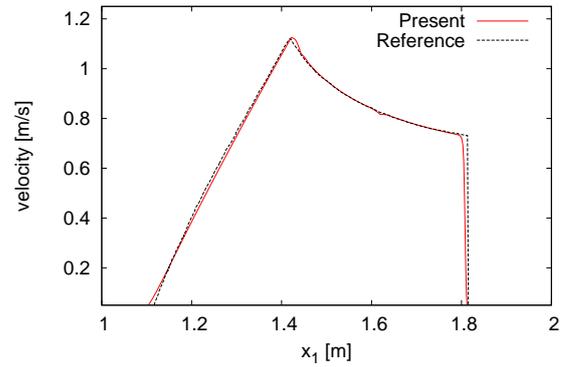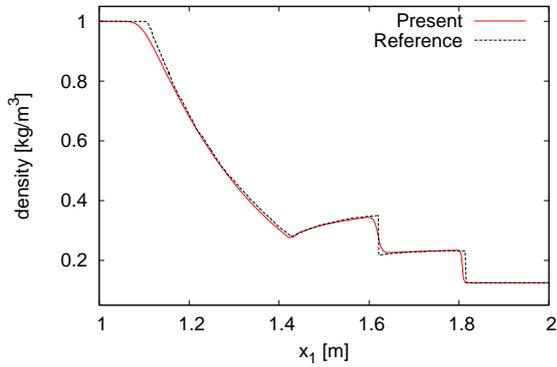
Figure 8. Numerical results and reference values [Toro(1997)] ($x_2 = 0.5l_2$ **and** $t = 0.25$ **[s]**)

Explosion Test in Three-Space Dimensions

As the final example of the high Mach number flows necessary to be solved by the large scale computations, the present computational method was applied to the three-dimensional shock-wave problem in the cubic domain with $l_1 = l_2 = l_3 = 2.0$[m], including a spherical region located in its center, as shown in Fig.9.

The initial conditions ($t = 0$ [s]) in the sphere of radius $r = 0.4$ [m] and the region outside the sphere are given the same as the two-dimensional case listed in Eq. (21). In addition, the physical properties and the gravity effects are also treated in the same way as the two-dimensional problem. The initial x_3 velocity components are set at $u_{3in} = u_{3out} = 0$ [m/s].

Th boundary conditions are given as follows: A free-slip boundary condition is imposed on all walls. On all walls, $\partial T / \partial n = 0$, $\partial P / \partial n = 0$ and $\partial \rho / \partial n = 0$.

All convection terms on the right-hand side in Eqs.(12), (16) and (19) are evaluated with the fifth-order compact upwind TVD scheme [Yamamoto and Daiguji(1993)].The time step Δt is 1.0×10^{-4} [s]. The number of computational cells is $802 \times 802 \times 802$. In order to solve this large scale problem efficiently, the computation was parallelized by flat MPI on the basis of a domain decomposition method.

Figure 10 shows a comparison between the present numerical results at $t = 0.25$ [s] and the analytical solutions of [Toro(1997)]. While some discrepancies are found in the distribution of the internal energy e in the range of $1.61 \leq x_1 \leq 1.79$ and near the discontinuous distributions of density and pressure compared with the analytical solution, the outline of the predicted results reasonably agree with the reference values. Therefore, it was shown that the present numerical method can be applied to three-dimensional problems as well.

In addition, the efficiency of the parallel computations was examined by changing the core numbers. Fig.11 shows the speed-up ratios on the basis of 32 cores in the Cray-XE6 in Kyoto University (AMD Opteron 2.5GHz, 32 cores / node, 64 GB memory / node). The maximum core number is 512 (16 nodes). As shown in Fig.11, the tendency of the increasing speed-up ratio is near the linear line. Thus, the present flat MPI parallelization enables us to decrease the elapse time satisfactorily.
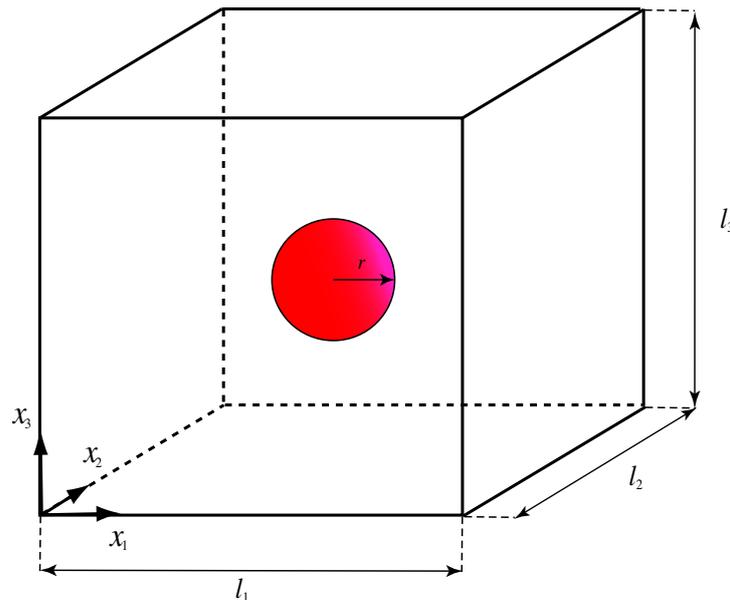
Figure 9. Calculation area for spherical explosion
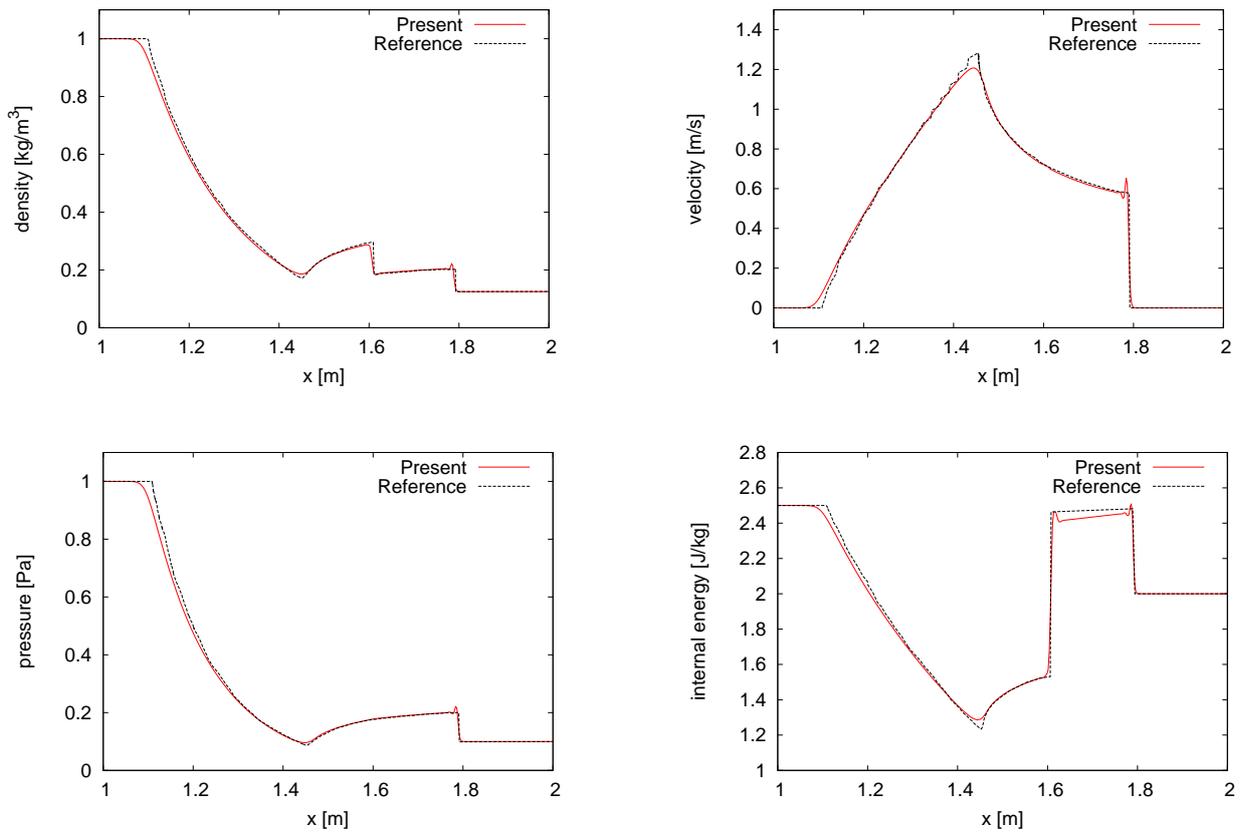
Figure 10. Numerical results (at $t = 0.25$ [s])
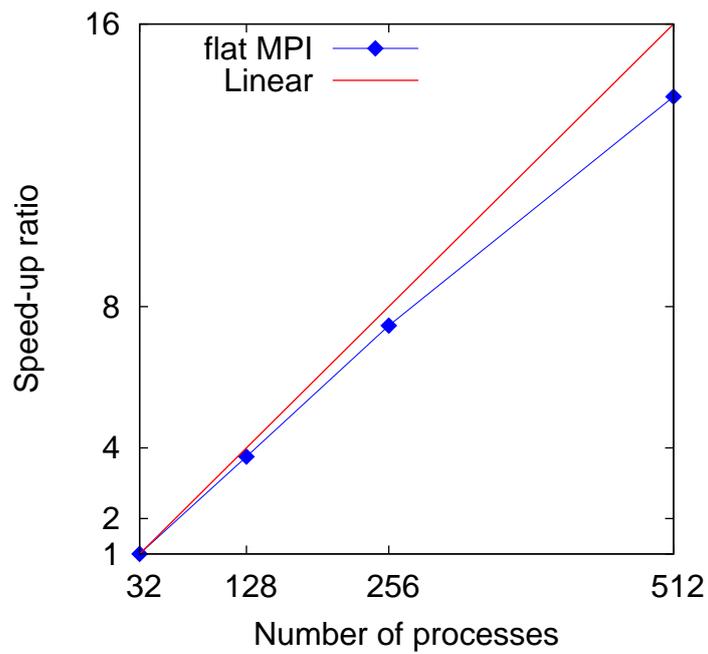
Figure 11. Speed-up ratios of flat MPI

Natural Convection in a Square Cavity

On the other hand, in order to confirm the applicability of the present method to the flows with low Mach numbers, it was applied to the natural convection in a square cavity with differentially heated side walls. While the compressibility is not so dominant in the example employed in this section, this is a suitable benchmark problem to confirm the applicability of the computational method.

As shown in Fig.12, the computational domain is a square cavity including the fluid of Prandtl number $Pr = 0.71$. The length of the square cavity is $l_1 = l_3 = 0.1$[m]. The acceleration of gravity affects in $-x_3$ direction. The the Rayleigh number Ra and Prandtl number Pr are defined as

$$Ra = \frac{Bg\Delta T l_3^3}{A\nu} \tag{22}$$

$$Pr = \frac{\nu}{A} \tag{23}$$

where A, B, μ, ν are thermal diffusivity, coefficient of thermal expansion, viscosity and kinematic viscosity, respectively.

The dimensionless numbers and the average Nusselt number are defined as

$$T^* = \frac{T_{cal} - T}{\Delta T} \quad , \quad x_1^* = \frac{x_1}{l_3} \quad , \quad x_3^* = \frac{x_3}{l_3}, \quad , \quad u_1^* = \frac{u_1 l_3}{A} \quad , \quad u_3^* = \frac{u_3 l_3}{A} \tag{24}$$

$$\overline{Nu} = \frac{1}{l_3} \int_0^{l_3} Nu \, dx_3 \bigg|_{x_3 = 0 \, or \, 0.1} \tag{25}$$

where T_{cal} is the predicted results of the temperature.

The initial conditions are given as follows: the velocity component u_i and the temperature T are set to zero and 300 [K] in the cavity. The pressure is set by considering gravity and the density is given by the equation of state with the initial pressure, temperature, the ratio of specific heats $\gamma = 1.4$.
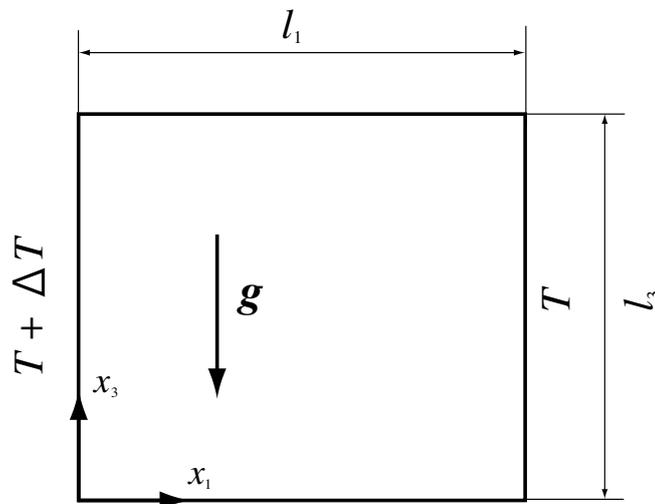
Figure 12. Calculation area for natural convection

The boundary conditions are given as follows: A non-slip boundary condition is imposed on all walls. $T = 300 + \Delta T$ [K] at $x_1 = 0$ [m] and $T = 300$ [K] at $x_1 = 0.1$ [m] where $\Delta T = 1$ [K]. The boundaries at $x_3 = 0$ [m] and 0.1 [m] are adiabatic, $\partial T / \partial x_3 = 0$ and $\partial \rho / \partial x_3 = 0$. On the vertical walls $\partial p / \partial x_1 = 0$, while p is set by considering gravity on the horizontal walls.

In the computations, The time step Δt is 8.0×10^{-7} [s]. The number of computational cells is 152×152. The predicted results were obtained at two Rayleigh numbers: $Ra = 10^4$ and 10^5. The coefficient of thermal conductivity $\kappa = 7.63467 \times 10^{-2}$ [W/(m·K)] and the coefficient of viscosity $\mu = 5.42061 \times 10^{-5}$ [Pa·s] at $Ra = 10^4$, while $\kappa = 2.41429 \times 10^{-2}$ [W/(m·K)] and $\mu = 1.71414 \times 10^{-5}$ [Pa·s] at $Ra = 10^5$.

Figures 13 and 14 show the calculated results with the present method. As shown in Figs.13 and 14, the isotherms and the isovels at steady state qualitatively agree with the predicted values of [Davis(1983)].

Table 1 shows the magnitude and the location of the maximum velocity u^*_{1max} along the vertical centerline, the maximum velocity u^*_{3max} along the horizontal centerline and the average Nusselt number $\overline{Nu}$ for Ra of 10^4 and 10^5. In Table1, the predicted results are compared with the benchmark results of [Davis and Jones(1983)] and [Dixit and Babu(2006)]. While u^*_{1max} at $Ra = 10^5$ at steady state is larger than that of [Davis and Jones(1983)], the other results reasonably agree with them.From these results, in addition to the high Mach number flows as shown in the above sections, it was shown that the present numerical method enables us to deal with the flows in low Mach numbers affected by viscosity and thermal diffusivity. It can be concluded that the numerical algorithm proposed in this paper is widely applicable to the flows from low to high Mach numbers.
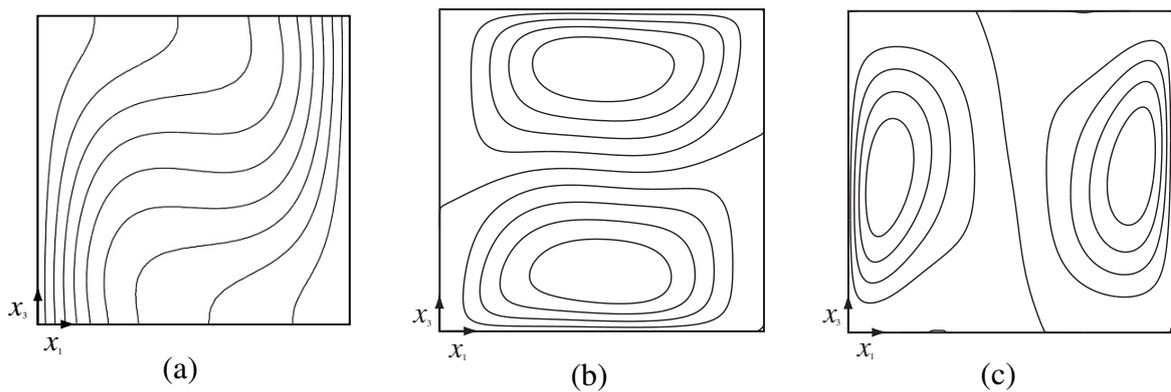
Figure 13. Contour maps for predicted results at $Ra = 10^4$: (a) T^*, (b) u^*_1, (c) u^*_3
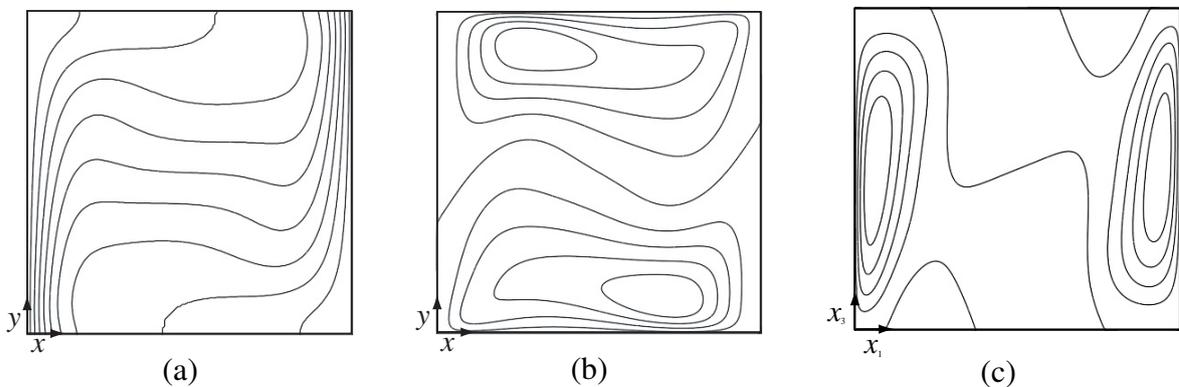
Figure 14. Contour maps for predicted results at $Ra = 10^5$: (a) T^*, (b) u^*_1, (c) u^*_3

Table 1. Comparison among numerical results

	Present		[Davis and Jones(1983)]		[Dixit and Babu(2006)]
Ra	10^4	10^5	10^4	10^5	10^4
u^*_{1max}	16.193	43.87	16.178	34.73	16.179
x^*_3	0.823	0.890	0.823	0.855	0.824
u^*_{3max}	19.649	68.52	19.617	68.59	19.619
x^*_1	0.117	0.063	0.119	0.066	0.121
$\overline{Nu}$	2.245	4.528	2.243	4.519	2.245

Conclusions

In this paper, the computational algorithms and numerical discretizations were newly proposed for governing equations of compressible fluids in an attempt to establish the numerical method which can be widely applicable to the flows from low to high Mach number. In this method, some numerical techniques proposed for the incompressible fluids are utilized on the collocated grid system on the basis of FVM. As a result, the conservation of variables are satisfied and the equation of state is established for $n+1$ time step variables. In addition, the parallelization with the domain decomposition method is easily employed due to the collocated grid system.

As a result of the computation of some benchmark problems, it has been shown that the present computational method enables us to capture shock discontinuity without artificial viscosity. It was also confirmed that the present method is applicable to the natural convection flow with low Mach number affected by viscosity and thermal conductivity.

References

S. Ushijima and I. Nezu (2002) Higher-order implicit (C-ISMAC) method for incompressible flows with collocated grid system, *Journal of JSCE*, **719**, 21-30.

B. R. Shin, T.Ikohagi, and H. Daiguji (1993) An unsteady implicit SMAC scheme for two-dimensional incompressible Navier-Stokes equations, *JSME International Journal*, **36**, 598-606

S. Yamamoto and H. Daiguji (1993) Higher-order-accurate upwind schemes for solving the compressible Euler and Navier-Stokes equations, *Computers Fluids*, **22**, 259-270.

Gray A. Sod (1978) A survey of several finite difference methods for systems of nonlinear hyperbolic conservation laws, *Journal of Computational Physics*, **27**, 1-31.

E. F. Toro (1999) *Riemann Solvers and Numerical Methods for Fluid Dynamics*, 2nd edn, Springer-Verlag, Berlin.

G. de Vahl Davis (1983) Natural convection in a square cavity : a bench mark numerical solution, *International Journal for Numerical Methods in Fluids* **3**, 249-264.

G. de Vahl Davis and I. P. Jones (1983) Natural convection in a square cavity : a comparison exercise, *International Journal for Numerical Methods in Fluids*, **3**, 227-248.

H. N. Dixit and V. Babu (2006) Simulation of high Rayleigh number natural convection in a square cavity using the lattice Boltzmann method, *International Journal of Heat and Mass Transfer*, **49**, 727-739.

Comparative Study of Sparse Matrix Storage Format in the Finite Element Analysis of Thermal-Structure Coupling Problems

Abul Mukid Mohammad Mukaddes[1], Ryuji Shioya[2] and Masao Ogino[3]

[1]Department of Industrial and Production Engineering
Shahjalal University of Science and Technology, Bangladesh
Email: mukaddes1975@gmail.com
[2]Faculty of Information Sciences and Arts
Toyo University, Japan
Email: shioya@toyo.jp
[3]Information Technology Center, Nagoya University, Japan
Email: masao.ogino@cc.nagoya–u.ac.jp

Corresponding author: mukaddes1975@gmail.com

Abstract

The aim of this research is to compare different sparse matrix storages schemes in the finite element analysis of thermal-structure coupling problems. Thermal-structure coupling approach has been developed using the ADVENTURE System. The approach relies upon the existing module ADVENTURE Thermal, which handles the heat conduction problems to have the temperature distribution in the solid model and a module named ADVENTURE Solid which takes care of the stress analysis. Quite commonly, the matrix that participates in the finite element computation of thermal and structural problem is sparse. The present adventure modules which are based on the domain decomposition method use the sparse matrix-vector multiplication (SpMxV) as their basic operation. Sparse matrix by definition, are populated primarily with zeroes and thus special storage formats are used to enable efficient storage and computational operations. These representations usually store the non-zero values of the matrix with additional indexing information about the position of these values. A memory efficient storage format is proposed in this research. In the proposed technique, the inherent block sizes present in the sparse matrix are exploited to reduce the memory requirement as well as computation time. A SpMxV library has been developed that could be used both thermal and structural problems. Impacts of sparse matrix storages formats on the total execution time of the solver are evaluated. A 3D blast furnace cooling stave is analysed efficiently in terms of computation time and memory using the developed approach.

Keywords: thermal-structure coupling; sparse matrix; indexing; cooling stave; ADVENTURE System

Introduction

With the advent of high performance computer, efficient modules for the finite element analysis (FEA) to solve large scale problems are the present demand of FEM users. The ADVENTURE [adventure] is open source CAE software that has been developed for large scale analysis and design of computational mechanics system. This software is able to analyze three-dimensional (3-D) finite element models of arbitrary shape with 10-100 million degrees of freedom (dof). We have been developing the ADVENTURE system for the future high performance computer as a memory and time efficient FEA module. In this research a thermal-structure coupling system has been developed using two of ADVENTURE modules, ADVENTURE_Thermal and ADVENTURE_Solid. The

developed system could be used to analyze heat transfer problems that have complex geometries for the temperature distribution. The predicted temperature combined with the applied external load is then used to compute the deformation and thermal stresses of the model. Both thermal and solid modules use the parallel finite element method called domain decomposition method to solve problems in parallel computers.

The performance of the domain decomposition method is dominated by the sparse matrix-vector multiplication (SpMxV), $y \rightarrow y + Ax$ where, A is a sparse matrix and x, y are dense vectors [Mukaddes (2014)]. The method is also influenced by the preconditioning techniques [Ogino (2011)]. Sparse matrix by definition, are populated primarily with zeroes and thus special storage schemes are used to enable efficient storage and computational operations. These representations usually store the non-zero elements of the matrix with additional indexing information about the position of these values. A variety of compressed sparse row (CSR) format is used to store and manipulate the sparse matrix for the thermal problem [Mukaddes (2014)]. The matrix in the structural problems has inherent block shape. In order to reduce the memory requirement, exploiting the block shape of the matrix could be a beneficial choice. In this research, a diagonal block compressed sparse row (DBCSR) format for the structural problem are proposed and compared with other formats. Here instead of entire rows or columns of a matrix, block algorithms operate sub-blocks or data blocks.

The proposed storage formats are evaluated considering two models: High Temperature Test Reactor (HTTR) and 3D cooling stave [Kumar et al. (2012), Lijun et al. (2006)]. The numerical results obtained are presented and discussed.

2. Thermal-Solid coupling analysis

The developed system is conducted to predict temperature distribution in solid models and then to investigate the thermal expansion or deformation due to the temperature change. Analysis steps are as follows.

1) Read the input data for the heat conductive analysis and decompose the model by ADVENTURE_Metis.

2) Analyze heat conduction problems using ADVENTURE_Thermal.

3) Gather temperature of all nodes of the model from outputs of heat conduction problems.

4) Read temperature of all nodes and other input data for structural analysis and then decompose the model by ADVENTURE_Metis again.

Figure 1 shows the flow chart of thermal-solid coupling analysis with the developed system. The name of the ADVENTURE module used in each analysis is shown in parentheses.
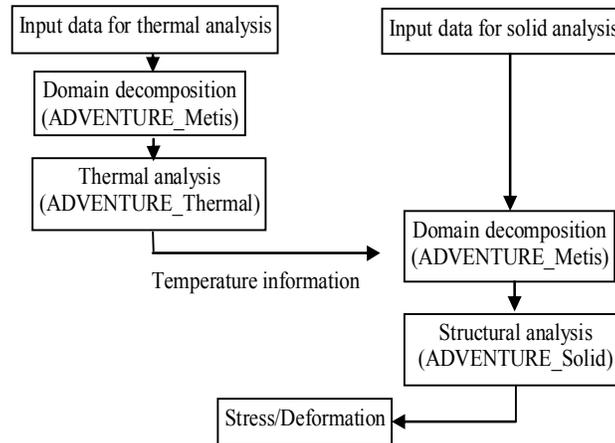
Figure 1 Flow chart of thermal-solid coupling system.

3. Sparse storage formats and their implementations

Sparse matrix storage schemes are described in this section. Consider the lower part of a typical symmetric matrix A. This matrix can be stored using different storage schemes. Several sparse matrix storage schemes are studied in this research and evaluated in both thermal and structural analysis. They are Compressed Sparse Row (CSR), Compressed Sparse Column (CSC), Variable Block Compressed Sparse Row (VBCSR) and Diagonal Block Compressed Sparse Row (DBCSR). The DBCSR are proposed in this research and compared with other storage format. For simplicity Skyline storage, CSR storage and DBCSR storage formats are explained for the example matrix A.

Matrix A

4.1 Skyline or Variable Band (SKY)

The Skyline representation becomes popular for direct solvers especially when pivoting is not necessary. This is the most common matrix storage format. The matrix elements are

stored using three single arrays: *data, row_index and row_pattern.* The array *data* stores the elements of the matrix *A* row by row, *row_index* contains column number of first element of each row and *row_pattern* array points to the start of every row.

Table 1: Skyline format

data:

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	**0**	**0**	**0**	**0**	**0**	**0**	40	41	
42	43	**0**	**0**	**0**	**0**	**0**	**0**	44	45	46	47	48	**0**	**0**	**0**	**0**	**0**	**0**	49	50	51		

row_index:

1	1	1	1	1	1	3	3	3	1	1	1

row_pattern:

1	2	4	7	11	16	22	26	31	37	47	58	69

4.2 Compressed Sparse Row (CSR)

Compressed sparse row format [Saad (1994)] is popular and the most general purpose storage format for the sparse matrix. The elements are stored using three arrays: *data*, *row_pattern* and *col_index.* The single array *data* of length number of nonzero (*nnz)* contains the non-zero elements of *A* row by row, *col_index* of length *nnz* contains the column number which correspond to the non-zero elements in the array *data.* The integer vector *row_pattern* of length *nrow+1* contains the pointers to the beginning of each row in the array *data* and *col_index.* With the *row_pattern* array we can easily compute the number of non-zero elements in the ith row as *row_pattern*[i+1] - *row_pattern*[i]. The last element of *row_pattern* is *nnz.* The CSR representation of the symmetric matrix *A*:

Table -2 CSR format

data:

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41							
42	43	44	45	46	47	48	49	50	51														

col_index:

1	1	2	1	2	3	1	2	3	4	1	2	3	4	5	1	2	3	4	5	6	4	5	6
7	4	5	6	7	8	4	5	6	7	8	9	1	2	3	10	1							
2	3	10	11	1	2	3	10	11	12														

row_pattern:

1	2	4	7	11	16	22	26	31	37	41	46	51

4.3 Diagonal Block Compressed Sparse Row (DBCSR)

 The DBCSR exploit the 3x3 block shape of the matrix. In this format the diagonal block can be stored separately in a *diag* array which does not require the indexing. The off-

diagonal elements are stored in the data array. *index_brow* represents the position of first element of the first block in the data array. The last element of the *index_brow* is the number of elements in the off-diagonal blocks. *Index_bcol* stores the column number of the first elements of each off-diagonal blocks. Such indexing reduces the working sets as well as memory requirements.

Table-3 DBCSR

diag:

1	2	3	4	5	6	10	14	15	19	20	21	25	29	30	34	35	36	40	44	45	49	50	51

data:

7	8	9	11	12	13	16	17	18	22	23	24	26	27	28	31	32	33	37	38	39	41	42	43	4

index_brow:

1	10	19	27

index_bcol:

1	4	1

4.4 Working sets of sparse matrix storage formats

The skyline format takes more memory than others as it needs to store some unnecessary zero elements. The CSR requires less memory than skyline, as it does not store zero elements of the matrix. DBCSR reduces the memory requirement of the indexing part of CSR. The working sets of three storage formats are given below.

Table-4 Working sets

SKY	CSR	DBCSR
elm + 4(row+1+row)	8nnz + 4(row+1+nz)	8nnz + 4(brow +1 + brow)

elm: num. of matrix elements; nnz: num. of non-zero; row: num. of rows; brow: num. of block rows

5 Numerical results and discussions

Evaluation of sparse matrix storage schemes:

A large scale High Temperature Test Reactor (HTTR) model is used to evaluate the sparse matrix storage schemes. The computational environment is Intel Corei7-960 (3.20GHz/L2 256KB/L3 8MB/QuadCore). The DBCSR are evaluated and compared with other sparse matrix storage fromats. Using the CSR type storage format computation time is reduced to around 50% and required memory is reduced to around 45% comapared to the skyline storage format. Again, DBCSR shows better performance over the CSR format. It reduces 21% computation time and 15% required memory compared to the CSR format.

Figure-2 Computation time for thermal (left) and sturcture (right)

Figure-3 Required Memory for thermal (left) and structural (Right)

Cooling stave analysis:

In the present paper, a cooling stave of a blast furnace has been modeled and analyzed using the developed system. Parts name are shown in figure 4a. For the finite element analysis, the 3D geometry of the cooling stave is made using the commercial CAD software, Meshman. After that the model is exported to .iges file. Then the file is imported to the ADVENTURE Systems. For simplicity, a part model of a cooling stave (figure 4b) is analysed. Results of full model are given later. The design parameters and material properties are taken from [Kumar et al. (2012)]. The boundary conditions for thermal and solid are set up as follows:

- Air temperature is 323 K, water temperature is 303 K.

- Heat convection coefficients: between furnace shell and atmosphere-12 W/(m² K), between water and inner sides of the furnace shell -8000 W/(m²K).
- The upper and lower surface are fixed.

The cooling channels are shown in figure 4b, where convection boundary conditions are considered. The figure 3a shows the temperature distribution after the thermal analysis using the ADVENTURE_Thermal. The ADVENTURE_PostTool is used to visualize the temperature, displacement and stress. The temperature information is used in ADVENTURE_Solid as loads to measure the thermal expansion and stresses. The thermal expansion (X 100) in the y direction is shown figure 3d and corresponding nodal equivalent stress is shown in figure 4a. Numerical results depict that the stress intensity on the cold surfaces is mainly affected by the cooling water and much higher on the hot sides.
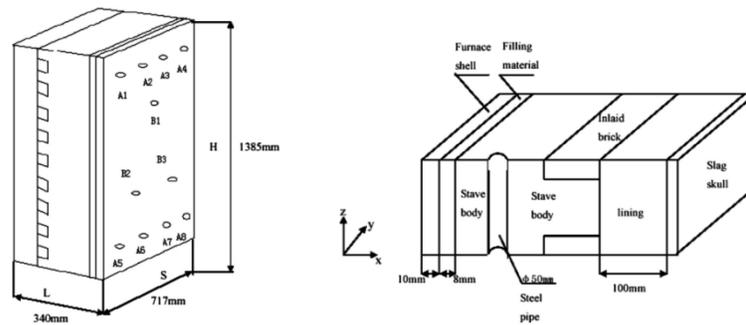
Figure-4 Cooling stave full (left) and part (right)
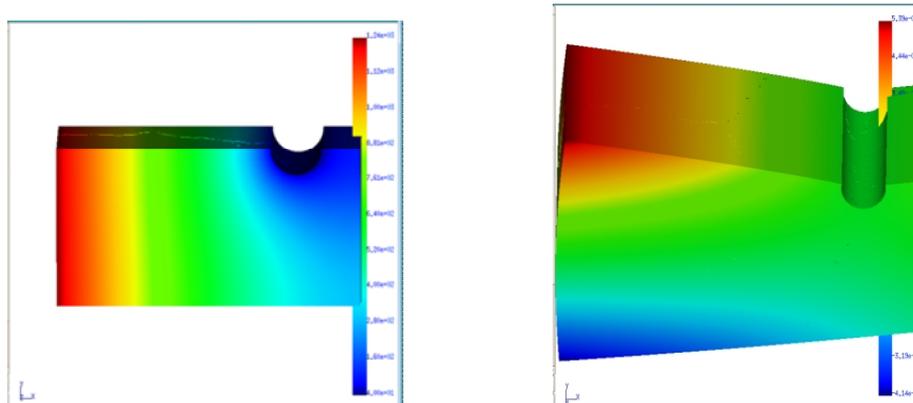
Figure 5: Temperature distribution (left) and expansion (x100, right)

(a)   (b)   (c)
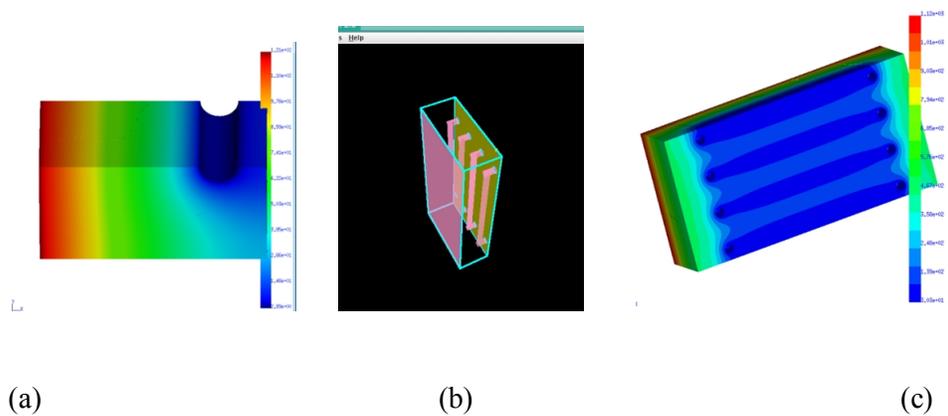
Figure 5: (a) Equivalent stress (b) Cooling channel (c) Temperature distribution of full model

Finally, a full model of cooling stave is modelled and analysed as shown in figure 5b. After thermal analysis of the full model it is found that, the temperature on the top and bottom surfaces are higher than other regions and maximum values are on the hot sides (figure 5c).

6. Conclusion

In this research, the developed thermal-solid coupling system is successfully implemented to analyze the cooling stave of a blast furnace. The intensity of stress is reduced due to the cooling system. The computational cost of the developed system is improved dramatically by employing several sparse matrix storage formats. Several sparse matrix storage formats are implemented and compared. DBCSR type storage format shows better performance compare to the previous Skyline and CSR storage format. Future research is to compare different blocking strategies of DBCSR storage format.

Acknowledgements

References

http://adventure.sys.t.u-tokyo.ac.jp/

Kumar, A., Bansal, S. N. and Chandraker, R., (2012), Computational modeling of blast furnace cooling stave based on heat transfer analysis, *Material Physics and Mechanics*, 15, 46-65.

Mukaddes, A M M., Ogino, M. and Shioya, R. (2014), Performance study of domain decomposition method with sparse matrix storage schemes in modern supercomputer, *International Journal of Computational Method (in press)*.

Lijun, W, Weiguo, Z. Huier, C. Yunlong, S., Xiaojing Li and Canyang, S. (2006), The study of cooling channel optimization in blast furnace cast steel stave based on heat transfer analysis. *International Journal of Advanced Manufacturing Technology*, 29, 64-69.

Ogino, M., Shioya, R. and Kanayama, H. (2008) An inexact balancing preconditioner for large-scale structural analysis, *Journal of Computational Science and Technology*, 2- 1, 150-161.

Saad, Y., (1994) SPARSEKIT: A basic tool kit for sparse matrix computations. Technical report, Computer Science Department, University of Minnesota, Minneapolis, MN 55455, Version 2.

Shahnaz, R. and Usman, Anila (2011), Blocked-based sparse-matrix vector multiplication on distributed memory computer, *The International Arab Journal of Information Technology,* 8- 2, 130-136.

Challenge of Hydrodynamic Analysis for a Structure in Waves

*Frank Lin, Norbert Bakkers, Michael Johnson, Yongwon Lee, Neil Southall, Zhenhong Wang, and Nigel White

Lloyd's Register Group Ltd

*Corresponding author: frank.lin@lr.org

Abstract

Hydrodynamic analysis is one of the key steps in safety assessment of a structure in waves. Many options are available for answering challenge raised from marine and offshore energy industry, from costly three dimensional CFD to the efficient but not perfect boundary element models. Focus on the boundary element methods, analysis methods for the interaction of waves and structures are discussed. Those boundary element models cover frequency domain and time domain, linear and non-linear. Special attention is pay to the problems encountered in those models and approaches we adopted for their engineering solution.

Keywords: Boundary element method, Frequency domain, Time domain, Linear, Non-linear, multi-level.

Introduction

Hydrodynamic analysis with reliable accuracy is the first step in a successful structure assessment. This analysis in marine and offshore industry is usually dominant by the interaction of ocean waves and floating or fixed structure, and it seeks for a solution of a gravitational water wave field in an infinite fluid domain around the structure. Varies numerical methods have been applied in this industry area, such as RANS, SPH, Rankin source distribution method, Green's function based boundary integration method, and so on. They can be categorized in CFD method class and boundary element method (BEM) class. RANS, as the typical CFD method, is the most robust method in this area. It performs the time domain simulation and has the capability to solve most of the problems, but the high computation cost is still the main obstacle to allow it been used in routine seakeeping analysis for design and design appraisal. Comparing to CFD model, boundary element class is efficient and has different models for analysis in time domain and frequency domain, and for analysis of linear and non-linear problems. The most efficient tool in this class is the Green's function based linear frequency domain model. It can solve a few thousands of regular wave cases in one day on a high-end laptop. To take advantage of this efficient, time domain boundary element tool based on frequency domain analysis results is developed to capture the so-called geometry non-linear which dominants the solution of ship/offshore-structure response in large waves. The most expensive boundary element tool is the Rankin source/panel model which can solve nonlinear seakeeping problem and has higher uncertainty and human effect in model setting, but it is still much cheaper in use comparing to CFD models. What numerical model should be applied for a specific problem is the question that every engineer needs to answer. How to extend the existing model for more complicate analysis is the challenge for researchers in industry. In this paper, we discussed some of practices in Lloyd's Register dealing with linear and nonlinear hydrodynamic assessment.

Nonlinear Viscous Damping in Potential Flow Modeling

Frequency domain BEM model is a linear analysis tool because hydrodynamic forces in this model, like wave exciting force, wave making added-mass and damping, are linear. On the other hand, non-linear factors can be involved in ship motion equation as external force and modify the results of hydrodynamic pressure and load with effects of the nonlinear factor. A typical example is the viscous roll damping. For roll motion, wave making added-mass and damping is not the dominant component for ships with conventional hull form and the hydrodynamic force from viscous flow

becomes important. A common approach is to enforce a nonlinear viscous damping moment in roll motion equation to correct the motion prediction. An analysis example is shown in Figure 1. The normalized roll motion result without viscous roll damping (VRD) is plotted in the left plot of the figure. Blue marks show the experimental data and red line is from the computation of WAVELOAD-FD, a seakeeping analysis software package of Lloyd's Register. The predicted roll motion peak at the roll natural frequency is of 32.5 and 15 times larger than the experimental observation. After adding the VRD effects in WAVELOAD-FD model, the predicted roll motion is in a good agreement with the observation as shown by the red line in the middle plot of the figure. The non-linear Ikeda roll damping model is applied in this example. In the right plot of this figure, the pressure distribution on surface of hull and bilge keel is presented at a time when the ship is rolling counter-clock wise. The orange colour indicates a higher pressure area and the light blue for the lower pressure area. The non-linear viscous damping is involved in the pressure computation.
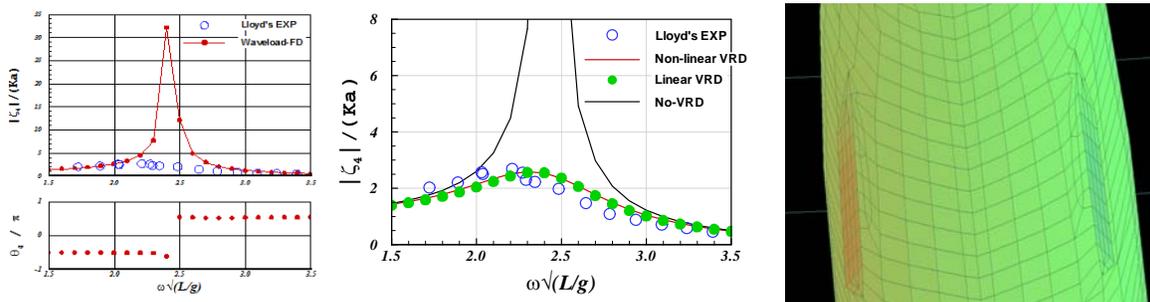
Figure 1. Nonlinear viscous damping in roll motion from a boundary element model
Left: without VRD; Middle: with VRD; Right: Pressure on hull and bilge keel

This example demonstrates that some nonlinear factor can be correctly taken into account in linear BEM model. For this viscous roll damping case, a further study revealed that the VRD can also be represented by an equivalent linear roll damping model as shown by the green marks in the middle plot of Figure 1.

Viscous flow damping also plays import role on structures with tubular members, like some offshore rags and pipelaying vessels. A pipelaying vessel assessment is used here to demonstrate a combination of nonlinear hydrodynamic model and linear boundary element seakeeping approach. Viscous flow will affect not only the roll but also other motion modes for this case. The panel model of the vessel with the stinger is shown on right of Figure 2, and the stinger configuration and force definition are presented on left of the figure.
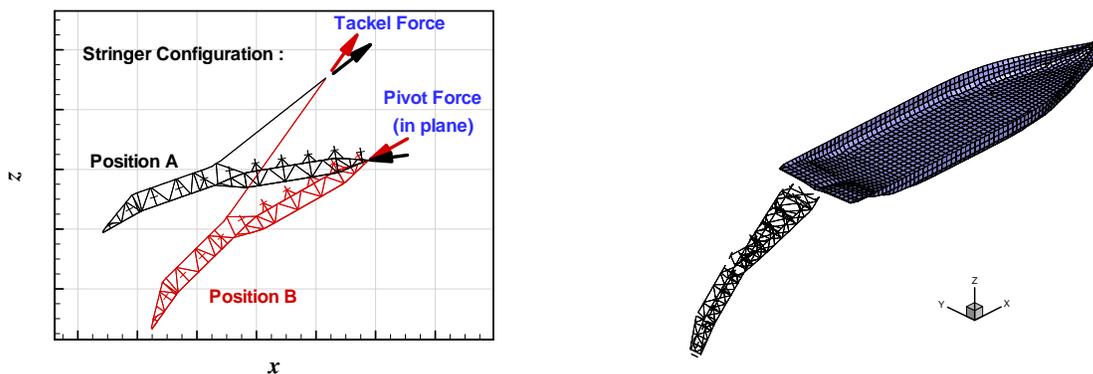
Figure 2. Model of a pipelaying vessel with stinger
Left: two position of the stinger; Right: under water part of the vessel and stinger

A CFD model for this problem will not be a practice choice and a model based on nonlinear Morrison formula would be considered. The force in the normal plane of each tubular member can be estimated from the relative velocity and acceleration between the structure motion and the flow around it. It is easy to compute this Morrison force for a fixed structure, but not for a floating one, as the vessel motion and Morrison force are coupled. In many available codes, the Morrison force model is involved in time domain BEM model. The time domain boundary element computation takes much shorter time than a CFD analysis, but its computing time is still beyond the acceptable level for design or design appraisal work. To answer the requirement, a Morrison force model module has been added in the frequency domain BEM code of Lloyd's Register, WAVELOAL-FD. The solution of this Morrison force coupled analysis is obtained from an iteration process. The Morrison force is treated as an external force in ship motion equation. In the first iteration, ship motion is obtained without Morrison force and then this ship motion is used to compute the first estimation of the Morrison force; the computed Morrison force is taken into account in ship motion solution of second iteration and repeat the 1st or previous iteration computation again for the new Morrison force. This iteration continues until both ship motion and Morrison force converged. Motion RAO of the vessel at zero ship speed and 150 degrees of heading is shown in Figure 3. The stinger decreases the ship motion and shifts the natural frequency of roll and pitch to high frequency side, and obviously the Morrison force coupling effect needs to be considered in the analysis.
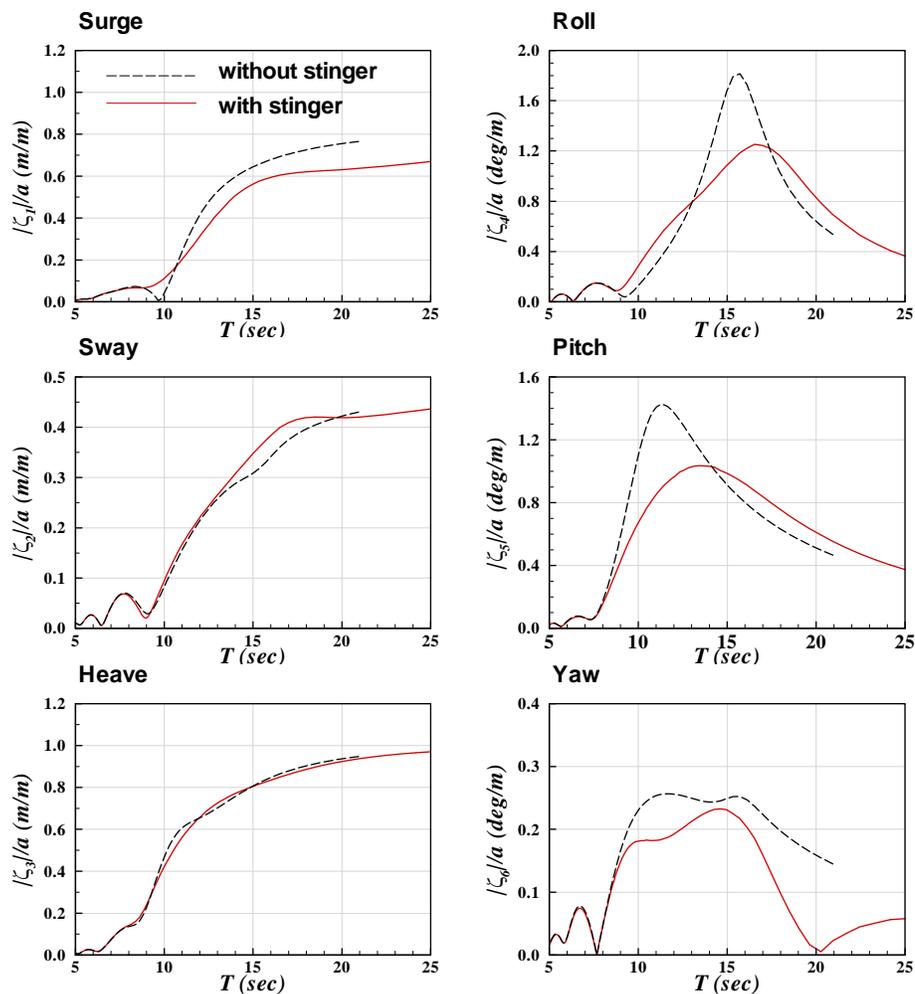
Figure 3. Motion RAO of the pipelaying vessel at zero forward speed and 150 degrees of heading
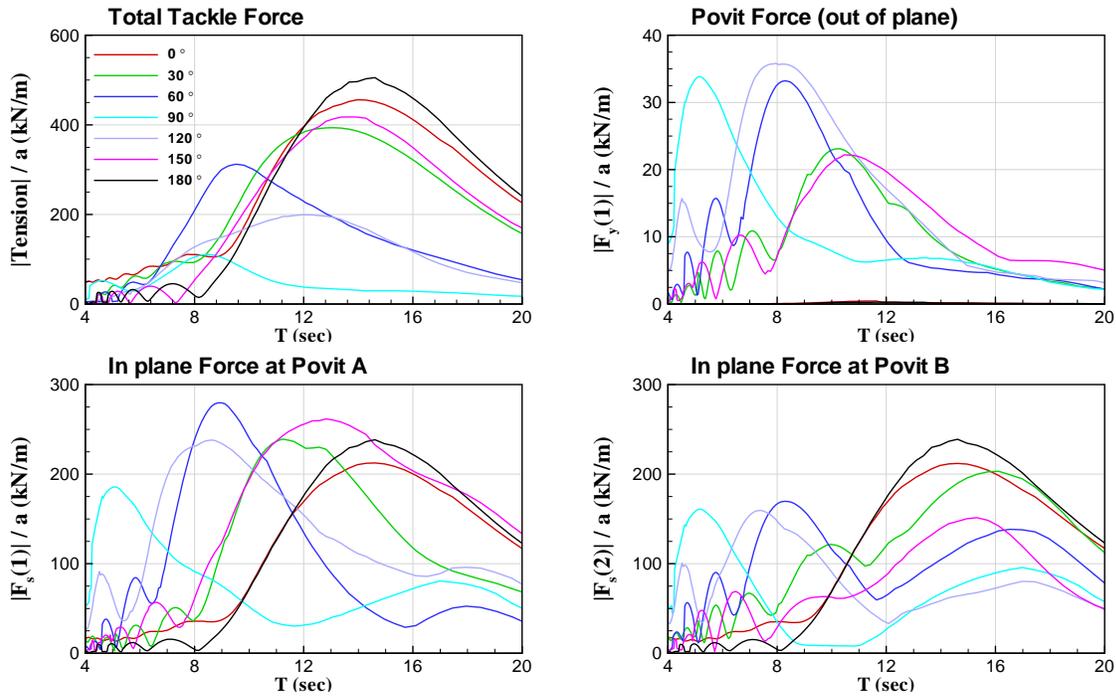
Figure 4. RAO of tackle force and pivot force

The stinger force result has been used to in a short-term statistic computation for a maximum value check. We found from Figure 5 that, in a high sea state where Tp=10 sec., the maximum stinger force based on frequency domain analysis is significantly large than those obtained from a time domain boundary element computation. The reason is that only one seed is used in the time domain analysis due to its long computation time. In general more seeds are required to obtain a reliable time domain simulation.
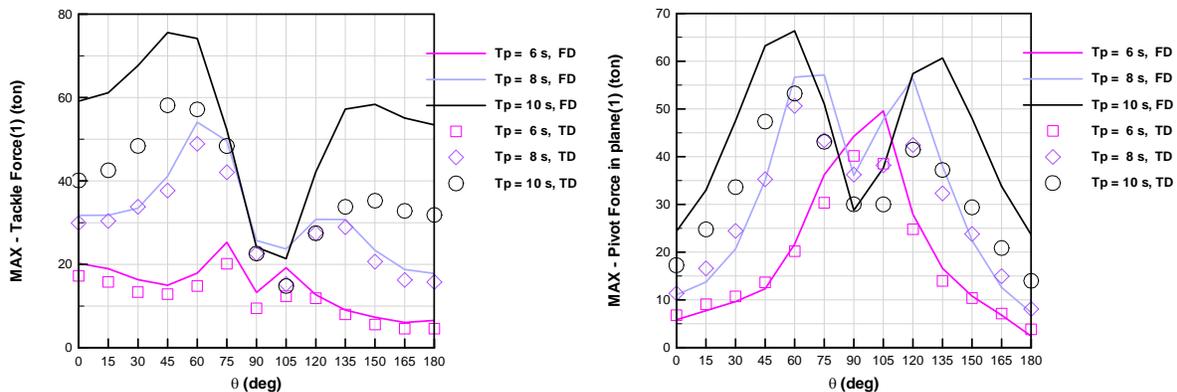
Figure 5. Comparison of 3 hours maximum force by the time domain (TD) and frequency domain (FD) analyses

Tank Sloshing of LNGC/FLNG

The nonlinear viscous flow force has been successfully involved in linear boundary element model in examples of previous section. But we do not always have luck to do so. Sloshing load on wall of partially filled tank is a good example. Structural damage, especially the fatigue one, on the tank

wall is the major threaten for a LNGC and sometimes FLNG. The tank wall damage is induced by the sloshing load due to the liquid flow inside the tank. A large LNGC can have a length of 350 meters; one large LNG tank can be 80,000 cubic meters. For this scale of the vessel and tank, a full ship CFD model, including both fluid domain around the ship and liquid domain inside tanks, will be out of consideration. Instead of that, a combined frequency domain boundary element model and CFD tank model will be selected. In the sloshing coupled boundary element model, an individual tank boundary element model will be adopted to solve the so-called radiation problem of the liquid flow inside a tank. This solution is still the velocity potential one exclude the non-linear flow effect. Adding those added-mass and damping from each partially filled tank in ship motion equation, the sloshing coupled ship motion can be solved. A model of a LNGC with two partially filled tanks is shown on left of Figure 6; the roll motion RAO at zero ship speed in beam seas from an analysis with and without tank flow effect is shown in the right of the figure. The ship motion is totally different when effect of liquid flow in the two tanks is involved; the roll motion from the sloshing coupled model has two peaks instead of one. In Figure 7, other two lateral motions, sway and yaw, in the same condition are presented. The filling ratio is 50% in both tanks. Red lines with name "FDWL" is the results obtained by WAVELOAD-FD, the sloshing coupled BEM model in frequency domain; and the blue marks are the experimental results. The RAO of lateral and vertical total force on the fore tank of the model are presented in Figure 8. The numerical results of the forces correlate with experiments well. These example shows that the inviscid linear BEM model works well for the global responses, ship motion and total tank force. But this model has a time harmonic tank wall pressure prediction and cannot predict the sloshing pressure peaks in a reliable accuracy. A CFD model is then required. An OpenFOAM based tank sloshing CFD tool, Aquarius, has been developed in Global Technology Center of Lloyd's Register in Southampton. The sloshing coupled ship motion will be computed first by WAVELOAD-FD, and the resultant ship motion will be used to drive the CFD tank model to simulate the pressure distribution due to the sloshing. In Figure 9, a 3D and 2D flow pattern obtained by Aquarius are presented, and pressure time history at different tank wall locations are plotted in Figure 10. The sharp peak of the pressure due to sloshing has been well captured.

This gives an example of using different level of numerical models in one hydrodynamic assessment for industry application.
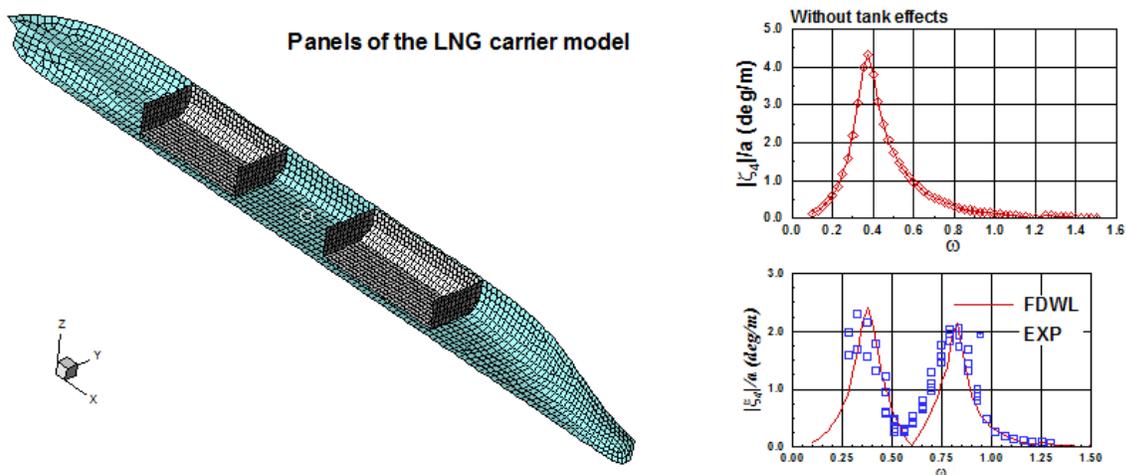
Figure 6. Model of a LNGC with two tanks (top left) and roll RAO without (top right) and with tank flow coupling (bottom right); at zero forward speed in beam sea.
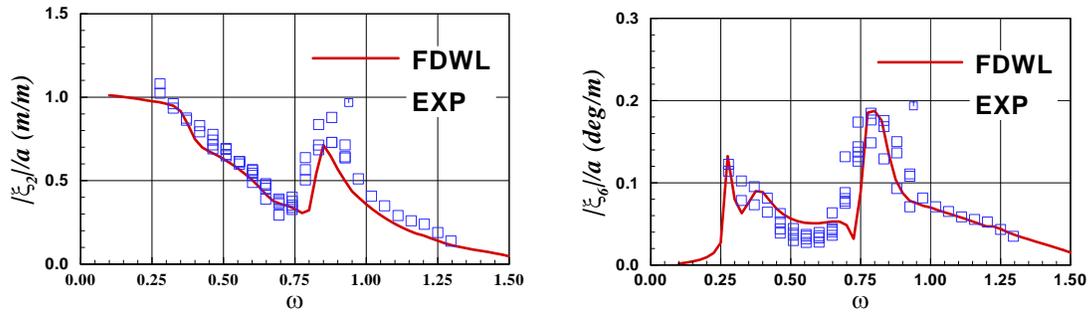
Figure 7. Sloshing coupled motion: sway (left) and yaw (right) at zero forward speed in beam seas.
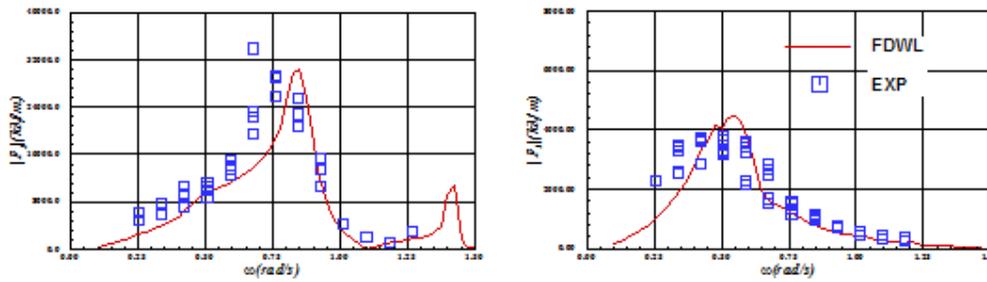
Figure 8. RAO of total tank force: lateral (left) and vertical (right) at zero speed in beam sea.
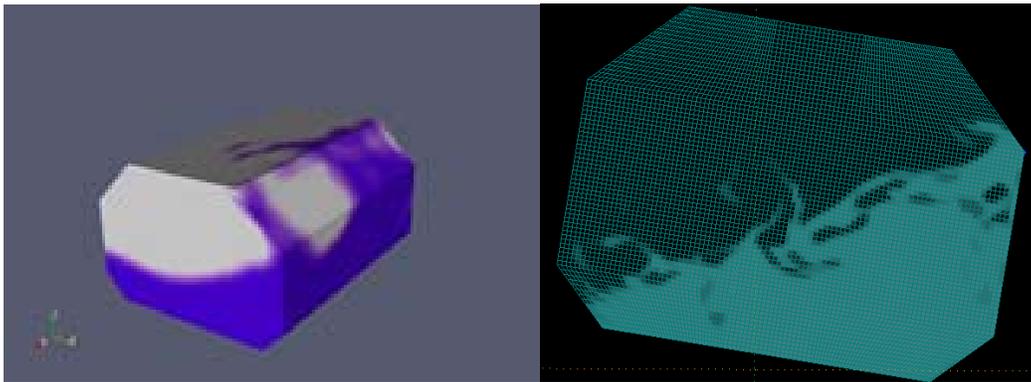
Figure 9. Flow by Aquarius CFD tank model
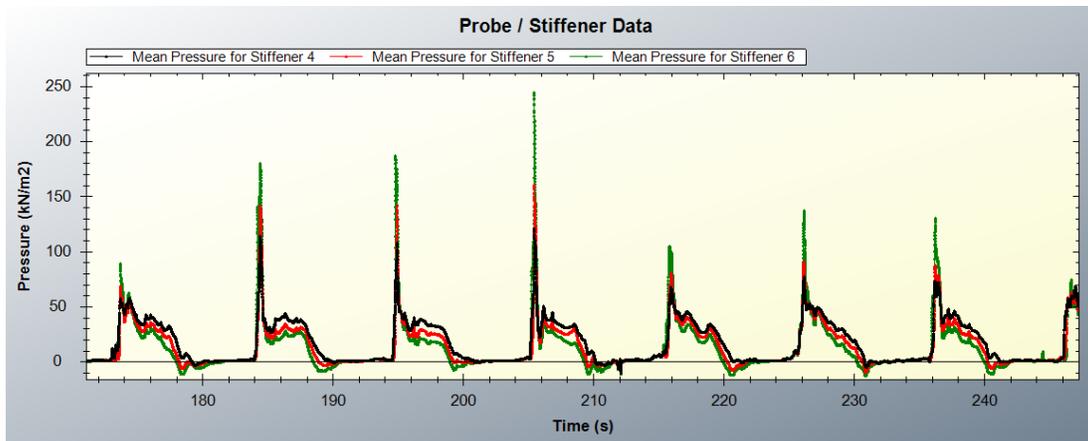
Figure 10. Time history of pressures at specific locations on tank wall by Aquarius

Multi-level Time Domain Approaches for Nonlinear Load Analysis

Frequency domain BEM model has the best efficiency and is used as the main analysis tool in the industry so far. But the linear character of this model has the limitation of its application for small wave and small ship motion. Vertical bending moment and shearing force are used as the typical load in many design and design appraisal works. The critical bending moment and shearing force are estimated from analyses in large wave conditions. Linear BEM model has a basic assumption that there is no vertical geometry change of hull surface around water line, therefore the bending moment in hogging and sagging condition has same amplitude by linear BEM model.   On the other hand, hull geometry does change along vertical at least at bow and stern part. As a result, the amplitude of bending moment in hogging and sagging condition is different in large wave conditions. The time domain model based on database of frequency domain analysis is in general required for critical load assessment. As explained in the stinger force analysis, a reliable time domain simulation in a high sea state needs a large number of seeds, which could be as large as 20 for some cases. The multi-seed time domain analysis may then be too time consuming and becomes unrealistic. A simplified time domain approach, so called intermittent, has been proposed by Lloyd's Register. In an intermittent model, ship motion will be kept the same to those obtained by a frequency domain model, and the frequency domain results of ship motion and pressure are transferred to time domain for a specific regular wave condition or sea state, hydrodynamic pressure on hull surface is then been corrected at each time step. By intermittent correction, non-zero pressure will be added on the mean dry hull surface if it is under water surface at that time according to the ship position and the height of total waves. Correction will also be applied on mean wetted hull surface to make the total pressure being zero if that part of hull moves out of water at the time. The corrected hydrodynamic response will be used to compute the loads, bending moment and shearing force. An analysis in a 24.1 meters wave height regular wave condition is presented in Figure 10 through 12. Pressure distribution by the linear BEM model at a time of hogging condition is plotted on left of Figure 10, the one with intermittent correction is plotted on the right. The linear BEM model has negative total pressure around bow and has zero pressure above the waterline. On the other hand, the intermittent model does not have negative pressure and has non-zero pressure above the water line around the middle ship. The pressure around bow has significant contribution to vertical wave making bending moment (VWBM), while pressure on vertical hull surface will have no contribution to VWBM. As a result, the linear BEM model will over-estimate VWBM due to the negative pressure and its VWBM result would be larger than that by intermittent model.
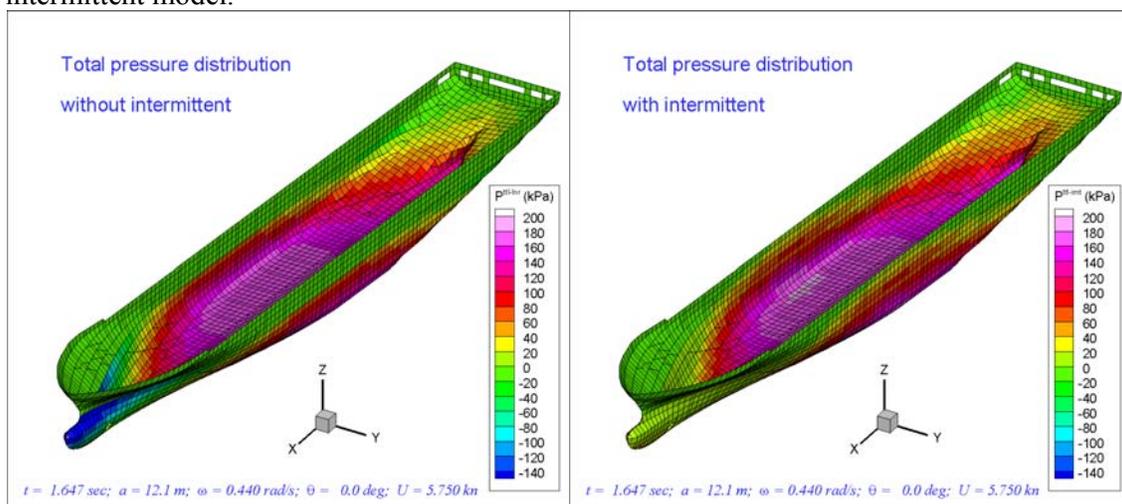
Figure 10. Pressure distribution of linear model and intermittent model in a hogging condition

Figure 11 shows the pressure distribution of the same vessel in same waves but at a time of sagging condition. At this time the linear BEM model has a portion of negative pressure are on vertical hull surface around middle ship area, but the intermittent model got more pressure on bow and stern above the waterline. An increase of VWBM can be expected in the results of intermittent model.
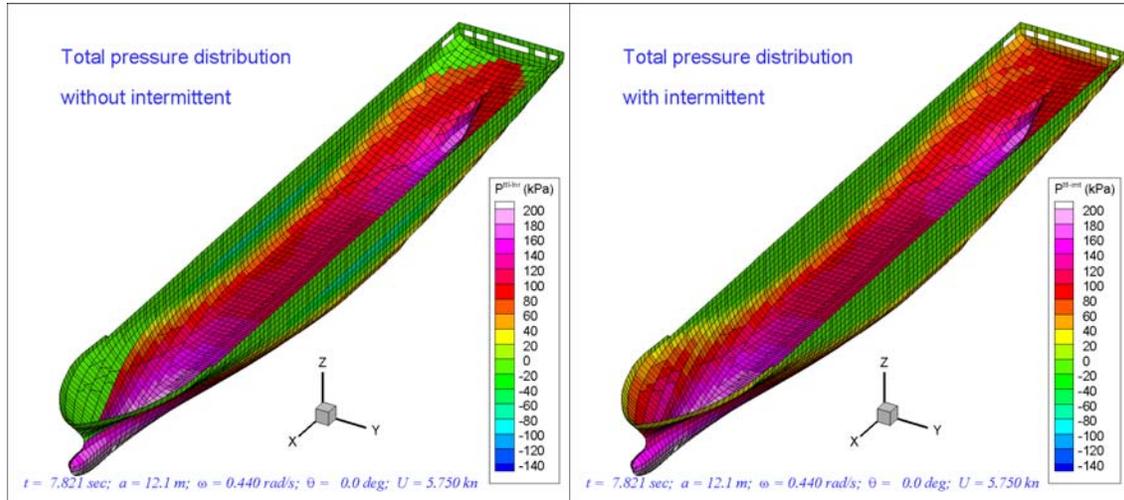
Figure 11. Pressure distribution of linear model and intermittent model in a sagging condition

The longitudinal distribution of VWBM for hogging and sagging conditions obtained by the linear BEM model and nonlinear intermittent model are presented in Figure 12. The results of linear VEM model are in the colour of green, results of intermittent model in colour of black, and the correction part from the intermittent are in red. The hogging results are presented by lines with marks and sagging ones by lines only. The linear BEM model shows a symmetric hogging/sagging result, and results of intermittent model are not. As we expected, intermittent model has a smaller VWBM in hogging, up to 15% on some locations; while for sagging condition, intermittent model got a maximum VWBM around 75% larger than that by the linear model for this extreme high wave
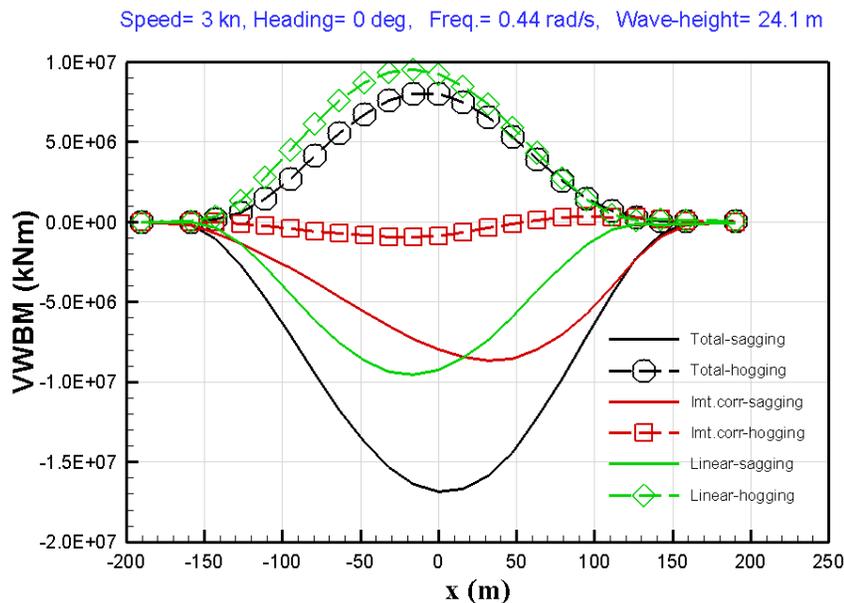
Figure 12. Longitudinal distribution of VWBM in a large wave condition

condition, H=24.1 meters. The VWBM results obtained by intermittent model is closer to the experiments and sea trail data, and could provide a reference load for design and design appraisal. Comparing to time domain BEM model, computation time by an intermittent model is ignorable.

Conclusions

Applications of different approaches based on linear boundary element method and other nonlinear models like CFD, viscous Morrison force and viscous damping, are presented in this work. A simplified nonlinear time domain correction method, intermittent, for design bending moment and shearing force has also been presented. Through these examples, we can see the efforts for improving the efficient numerical tool to answer the requirement from the marine and offshore industry.

References

Lin, F. (2011): Linear and Non-linear Hydrodynamic Analysis for Structural Load of a Pipe-laying Ship, *International Journal of Shipping and Ocean Engineering*, **1, 1-7**, pp 1–7.
Lin, F., Ge, C., Polezhayeva, H.A., and Byers, R. (2009): Coupled Seakeeping Analysis and Sloshing Load for LNG and FPSO Vessels, ISOPE 2009, Osaka, Japan, TPC-252
Faltinsen, O.M. (1990): Sea Loads on Ships and Offshore Structures, Cambridge University Press.
Lloyd, A.R.J.M, and Crossland, P. (1989): Motions of a Steered Model Warship in Oblique Waves, *TRINA* **Vol. 131**, London, UK.
Ikeda, Y., Komatsu, K., Himeno, Y. and Tanaka, N. (1977) On Roll Damping Force of Ship – Effect of Fraction of Hull Surface Pressure Created by Bilge Keels, *Journal of the Kansai Society of Naval Architects*, Japan, No. 165.
Ikeda, Y., Himeno, Y. and Tanaka, N. (1976) On Roll Damping Force of Ship – Effect of Fraction of Hull and Normal Force of Bilge Keels, *Journal of the Kansai Society of Naval Architects*, Japan, No. 161.

Levitron: an exotic toy of nonlinear and linearised dynamics

*Elvio Bonisoli[1] and Cristiana Delprete[2]

[1]Department of Management and Production Engineering,
Politecnico di Torino, Corso Duca degli Abruzzi, 24 - 10129, Torino, Italy
[2]Department of Mechanical and Aerospace Engineering,
Politecnico di Torino, Corso Duca degli Abruzzi, 24 - 10129, Torino, Italy

*Corresponding author: elvio.bonisoli@polito.it

Abstract

The Levitron is a revolutionary toy that continues to astonish beginners and experts of spinning tops. Permanent magnets demonstrate experimentally that can levitate practically without any dissipative effects in the air, but the complexity of the dynamic equations of this famous and exotic toy are relevant. In particular the stability region and the related boundary conditions are surprising and a single model shows difficulties to be consistent for all kind of its dynamics. Starting from a specific experimental test bench and data processing on movies, the paper presents an unique nonlinear magneto-rotordynamic model that allows obtaining the nonlinear equations of motion of all rigid body modes of the Levitron, and with which it is possible to describe the complete dynamic behaviour of the spinning top and to highlight the presence of stability fields related to its spin speed and vertical position of levitation. The advantage of this unique model is also its property to describe and to underline the intrinsic linearised and nonlinear dynamics and the capabilities of this exotic toy to extend the characteristic of a nonlinear system dependent on large displacements and spin speed. By means of the numerical integration of the equations of motion, the spatial trajectories of the spinning top have been computed and validated by comparison with the experimental test results.

Keywords: Levitron, rotordynamics, magnetic levitation, nonlinear dynamics

Introduction

The Levitron is a revolutionary toy that continues to astonish beginners and experts of spinning tops, as the device seemed to violate the famous Earnshaw's theorem of magnetic levitation instability [Earnshaw (1842)]. Although it is well known that the gyroscopic effect stabilizes the rotordynamic behaviour of this magnetic spinning top, consistent analytical models and exhaustive explanations about limit conditions on its stability are not completely presented. Permanent magnets demonstrate experimentally that can levitate practically without any dissipative effects in the air, but the complexity of the dynamic equations of this famous and exotic toy are relevant. In particular the stability region and the related boundary conditions are surprising and a single model shows difficulties to be consistent for all kind of its dynamics. Dynamic properties such as angular speed ranges and geometrical subspace where the stability may be reached, physical masses and magnetic limits, need nonlinear models and modal approaches to identify different behaviours.

In 1996 one of the earliest and most cited paper on the Levitron is published by [Berry (1996)] and it first expressed a theory based on this type of magnetic levitation. The paper defines the vertical stability range for a magnet with a disk-shaped base, the admissible range for the mass of the spinning top, the rotational speed range where it is possible to find a stable behaviour, and the stability conditions in the horizontal plane. Analogously [Simon et al. (1997)] investigates the Levitron dynamic behaviour with a square permanent magnet base and with a circular ring base and demonstrates that both the bases work in a similar way. Also [Jones et al. (1997)] propose a simple

dipole interaction model to investigate the stable behaviour of the Levitron and analyses different assumptions for the orientation of the spinning top for estimating the upper and lower limits of the spin speed for a stable equilibrium, neglecting an unique model.

[Gans et al. (1998)] present a complete, coupled, non-dissipative Hamiltonian system to describe the Levitron dynamic behaviour; the paper points out a region of a two-dimensional manifold of initial conditions for which levitation is permitted and identifies three distinct failure modes that correspond to an insufficient initial spin speed, a too large initial tilt and a too large initial spin speed. [Gov et al (1999)] describe the Levitron problem with a more dynamic approach and defines the stability field along the vertical axis in terms of spin speed. In [Flanders et al. (1999)] the expressions of the minimum speed precession, depending on the transversal and the polar moments of inertia of the spinning top, are reported; the papers try to describe the maximum height reached by the spinning top with complex physical-magnetic approaches.

Finally in [Genta et al. (1999)], by means of a nonlinear rotordynamic model and without introducing any simplification, the equations of motion of all the rigid body modes of the spinning top are obtained. Computing the linerised natural frequencies, thus using the modal approach, that characterize the roto-translational vibrations of the rotor in the plane and the precession motion of its axis, the spin speed conditions to assure the levitation stability are obtained and some results from a numerical integration of the equations of motion are presented. [San Miguel (2005)] arrives at results similar to [Genta et al. (1999)]; it shows the results of three different methods to integrate the equations of motion obtained using a mechanical approach.

[Krechetnikov and Marsden (2006)] discuss the instabilities caused by the non-conservative forces of dissipative and positional type and uses the results of two classical theorems to interpret the Levitron behaviour. It is emphasized that dissipation is fundamental for the stabilization of the spinning top, but in the present authors opinion this point of view may not be completely shared.

All literature papers mentioned so far consider a uniform magnetic induction field or a magnetic induction field linearised at the levitation point, and obtain the rotordynamic equations arbitrarily uncoupling the behaviour of planar and vertical degrees of freedom.

In [Bonisoli et al. (2011)] is presented a way to take into account a nonlinear magnetic model based on the analogy of the equivalent solenoids [Bonisoli and Vigliani (2006)] applied to the Levitron; [Genta et al. (1999)] is revisited presenting a nonlinear and a linearised analyses to describe the physical causes for the existence of two spin speed thresholds of stability; some numerical simulation are performed to underline the limits of the linearised analysis in the coupling between different mode shapes to influence the dynamic behaviour.

In the present paper, the experimental outcomes are initially analysed through a dedicated test bench suitable to provide 3D data for the model comparison. With an unique magneto-rotordynamic model, the analysis of [Bonisoli et al. (2011)] are completely developed taking into account also the air dissipative effects present on the spinning top and the influence on the dynamic behaviour of the device is evaluated and compared with experimental results. The model adopted for both nonlinear and linearised analyses allows evaluating stability conditions, modal linearised dynamic behaviour and nonlinear properties. With respect to the vertical equilibrium position, two different spinning top collapse trajectories can be simulated due to a too slow or a too fast spin speed. The possibility to compute the solutions by using a time-numerical integration for any point in the space and to estimate the magneto-static force intensity by using a nonuniform magnetic model, based on the analogy of the equivalent solenoids, allows calculating the spinning top spatial trajectories. The spectral analysis of the computed time histories are compared with the experimental measurements and the complex nonlinear dynamic nature of the Levitron is discussed. In particular, the transient analyses in time-frequency domains allow verifying the natural frequencies of the linearised model and can demonstrate the nonlinear coupling between modes.

The complete and more detailed comparison between nonlinear and linerised dynamics both numerically and experimentally is under evaluation of the submitted paper [Bonisoli and Delprete (2014)].

Experimental outcomes

The spinning top is constituted by a disk of rare-earth magnetic material (estimated residual magnetization $B_r = 1.13$ T) with vertical magnetization and some paramagnetic/diamagnetic rings that are used to calibrate its overall weight and consequently to adjust its floating height over the base. The disk diameter is 30 mm, the overall height is 33 mm and the spinning top mass is 29.4 grams. The larger magnetic base with vertical magnetization and disposed in opposition with respect to the spinning top is of toroidal form (external diameter $\varnothing_{ext} = 110$ mm, internal hole $\varnothing_{int} = 60$ mm, height $h = 20$ mm) and residual magnetic induction $B_r = 0.254$ T experimentally measured through Hall effect sensor.

The experimental investigation of the spinning top trajectories was done through the analysis of the trajectories by using a dedicated test bench and a video developed with dedicated programs.

Considering an arbitrary Cartesian inertial reference system ($\bar{x}$, $\bar{y}$, $\bar{z}$), with centre in the mean value of the acquired data, the information about the coordinates as a function of time are recorded without synchronizing two digital cameras to the spinning top from two different perspectives, but using a set-up with two mirrors that allow to obtain the three coordinates, in function of time, on the same frame of a unique and economic digital camera Nikon Coolpix 5200. The sketch of the video system is reported in Fig. 1 (left) and one of the frames acquired by the digital video camera is visible in Fig. 1 (right).
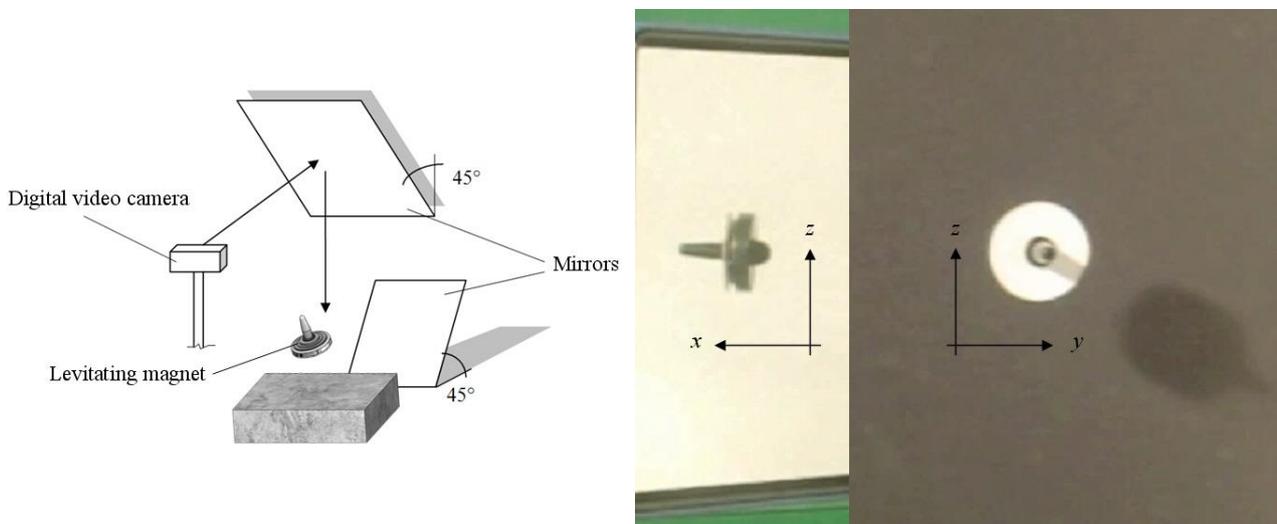
Figure 1. Experimental test bench (left) and frame acquired through the digital video camera (right).

Each frame has a resolution of 320×240 pixel and the movie has an image frequency of 25 Hz. Thus the acquisition setup has low performance, but for the aim of acquisition, it is sufficient as described later.

For demonstrating the amplitude and the coupling of translation and rotational degree of freedom of the spinning top, an example of three frames in perspective are shown in Fig. 2.

In order to highlight the contours of the spinning top and the evaluation of the barycentre displacements, in the top view and in the profile view a visual contrast "black and white", using

high contrast on the movies frames, is adopted to increase the contours definition and to determine in each frame the correct position of the centre of gravity of the spinning top.

An example of the two regions of the spin in the horizontal and vertical plane are shown in Fig. 3. Due to the definition of the barycentre, it results:

$$x_G = \frac{S_{z,front}}{A_{front}} , \quad y_G = \frac{S_{z,top}}{A_{top}} , \quad z_G = \frac{S_{y,top}}{A_{top}} \tag{1}$$

where S are the first moment of area with respect to the reference system and the specific axis and A are the relative areas of the two regions *front* or *top*.
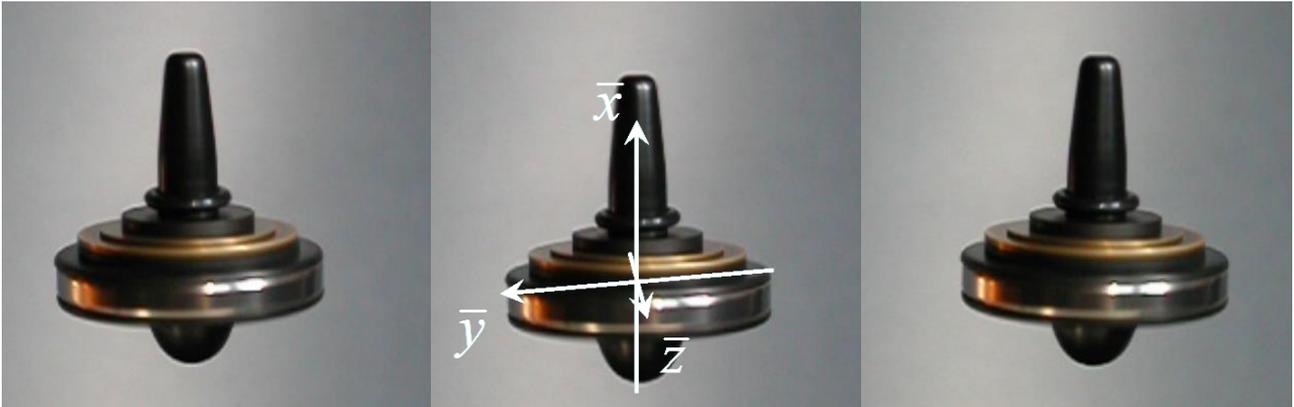
Figure 2. Three frames in perspective, showing large amplitude dynamics of the spinning top.
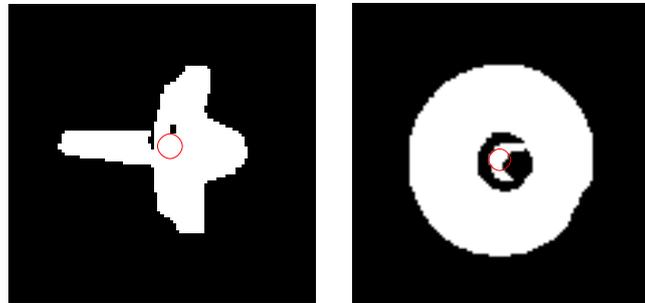
Figure 3. Evaluation of the centre of gravity through high contrast vertical (left) and planar (right) digital process of the same frame.

During the acquisition the minimum and maximum trajectory increment result in the range $0.023 \div 0.646$ mm, with a mean increment of 0.243 mm. According to the frequency rate of 25 Hz of the movie, the measured translational velocities are in the range $0.6 \div 16$ mm/s and the mean value is 6.085 mm/s. It is worth noting that the minimum spatial resolution obtained is about 10 µm, demonstrating the effective strategy to obtain the experimental trajectory of the spinning top.

In Fig. 4 the experimental trajectories are reported and in the plane the detected "flower" shape, composed of "petals", is visible.
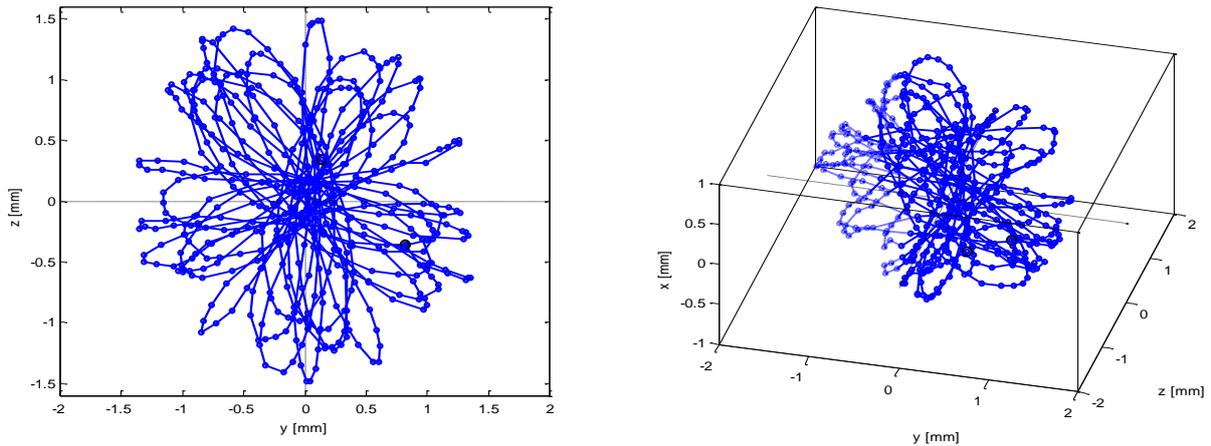
**Figure 4. Experimental trajectories of the spinning top
in the plane *yz* (left) and in three-dimensional trajectory (right).**

Due to the quasi stationary dynamics, it is possible to analysed acquisition data of some minutes. In particular, considering a data movie of 60 seconds, the investigation of the experimental measurements in the frequency domain shows important information on the Levitron dynamics.

It is evident the coupling between various degrees of freedom of the system, which confirms its nonlinear behaviour. A Hamming windowing function was used to calculate the frequency response of the spinning top. Fig. 5 shows the frequency analysis through the power spectrum density (PSD) related to the vertical $\bar{x}$ axis.

With a trial and error approach in Fig. 5 a research of superharmonics and linear combination of the picks is shown. Their names are defined according to the model analysis presented in the next chapter.
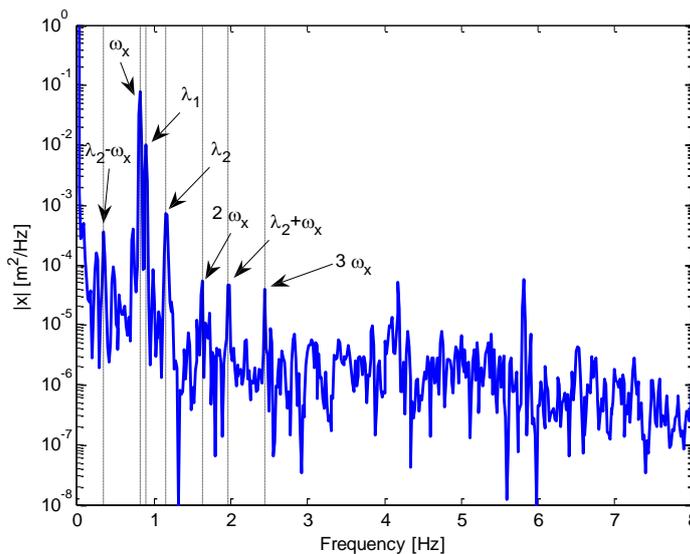
Figure 5. Fourier analyses on experimental vertical $\bar{x}$ data.

It is evident how this exotic toy represents an interesting example of a dynamic multi-degree-of-freedom system where linear and nonlinear effects are evinced; well-known modal approach and

nonlinear techniques can be implemented starting from a mechanical-physical model that is presented in the following chapter.

Numerical model

The numerical model assumes a spinning top with inertial properties according to the experimental device. For simplification in the magnetic model, the magnetic base is assumed to be of a prismatic shape (30×80×80 mm, residual magnetic induction $B_r = 1.13$ T), as the first devices available on the market. Other magnetic bases have been developed with a disk or ring shape, but the magnetic field characteristics are similar and the main parameter for the floating properties is a high ratio between the surface of the magnetic base and the diameter of the spinning top.

According to a mobile reference system coincident to the barycentre of the levitating body (see Fig. 2), to describe the mechanical system under investigation a Cartesian inertial reference system, with centre in the centre of gravity of the magnetic base is used. The magnetic field equations are expressed in the main inertial reference system, while the relations expressed according to the mobile reference system can be traced back to the main one by means of three successive rotations in accordance with the Cardano convention. The translation of the centres of the two reference systems is neglected and, for the used Lagrangian approach, only the axes orientations are necessary and not their relative distance. To obtain the mobile triad from the inertial triad (x, y, z) a first rotation by the angle ψ around the y axis, a second rotation by the angle χ around the new z axis and a final rotation by the angle φ about the last x axis leads to the final mobile reference system.

The rotordynamic behaviour of the spinning top can be described by using the Lagrangian approach presented in [Genta et al. (1999); Bonisoli et al. (2011)] which produces the following system of n differential equations:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial q_i} = Q_i \tag{2}$$

where the generalized coordinate, velocity and force are respectively indicated as q_i, $\dot{q}_i$ and Q_i, and the potential energy is related to the magnetic effects (magnetic induction) through the well-known equations of the magnetic forces and torques:

$$\mathbf{F} = -\nabla U = -\nabla(\boldsymbol{\mathcal{M}} \cdot \mathbf{B}) \quad \text{and} \quad \mathbf{T} = \boldsymbol{\mathcal{M}} \times \mathbf{B} \tag{3}$$

where $\boldsymbol{\mathcal{M}}$, $\mathbf{B}$ are respectively the residual magnetization vector and magnetic induction vector.

They are applied to the spinning top centre of gravity and that the spinning top, assumed point-like, feels the effect of a magnetic field strongly nonuniform and three-dimensional.

If no simplification is made during the mathematical development of the equations of motion (2), the following nonlinear system can be obtained:

$$\begin{cases} m\ddot{x} + c_{tr}\dot{x} + MV\left(\cos(\psi)\cos(\chi)\dfrac{\partial B_x}{\partial x} + \sin(\chi)\dfrac{\partial B_y}{\partial x} - \sin(\psi)\cos(\chi)\dfrac{\partial B_z}{\partial x}\right) = -mg \\[4mm] m\ddot{y} + c_{tr}\dot{y} + MV\left(\cos(\psi)\cos(\chi)\dfrac{\partial B_x}{\partial y} + \sin(\chi)\dfrac{\partial B_y}{\partial y} - \sin(\psi)\cos(\chi)\dfrac{\partial B_z}{\partial y}\right) = 0 \\[4mm] m\ddot{z} + c_{tr}\dot{z} + MV\left(\cos(\psi)\cos(\chi)\dfrac{\partial B_x}{\partial z} + \sin(\chi)\dfrac{\partial B_y}{\partial z} - \sin(\psi)\cos(\chi)\dfrac{\partial B_z}{\partial z}\right) = 0 \end{cases} \tag{4a}$$

$$\begin{cases} I_T \cos(\psi)^2 \ddot{\psi} + I_P \cos(\chi)\dot{\chi}\dot{\phi} + (I_P - 2I_T)\sin(\chi)\cos(\chi)\dot{\psi}\dot{\chi} + c_{rot}\dot{\psi} + \\ - MV\big(\sin(\psi)\cos(\chi)B_x + \cos(\psi)\cos(\chi)B_z\big) = -mgl_g\sin(\psi)\cos(\chi) \\ \\ I_T\ddot{\chi} + (I_T - I_P)\sin(\chi)\cos(\chi)\dot{\psi}^2 - I_P\dot{\psi}\dot{\phi}\cos(\chi) + c_{rot}\dot{\chi} + \\ - MV\big(\cos(\psi)\sin(\chi)B_x - \cos(\chi)B_y - \sin(\psi)\sin(\chi)B_z\big) = -mgl_g\cos(\psi)\sin(\chi) \\ \\ I_P\sin(\chi)\ddot{\psi} + I_P\ddot{\phi} + I_P\cos(\chi)\dot{\psi}\dot{\chi} + c_{rot}\dot{\phi} = 0 \end{cases} \tag{4b}$$

where it is assumed a spinning top mass $m = 29{,}4$ g, moments of inertia $I_P = 2{,}346$ kg mm^2 and $I_T = 1{,}292$ kg mm^2, V is the magnetic volume of the spinning top, M is its magnetization, l_g is the distance between centre of gravity and magnetic volume centre of the spinning top and g is the constant of gravity.

Three different causes of nonlinearity are evinced: order of degrees of freedom (i.e. $\dot{\chi}\dot{\phi}$ term), trigonometric nonlinearities (i.e. $\cos(\psi)$ and other similar terms), distribution of the magnetic induction field (i.e. $\partial B_x/\partial x \neq$ constant).

Dissipative effects, such as the aerodynamic drag torque of the spinning top, is taken into account through the linear damping coefficients c_{tr} and c_{rot} respectively for the translational and rotational behaviour.

The nonlinear magnetic model is based on the analogy of the equivalent solenoids [Bonisoli and Vigliani (2006)] and it allows to map the magnetic induction, its derivatives in the entire spatial domain of the spinning top. Thus it provides the generalised elastic forces. Fig. 6 shows the integration approach on a generic prismatic geometry and the vertical magnetic induction gradient in the plane containing the vertical equilibrium point.
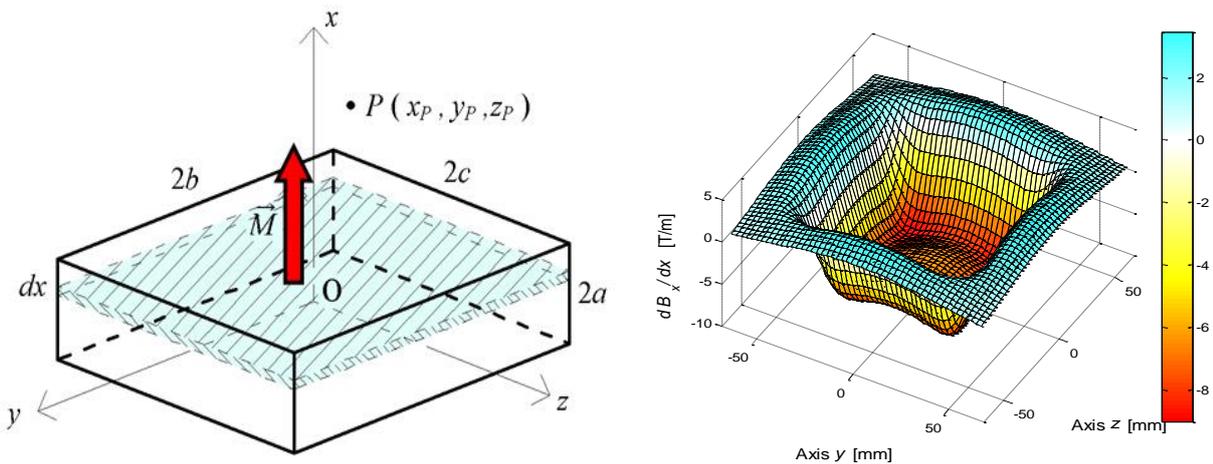
Figure 6. Magnetic model as integrating planar contributions (left) and vertical magnetic field gradient on plane $x_{eq} = 29{,}6$ mm (right).

Using the polynomial formulation of the magnetic induction field, the equations of motion can be linearised with respect to the equilibrium point $x_{eq} = 29{,}6$ mm where the gravity force is equal to the magnetic repulsive force of the magnetic base. The marked degrees of freedom are referred to

the equilibrium position of the spinning top, e.g. $\bar{x} = x - x_{eq}$, it is assumed the spin speed $\dot{\varphi}_0 = \omega$ and the linearised system is expressed as:

$$
\left|
\begin{array}{l}
m\ddot{\bar{x}} + c_{tr}\dot{\bar{x}} + \dfrac{MVB_r}{4\pi} 2A_2 \bar{x} = 0 \\[2ex]
m\ddot{\bar{y}} + c_{tr}\dot{\bar{y}} + \dfrac{MVB_r}{4\pi} 2B_0 \bar{y} + \dfrac{MVB_r}{4\pi} D_0 \bar{\chi} = 0 \\[2ex]
m\ddot{\bar{z}} + c_{tr}\dot{\bar{z}} + \dfrac{MVB_r}{4\pi} 2B_0 \bar{z} - \dfrac{MVB_r}{4\pi} D_0 \bar{\psi} = 0 \\[2ex]
I_T \ddot{\bar{\psi}} + c_{rot}\dot{\bar{\psi}} + I_P \omega \dot{\bar{\chi}} + \left(mgl_g - \dfrac{MVB_r}{4\pi} A_0 \right)\bar{\psi} - \dfrac{MVB_r}{4\pi} D_0 \bar{z} = 0 \\[2ex]
I_T \ddot{\bar{\chi}} + c_{rot}\dot{\bar{\chi}} - I_P \omega \dot{\bar{\psi}} + \left(mgl_g - \dfrac{MVB_r}{4\pi} A_0 \right)\bar{\chi} + \dfrac{MVB_r}{4\pi} D_0 \bar{y} = 0 \\[2ex]
I_p \ddot{\bar{\phi}} + c_{rot}\dot{\bar{\phi}} = 0
\end{array}
\right.
\tag{5}
$$

The first and the last equation of system (5) are uncoupled with respect to the other equations, that represent the vertical behaviour (in x direction) and the planar behaviour of the spinning top, and they can be studied separately. In particular, the planar behaviour can be studied by using the following vector that contains the complex translational and rotational coordinates:

$$
\mathbf{q} = \left\{ \begin{array}{l} r \\ \eta \end{array} \right\} = \left\{ \begin{array}{l} \bar{y}+i\bar{z} \\ \bar{\chi}-i\bar{\psi} \end{array} \right\} = \left\{ \begin{array}{l} r_0\,e^{st} \\ \eta_0\,e^{st} \end{array} \right\}
\tag{6}
$$

where the eigenvalue $s = \sigma \pm i\lambda$ is respectively composed of the decay rate σ and the natural frequency λ.

The planar behaviour of the system can then be expressed in the following matrix form:

$$
\begin{bmatrix} m & 0 \\ 0 & I_T \end{bmatrix}\ddot{\mathbf{q}} + \left(\begin{bmatrix} c_{tr} & 0 \\ 0 & c_{rot} \end{bmatrix} - i\,\omega \begin{bmatrix} 0 & 0 \\ 0 & I_P \end{bmatrix} \right)\dot{\mathbf{q}} + \begin{bmatrix} \dfrac{MVB_r}{4\pi}2B_0 & \dfrac{MVB_r}{4\pi}D_0 \\[2ex] \dfrac{MVB_r}{4\pi}D_0 & mgl_g - \dfrac{MVB_r}{4\pi}A_0 \end{bmatrix}\mathbf{q} = \mathbf{0}
\tag{7}
$$

and it is used for the modal analysis to estimate the stability regions with respect to the spin speed ω.

Terms A_0, A_2, B_0 and D_0 are the Taylor's coefficients used in the polynomial formulation of the magnetic field, considering the corresponding linearised expressions. The components of the magnetic induction field and their corresponding derivatives are respectively approximated as:

$$
B_x \cong \frac{B_r}{4\pi}\left(A_0 + A_1\bar{x}\right),\ B_y \cong \frac{B_r}{4\pi}D_0\bar{y}\,,\ B_z \cong \frac{B_r}{4\pi}D_0\bar{z}
\tag{8a}
$$

$$
\frac{\partial B_x}{\partial x} \cong \frac{B_r}{4\pi}\left(A_1 + 2A_2\bar{x}\right),\ \frac{\partial B_y}{\partial y} \cong \frac{B_r}{4\pi}\left(D_0 + D_1\bar{x}\right),\ \frac{\partial B_x}{\partial y} = \frac{\partial B_y}{\partial x} \cong \frac{B_r}{4\pi}2B_0\bar{y} = \frac{B_r}{4\pi}D_1\bar{y}
\tag{8b}
$$

$$
\frac{\partial B_x}{\partial z} = \frac{\partial B_z}{\partial x} \cong \frac{B_r}{4\pi}2B_0\bar{z} = \frac{B_r}{4\pi}D_1\bar{z}\,,\ \frac{\partial B_z}{\partial z} \cong \frac{B_r}{4\pi}\left(D_0 + D_1\bar{x}\right),\ \frac{\partial B_y}{\partial z} = \frac{\partial B_z}{\partial y} \cong 0
\tag{8c}
$$

Comparison and discussion

The experimental spin speed and stability limits are globally compatible with the model predictions. The first comparison is about the predicted limits of stability of the spinning top.

For the investigated experimental setup and the corresponding numerical configuration, the system results stable in the horizontal plane if the spin speed of the spinning top is between the lower limit value ω_{min} = 96,5 rad/s (921 rpm) and the upper limit value ω_{max} = 239,3 rad/s (2285 rpm).

Fig. 7 reports the computed Campbell's diagram and root locus. Referring to the Campbell's diagram (Fig. 7 left), within the stability range [ω_{min} , ω_{max}] (drawn as two vertical black solid lines) four natural frequencies, corresponding to the four eigenvalues of eq. (7), are distinct, while before and after the stability range only a couple of natural frequencies exists; the reference stationary spin speed of the spinning top, neglecting the dissipative effect, is set at ω = 167 rad/s (1595 rpm) and it is shown as a vertical dashed green line. In the root locus (Fig. 7 right), the eigenvalues s form arcs of circular orbit in the complex plane; the drawn arrows are useful to show the eigenvalues directions when the spin speed increases: when the real part of all the eigenvalues is equal to zero, or less than zero if the dissipative effect is taken into account, a stable behaviour is present.
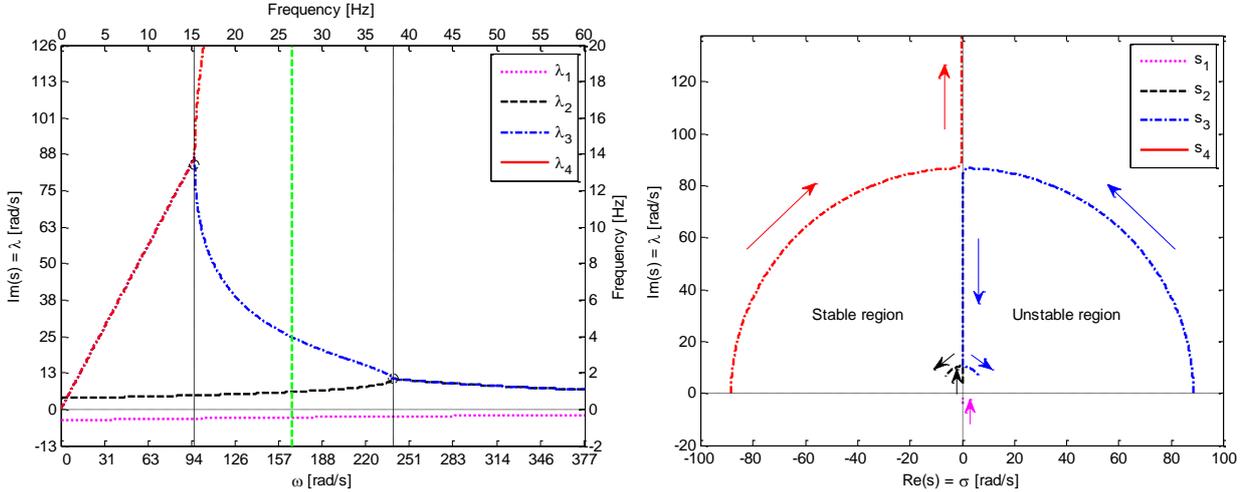
Figure 7. Campbell's diagram versus spin speed ω (left) and root locus (right).

For what the vertical stability is concerned, it is possible to study the first equation of the linearised system (5) separately by the others. For the investigated configuration, this equation represents a magnet-pendulum system with an oscillatory behaviour around the equilibrium height x_{eq} = 29,6 mm and a fundamental harmonic equal to:

$$\omega_x = \sqrt{\frac{MVB_r}{2\pi} A_2} \quad \text{with} \quad \frac{\partial^2 B_x}{\partial x^2} \cong \frac{B_r}{2\pi} A_2 > 0 \tag{9}$$

The sign of the Taylor's coefficient A_2 is linked to the second derivative of the magnetic induction field **B**; according to the static equilibrium of an elastic force, the negative slope of the elastic characteristic $\partial F_x / \partial x < 0$ allows the vertical stability condition.

The vertical stability of the investigated device configuration is then assured when the Taylor's coefficient A_2 is positive: for a vertical position lower than x_{min} = 24,9 mm A_2 is negative and the vertical stability does not exist, for a vertical position higher than x_{min} = 24,9 mm A_2 is positive and the vertical stability exists.

In Fig. 8 the elastic characteristic of the spinning top in shown; the configuration corresponding to the following numerical simulation is represented by the vertical green dashed line where the spinning top weight is balanced by the repulsive magnetic force; in this numerical simulation the ω_x value is about 11,24 rad/s that corresponds to a spinning top vertical frequency of 1,79 Hz. The vertical frequency of the spinning top is very sensitive to the weight; in particular increasing the weight of the spinning top, the frequency decreases to zero where the x_{min} condition in reached.
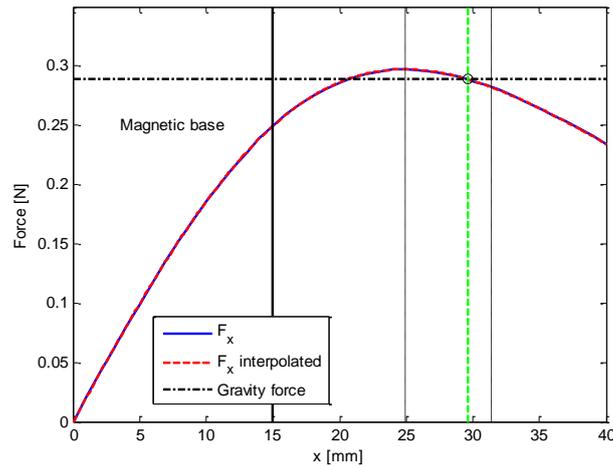
Figure 8. Vertical equilibrium of the spinning top.

The vertical stability has also an upper limit x_{max} for which only qualitative expressions are proposed in Literature. In the present paper, the upper limit x_{max} is also related to the eigenvalue analysis: it is determined by analysing the stiffness matrix reported in equation (7), of the linearised system.

If the inertial, viscous and gyroscopic effects are neglected, the static stability of this equivalent system, corresponding to the linearised system (7), corresponds to positive eigenvalues of the stiffness matrix. By analysing the sign of the two real eigenvalues, the static stability of the system can be investigated as reported in Fig. 9 where the eigenvalues trend is sketched in function of the spinning top levitation height: the first eigenvalue s_1 varies remaining always negative (corresponding to an unstable translational mode that is stabilized through the gyroscopic effect), while the second eigenvalue s_2 varies from positive values (corresponding to a stable precession mode) to negative values (corresponding to another unstable mode).

In particular the first static mode, related to s_1 and unstable, has discordant complex displacement and rotation. When the spinning top has a dynamic behaviour of this form, for positive displacement, i.e. increasing the distance with respect to the x axis, it has a rotation opposed to the magnetic flow field, thus, using the gyroscopic effect to reduce this rotation, a restoring torque is acting on the spinning top, it is moved towards the x axis and the global stability is achieved. The second static mode, related to s_2 and stable in the reference configuration, has concordant complex displacement and rotation. It is fundamental to have it stable, because it is aligned to the magnetic flow field inclination when the spinning top is moving from the x axis.

The transition limit of s_2 sign corresponds to the upper limit $x_{max} = 31,4$ mm of vertical stability of the investigated configuration. Both x_{min} and x_{max} limits are shown with vertical dashed black lines in Fig. 8 and Fig. 9, while the reference model is represented by the dashed green lines.
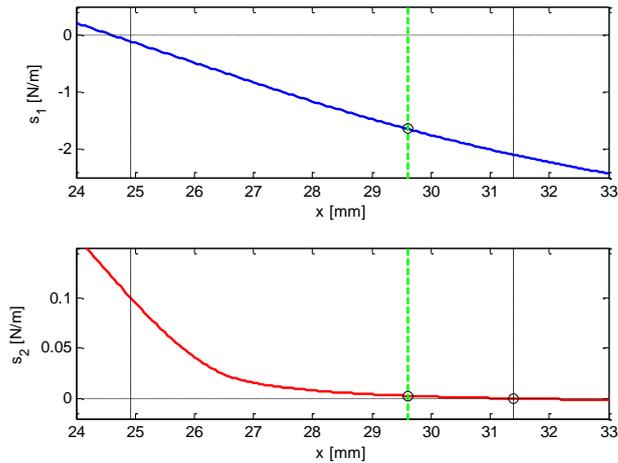
Figure 9. Eigenvalues s_1 and s_2 of the spinning top.

The numerical investigation of the spinning top trajectories was done through the solution of the nonlinear equations of motion (4) with a numerical integration procedure involving a variable step solver (ode45). Starting from the spinning top initial conditions and the chosen simulation time, each numerical simulation provides the trend of the coordinates of the centre of gravity of the levitating magnet; by using a representation on a single three-dimensional diagram the spatial trajectory of the magnet is so obtained.

The trajectory of the spinning top, obtained with initial conditions $x_0 = x_{eq} = 29{,}6$ mm, $\dot{y}_0 = 10$ mm/s and $\dot{\varphi}_0 = \omega = 167$ rad/s and simulation time $T = 10$ s, is reported in Fig. 10 (left); the spinning top can be thought as fluctuating between one extreme and the other of the drawn "flower" shape. The edge effects, with intensity greater than the magnetic force in the middle, increase the vertical component (x direction) of the centre of gravity when the spinning top is far away from the centre of the "flower", which also corresponds to the centre of the base magnet, and it leads to fold up the "petals". In the linearised model case the behaviour in the plane is uncoupled with respect to the vertical axis and, thus, the trajectory is planar and the "petals" are flat.
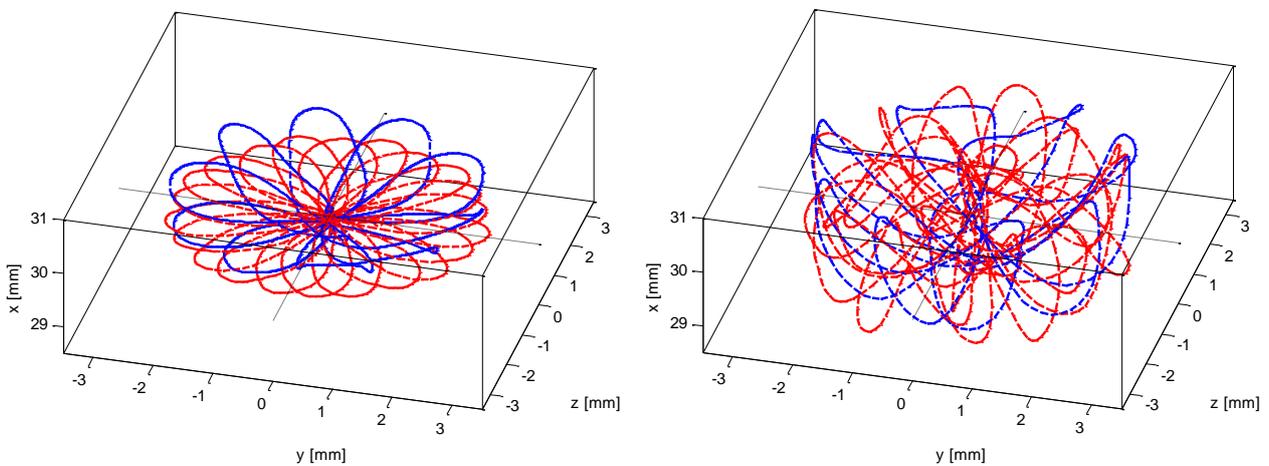
Figure 10. Numerical trajectories of the spinning top: nonlinear model (blue lines) and linearised model (red lines), starting from non-zero initial conditions without vertical perturbation (left) and with vertical perturbation (right); $\dot{\phi}_0 = \omega = 167$ rad/s.

By changing the initial conditions, the trajectory can become more complex than the reported "flower" shape and its domain can take a form other than a "bowl". Fig. 10 (right) shows the trajectory corresponding to the initial vertical condition $x_0 = x_{eq} - 1$ mm $= 28,6$ mm different from the equilibrium one, which presents large oscillations in the vertical direction. The increase in the amplitude of the oscillations is due to the magneto-static force, which generates the levitation, and that is similar to the elastic force of a nonlinear spring [Bonisoli and Vigliani (2007)] and is greater than the case of initial condition $x_0 = x_{eq}$. The planar behaviour of the linearised model is the same of the previous case, due to the vertical and planar uncoupled dynamics.

Considering now the aerodynamic drag torque added to both nonlinear and linearised models, the transient analyses in time-frequency domains allow verifying the natural frequencies of the linearised model and also can demonstrate the nonlinear coupling between modes. In Fig. 11 the progressive decreasing spin speed from $\omega = 167$ rad/s produces the eigenvalues changes in the time domain from the dashed green line to the critical value $\omega_{min} = 96,5$ rad/s of Figure 4 (left) in about 120 s. When natural frequency λ_2 touches λ_3 the behaviour becomes unstable and the spinning top falls. Theoretical natural frequencies ω_x in the vertical x axis, λ_1, λ_2, λ_3 and λ_4 in the plane yz are depicted with dashed white curves. The comparison between the two models shows that in the vertical behaviour secondary frequencies are $(\omega_x \pm \lambda_1)$, the superharmonic $2\omega_x$ and the secondary frequency $(\lambda_3 - \lambda_2)$. In the planar behaviour the secondary frequencies $(\lambda_2 + \omega_x)$, $(\lambda_3 \pm \omega_x)$ and $\left(\lambda_2 \pm \frac{1}{2}\lambda_1\right)$ are detectable. Naturally, in the linearised model only the five eigenvalues survive.
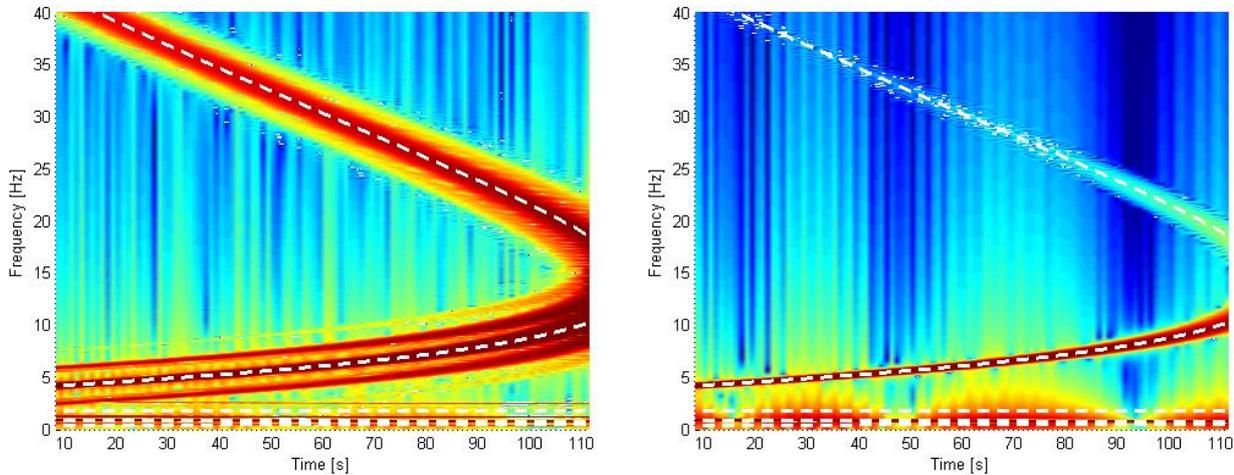
Figure 11. Time-frequency analyses of the spinning top transient behaviour on the vertical axis $\bar{x}$ from $\omega = 167$ rad/s of the nonlinear model (left) and linerised model (right).

With respect to the experimental data, due to the fact that the experimental base magnet is of toroidal shape and it is devoid of the central hole (unlike the numerical simulations), the magnetic field is extended for larger $\bar{x}$ (increasing the Levitron levitating effect) and the natural frequencies present some differences. In particular it has been evinced that ω_x is closer to the backward mode λ_1, also because it is deeply related to the spin weight, according to the note about Fig. 8, and it decreases, increasing its weight (see the tangent behaviour to the force characteristic of Fig. 8).

The presence of the couplings between ω_x and the natural frequencies of the linearised planar analysis is another important characteristic of nonlinearity that demonstrates the coupling between

the vertical and the planar behaviour of the spinning top, experimentally well visible in Fig. 4 and in the three frames of Fig. 2 where the spinning top increases its levitation height when it is far from the vertical axis.

Conclusions

The Levitron continues to surprise beginners and experts of spinning tops. It represents not only a toy, but also a mechanical-physical demonstrator of how nonlinear and linearised dynamics approaches can be able to describe reality. In particular the results obtained from the application of an unique nonlinear magneto-rotordynamic model can relate stability regions of the spin speed, both its vertical limits of levitation and the related boundary conditions through the linearization methodology. The nonlinear model demonstrates how this exotic toy merges the linear modes together and suggests interesting developments in nonlinear normal mode applications and to analyse the planar stability limits and large displacements dynamics, till unknown although the simulations can measure the nonlinear effects of the magnet base boundaries.

Finally, the authors consider the Levitron an interesting example to explain didactically linear and nonlinear dynamic properties.

References

Earnshaw, S. (1842) On the nature of the molecular forces which regulate the constitution of the luminiferous ether, *Trans. Cambridge Philos. Soc.* **7**(116), 97-112.

Berry, M. V. (1996) The LevitronTM: an adiabatic trap for spins, *In: Proceedings Royal Society London* **452**, 1207-1220.

Simon, M. D., Heflinger L. O. and Ridgway S. L. (1997) Spin stabilized magnetic levitation, *Am. J. Phys.* **65**(4), 286-292.

Jones, T. B., Washizu, M. and Gans, R. (1997) Simple theory for the Levitron, *Journal of Applied Physics* **82**(2), 883-888.

Gans, R. F., Jones, T. B. and Washizu, M. (1998) Dynamics of the LevitronTM, *J. Appl. Phys. D* **31**(6), 671-679.

Gov, S., Shtrikman, S. and Thomas, H. (1999) On the dynamical stability of the hovering magnetic top, *Physica D* **126**, 214-224.

Flanders, P., Gov, S., Shtrikman, S. and Thomas, H. (1999) On the spinning motion of the hovering magnetic top, *Physica D* **126**, 225-235.

Genta, G., Delprete, C. and Rondano, D. (1999) Gyroscopic stabilization of passive magnetic levitation, *Meccanica* **34**, 411-424.

San Miguel, A. (2005) Numerical integration for the dynamics of the heavy magnetic top, *Physics A* **335**, 235-244.

Krechetnikov, R. and Marsden, J. E. (2006) On destabilizing effects of two fundamental non-conservative forces, *Physica D* **214**, 25-32.

Bonisoli, E., Delprete, C. and Silvestri. M. (2011) Eigenvalues and nonlinear behaviour of Levitron, *In: Rotating Machinery, Structural Health Monitoring, Shock and Vibration* ISBN: 978-1-4419-9427-1, 245-256.

Bonisoli, E. and Vigliani, A. (2006) Passive effects of rare-earth permanent magnets on flexible conductive structures, *Mechanics Research Communications* **33**(3), 302-319.

Bonisoli, E. and Delprete, C. (2014) Nonlinear and linearised behaviour of the Levitron®, *Meccanica* (submitted).

Bonisoli, E. and Vigliani, A. (2007) Passive elasto-magnetic suspensions: nonlinear models and experimental outcomes, *Mechanics Research Communications* **34**(4), 385-394.

Application of Lattice Green Function & Lattice Boltzmann Model to Lithosphere-asthenosphere permeability

*Bojing Zhu,[1] Dave. A.Yune, [2] Yaolin Shi, [3] Dongsheng Sun,[4] Weiren Lin,[5]

[1] Rock Mechanics and carbon store lab, Department of Earth Sciences, Durham University, DH1 3LE, UK
[2] Department of Earth Science, University of Minnesota, Minneapolis, USA
[3] Key Laboratory of Computational Geodynamics of Chinese Academy of Sciences; College of Earth Science, University of Chinese Academy of Sciences, Beijing, 100049, China
[4] Key laboratory of Neotectonic movement and Geohazard, Ministry of Land and Resources; Geological Mechanics Institute, Chinese Academy of Geological Sciences, Beijing, 100081, China
[5] Kochi Institute for Core Sample Research, Japan Agency for Marine-Earth Science and Technology, Kochi 783-8502, Japan

*Corresponding author: bojing.zhu@durham.ac.uk; bojingzhu@gmail.com

Abstract

This work develops a model to measure the permeability property of the deep lithosphere-asthenosphere [~ 660 Km] under coupled lithostatic pressure [~ 10 GPa] and temperature [~ 1500 ºC]. First, the new novel lattice Green function & lattice Boltzmann multi-scale [200 pm ~ 100 μm ~ 10 cm] flow driven pore-network model is obtained for the first time, and the role of free, supercritical [374 °C & 22.1 MPa] and constitutional state water [H+, (OH)-, (H$_3$O)+, 1500 °C & 1~10 GPa] is analyzed. Then, the permeability property of the deep lithosphere-asthenosphere [~ 660 Km] is predicted and their dependence on pressure, temperature and chemical composition is explored.

Keywords: Lattice Green function; Lattice Boltzmann method; Flow driven pore-network damage theory; Permeability of lithosphere-asthenosphere; Parallel CPU-GPU platform

1. Introduction

Permeability of lithosphere-asthenosphere is closely linked to the rheology, viscous-elastic and strength properties of the deep earth [0 ~ －660 Km], reflected dehydration-melting and phase-transition-melting properties of the rock/mineral and their dependence on pressure, temperature and chemical composition. Permeability physical property is important to understand the interior structure and material convection in the earth [1-4]. Brace et.al [5] studied the permeability of granite under 10~400 MPa condition; Chu et.al [6] found that permeability decrease with an increase in pressure from 1.5 Mpa to 22 MPa, and Morrow et.al [7] measured permeability under 10~200MPa confining pressures; Lockner et.al [8] studied permeability from the 12 Km deep well under confining pressure 10~300 MPa, and pore pressure 112-117 MPa; Darot et.al [9] measured permeability of granite specimens heated up to 650 °C for various confining pressure10~100 MPa; Gorbatsevich et.al [10] investigated permeability of deep-seated rocks under PT-conditions in the Kola Super deep Borehole [SG-3, 6~8 Km]; Gleeson et.al [11] used an extensive compilation of results from hydrogeologic models [>5 Km] and provided the first global picture of near-surface[100 m] permeability; Behnsen and Faulkner [12] measure water permeability of phyllosilicate powders at room temperature and varying confining pressure up to 160 MPa with a constant pore pressure of 10 MPa; Armitage et.al [13] studied vertical and horizontal permeability across a range of effective pressure 0~70 MPa; In-situ measurements of permeability in the Nankai accretionary prism were under taken during a recent IODP expedition [14].

But to the rock/mineral in the deep lithosphere-asthenosphere (~ 660 Km), it is hard to measure permeability through lab experiments manner and classical theoretical analysis way and traditional numerical method, especially when the water is located at supercritical state (374 °C, 22.1 MPa)

and constitutional state (H+, (OH)-, (H$_3$O)+; 1500 °C, 1~10 GPa); In the lab experiment asperity, the existing design capacity and manufacturing ability can not provide such high pressure and temperature condition for the lab experiment apparatus and device; in the classical theoretical analysis asperity, the classical fluid-solid coupled mechanics/physics theory do apply to the ultra high temperature and pressure rock/mineral physical-chemical process in the deep earth; In traditional numerical method asperity, the traditional computational finite different element method and finite element method which based on the continue mechanics/physics can not calculate the lowest scale relationship between the molecular-ion scale rock/mineral structure and potential energy/force; All these reasons lead the permeability in the deep earth (~ − 660 Km) studies remain stagnate stage in the last 40 years.

In recent years, with the development of lowest level scales (atom-molecular-pico-nano-micro) computational porosity fluid flow mechanics/physics theory, especially for quantum lattice Boltzmann method[15-22] applied into physical transformation and chemical equilibrium/reactions problem, measuring/determining ultralow permeability in the deep lithosphere-asthenosphere became possible.

In this work, based on the quantum physical-chemical theory, the lattice Green function under coupled temperature and pressure potential is derived and combined with the previous work on the lattice Boltzmann distribution function [23-27], the new novel lattice Green function & lattice Boltzmann multi-scale flow driven pore-network model for measuring the permeability of rock/mineral in the deep earth is developed for the first time by the authors. Then, the permeability of lithosphere-asthenosphere under coupled LS (0~10 GPa), PS (0~600 MPa) and temperature (0~1500 ºC) is calculated and their dependence on pressure, temperature and chemical composition are analyzed, which can be helpful for understanding the strength weakening of the asthenosphere and the interaction between lithosphere-asthenosphere.

2. The mathematical model

As we know, when the fluid flow translates through ultralow permeability porous rock/mineral (UPPR/M) under ultrahigh pressure-temperature (UHPT), the dynamics viscous of fluid, effects of boundary layer and surface diffusion can not be neglected; the classical fluid-solid theory established on the continue mechanics do not suit for these complex rock/mineral physical-chemical process, and the molecular-ion scale potential energy/force can not be simulated through traditional finite different element and finite element method, the lattice green function & lattice Boltzmann method had be used to analyzed.

Put the typical nature core samples from upper-crust (0~20 Km), lower-crust (20~50 Km), upper mantle-lithosphere (50~150 Km) and lithosphere and asthenosphere (150~660 Km) into the parallel central processing unit & graphics processing unit (CPU&GPU) high performance computational system is the first step for the lattice Green function & lattice Boltzmann (LGF-LBM) model; Using micro XCT-400 CT system and focused ion beam system, the high resolution crosses-section images and micro surface images are obtained, then the 3D virtual digital rock are reestablished, and the LGF-LBM model are constructed. The left row (A1~E1), the medium row (A2~E2) and the right row (A3~E3) in the fig.1 show the initial CT data, the pretreated CT data (the threshold is arrange from 78~100) and the ES&NS model, respectively. In this process, the interior geometry architecture on low-level scales is translated from the pix RGB color format into the cell-mesh-grid model (cell-lattice green function method; grid-lattice Boltzmann method; mesh-LGF-LBM method).

Fig.1. LGF-LBM model for typical rocks/mineral of the deep earth.

A1~E1: represent center slice of physical object of granite. marble, sand, eclogite and olivine, respectively. A2~E2: represent relatively virtual digital models boundary of A1~E1. A3~E3: represent cross section mesh grid (1024×1024) for A1~E1.

Then, the initial/boundary conditions, the physical/chemical parameters of rock/mineral and water, and the low level scales geometry pore-solid structure are defined. The diameter and the length of the nature core samples are equal to 5cm and 10cm, respectively; the resolution of cross-section and the interval between cross-sections are equal to 2~10 µm. The general model was formulated by utilizing an automatic local amplification grid technique, and the initial ES&ES physical is equal to 1024×1024×1024 pixels, which does not include the extended variables (Fig. 2).

Fig.2. The physical objects and corresponding digital virtual physical 3D models

A~E: physical core samples of granite, marble, sandstone, eclogite and olivine, respectively; F~J: relatively digital

virtual physical 3D models of A~E, respectively.

There are four different scales for the UPPR/M in the deep earth. First, the pore/void size is located at atom-molecular-pico-nano scale level, the fluid viscous (as function of pressure-temperature [P-T]), the effects of boundary layer, and the unsteady fluid flow (eddy flow and turbulent flow) can not be neglected; Second, when the P-T conditions are high enough, the water role in the rock/mineral includes free-supercritical-constitutional state, the fluid flow particles are composed of four components [H_2O, H+, (OH)-, (H_3O)+]. Third, the micro pore is composed of four types [multi-grain gap, polycrystalline space, crystal space and crystal internal space[28] and the deformation of the micro-structure had to be considered; Last, the fluid flow permeability and diffusion include intermolecular collisions and diffusion (Fick's laws of diffusion), molecular collisions with interface (Knudsen diffusion), molecular and interfacial adhesive and viscous flow (Darcy and Forchheimer flow). With the scale decrease, the surface stress component became domain, the effect of the body stress component reducing, and the classical N-S equation is no longer applies (Fig. 3).

Fig.3. Micro fluid flow translate in the ultralow permeability under UHPT

Olivine, resolution 2 µm, cell-node-grid scale 1024^3, 1.4 GPa and 600 °C.

The UPPR/M consists of six constituents, i.e. the particles of solid skeleton (the 1st component), bound liquid film (the 2nd component), static pore-liquid (the 3rd component, free state water), and dynamic pore-liquid I (the 4th component, free state water), dynamic pore-liquid II (the 5th component, supercritical state water) and dynamic pore-liquid III (the 6th component, constitutional state water).

Here summation from 1 to 3(1 to 6) over repeated lowercase (uppercase) subscripts is assumed, and a subscript comma denotes the partial differentiation with respect to the coordinates. The dynamic governing equations and constitutive relations of UPPR/M under electro-magneto-thermo-force (EMTE) field can be expressed as [29, 30]

$$A_L \tau_{ij,j}^L + B_L + \Sigma_{iJ,i} + f_J = 0 \qquad (1)$$

where the extended porosity, the extended permeability, the extended stress and the extended body force of the UPPR/M can be defined as

$$
A_L = \begin{cases}
1 - (1-\alpha)\kappa & L = 1 \\
\phi_2^P & L = 2 \\
\phi_3^P & L = 3 \\
\phi_4^P & L = 4 \\
\phi_5^P & L = 5 \\
\phi_6^P & L = 6
\end{cases}
\tag{2}
$$

$$
B_L = \begin{cases}
\left[(1-\kappa)\rho_1 + \alpha\kappa\rho_2\right]\ddot{U}(x_i,t)_{1i} - \sum_{N=1}^{5}\left\{\nu_N\delta_N^2\chi^{-1}\chi_N^{-1}\left[\dot{U}(x_i,t)_{(N+1)i} - \dot{U}(x_i,t)_{Ni}\right]\right\} & L = 1 \\
\phi_2^P\rho_2\ddot{U}(x_i,t)_{2i} - \sum_{N=2}^{5}\left\{\nu_N\delta_N^2\chi^{-1}\chi_N^{-1}\left[\dot{U}(x_i,t)_{(N+1)i} - \dot{U}(x_i,t)_{Ni}\right]\right\} & L = 2 \\
\phi_3^P\rho_3\ddot{U}(x_i,t)_{3i} - \sum_{N=3}^{5}\left\{\nu_N\delta_N^2\chi^{-1}\chi_N^{-1}\left[\dot{U}(x_i,t)_{(N+1)i} - \dot{U}(x_i,t)_{Ni}\right]\right\} & L = 3 \\
\phi_4^P\rho_4\ddot{U}(x_i,t)_{4i} - \sum_{N=4}^{5}\left\{\nu_N\delta_N^2\chi^{-1}\chi_N^{-1}\left[\dot{U}(x_i,t)_{(N+1)i} - \dot{U}(x_i,t)_{Ni}\right]\right\} & L = 4 \\
\phi_5^P\rho_5\ddot{U}(x_i,t)_{5i} - \nu_5\delta_5^2\chi^{-1}\chi_5^{-1}\left[\dot{U}(x_i,t)_{6i} - \dot{U}(x_i,t)_{5i}\right] & L = 5 \\
\phi_6^P\rho_6\ddot{U}(x_i,t)_{6i} + \nu_6\delta_6^2\chi^{-1}\chi_6^{-1}\dot{U}(x_i,t)_{6i} & L = 6
\end{cases}
\tag{3}
$$

$$
\sum_{iJ} = E_{iJKl}Z_{Kl} = \begin{cases}
\sigma_{ij} & J = j = 1,2,3 \\
D_i & J = 4 \\
B_i & J = 5 \\
\vartheta_i & J = 6
\end{cases}
\tag{4}
$$

$$
f_J = \begin{cases}
f_j & J = j = 1,2,3 \\
-f_e & J = 4 \\
-f_m & J = 5 \\
-f_\vartheta & J = 6
\end{cases}
\tag{5}
$$

The more detailed explanation to the equation (4) and (5), please refer to the author's previous work[31, 32]. The drained porous solid frame and the drained bound liquid film have the same displacement and pressure; the static pore-liquid, the dynamic pore-liquid I have the same displacement. The pressure of saturated porous solid frame is the sum of static pore-liquid, dynamic pore-liquid I, dynamic pore-liquid II and dynamic pore-liquid III components pressure. The elastic wave are composed of volume wave, spin wave and deflection wave parts, and can be written as

$$
\frac{\partial^2 \varepsilon_{kk}}{\partial x_m \partial x_m} = \frac{1}{c_L^2}\frac{\partial^2 \varepsilon_{kk}}{\partial^2 t} \quad c_L^2 = \frac{\lambda + 2\mu}{\rho}
\tag{6}
$$

$$
\frac{\partial^2 \omega_k}{\partial x_m \partial x_m} - \frac{\eta}{\mu}\frac{\partial^4 \omega_k}{\partial x_m \partial x_m \partial x_n \partial x_n} = \frac{1}{c_T^2}\frac{\partial^2 \omega_k}{\partial^2 t} \quad c_T^2 = \frac{\mu}{\rho}
\tag{7}
$$

$$
\frac{\partial^2 e_k}{\partial x_m \partial x_m} - \frac{\eta}{\mu}\frac{\partial^4 e_k}{\partial x_m \partial x_m \partial x_n \partial x_n} = \frac{1}{c_T^2}\frac{\partial^2 e_k}{\partial^2 t} \quad c_T^2 = \frac{\mu}{\rho}
\tag{8}
$$

The motion equation of UPPR/M under transient dynamic wave loading can be written as

$$
\left(C_{ijkl}e_{kl} - M_{ij}\left[\phi(U_i - u_i)\right]_{,i}\right)_{ij,j} = \rho\ddot{u}_i + \rho_f\left(\frac{\partial\left[\phi(U_i - u_i)\right]}{\partial^2 t}\right)_{,i}
\tag{9}
$$

$$\left(M\left[\phi(U_i - u_i)\right] - M_{ij}e_{ij}\right)_{,i} = \rho_f \ddot{u}_i + M_{ij}\left(\frac{\partial\left[\phi(U_j - u_j)\right]}{\partial^2 t}\right)_{,j} + r_{ij}\left(\frac{\partial\left[\phi(U_j - u_j)\right]}{\partial t}\right)_{,j} \quad (10)$$

The constitutive relation equation of UPPR/M under transient dynamic wave loading can be written as

$$\left[\left(\lambda + \frac{Q^2}{R}\right)\delta_{ij}u_{i,i} + \mu\left(u_{i,j} + u_{j,i}\right) + Q\delta_{ij}U_{i,i}\right]_{,j} + F_i = -\omega^2\left(\rho_{11}u_i + \rho_{12}U_i\right) - i\omega^2 b\left(u_i - U_i\right) \quad (11)$$

$$\left[Qu_{i,i} + RU_{i,i}\right]_{,i} + F_i' = -\omega^2\left(\rho_{11}u_i + \rho_{22}U_i\right) - i\omega^2 b\left(u_i - U_i\right) \quad (12)$$

Then the wave displacement in UPPR/M can be defined as

$$U(x_i,t)_{iL} = \Phi^1(x_i,t)_{iL,i} + \varepsilon_{ijk}\Phi^2(x_i,t)_{jL,k} \quad (13)$$

where

$$\begin{aligned}&\Phi(x_i,t)_{iL}^{\alpha}\\ &= -A_i^{\alpha}E(\omega t)\left\{\delta_{3i}\left[\begin{matrix}E(-2\omega_L^{\alpha}x_1x_2x_3\sin\theta_{1L}^{\alpha}\cos\theta_{2L}^{\alpha}) + E(-2\omega_L^{\alpha}x_1x_2x_3\cos\theta_{1L}^{\alpha}\sin\theta_{2L}^{\alpha})\\ + E(-2\omega_L^{\alpha}x_1x_2x_3\cos\theta_{1L}^{\alpha}\cos\theta_{2L}^{\alpha}) + E(-2\omega_L^{\alpha}x_1x_2x_3\sin\theta_{1L}^{\alpha}\sin\theta_{2L}^{\alpha})\end{matrix}\right]\right.\\ &\quad\left.+ E(-2\omega_L^{\alpha}x_3\cos\angle(x_3,x_2)_{pL}^{\alpha}) + E(-2\omega_L^{\alpha}x_3\cos\angle(x_3,x_1)_{pL}^{\alpha})\right\}\\ &\times\left[\begin{matrix}E(\omega_L^{\alpha}x_1\sin\angle(x_3,x_2)_{pL}^{\alpha}) + E(\omega_L^{\alpha}x_2\sin\angle(x_3,x_1)_{pL}^{\alpha}) + E(2\omega_L^{\alpha}x_1x_2\sin\theta_{1L}^{\alpha}\cos\theta_{2L}^{\alpha}) + E(2\omega_L^{\alpha}x_1x_2\cos\theta_{1L}^{\alpha}\sin\theta_{2L}^{\alpha})\\ + E(2\omega_L^{\alpha}x_1x_2\cos\theta_{1L}^{\alpha}\cos\theta_{2L}^{\alpha}) + E(2\omega_L^{\alpha}x_1x_2\sin\theta_{1L}^{\alpha}\sin\theta_{2L}^{\alpha})\end{matrix}\right]\end{aligned} \quad (14)$$

$$\theta_{1L}^{\alpha} = \begin{cases}\angle(x_3,x_2)_{pL} + \angle(x_3,x_1)_{pL} & \alpha = 1\\ \angle(x_3,x_2)_{pL} - \angle(x_3,x_1)_{pL} & \alpha = 2\end{cases} \quad (15)$$

$$\theta_{2L}^{\alpha} = \begin{cases}\angle(x_3,x_2)_{sL} + \angle(x_3,x_1)_{sL} & \alpha = 1\\ \angle(x_3,x_2)_{sL} - \angle(x_3,x_1)_{sL} & \alpha = 2\end{cases} \quad (16)$$

$$\omega_L^{\alpha} = \begin{cases}\omega / v_P^L & \alpha = 1\\ \omega / v_S^L & \alpha = 2\end{cases}, \quad E() = e^{i()} \quad (17)$$

If we defined that solid skeleton component and liquid components are parallel and subjected to the same strain, the time-dependent rigidity modulus components G_L of UPPR/M is defined as

$$G_L = (1-f)G_0 + \alpha f(v_s^L \varepsilon_L^{-2})_{,t} \quad (18)$$

The permeability tensor can be defined as

$$\kappa_{ij} = -\frac{V_i \mu}{g\phi_{,j}} = -\frac{V_i \mu}{gJ_i} \quad (19)$$

The more detailed introduction can be found elsewhere in the literature[33-36].

The lattices point in $\Re^d$ can be defined as

$$\left\{ \sum_{i=1}^{d} n_i a_i : n_1,......,n_d \in \mathbb{Z} \right\} \subseteq \mathfrak{R}^d \tag{20}$$

the lattice Green function on the lattices point can be written as

$$p_n(x:z) = \sum_{n=0}^{\infty} p_n(x) z^n = \frac{1}{\pi^d} \int_x^{\pi}......\int_x^{\pi} \frac{d\kappa_1......d\kappa_n}{1 - z\lambda(\kappa)} \tag{21}$$

$$\lambda(\kappa) = \sum_{x \in \mathfrak{R}^d} p_i(x) e^{ix \bullet \kappa} \tag{22}$$

When the resolution between any two grids in the D3Q27 [31] numerical model is high enough to reach the quantum physical-chemical scale level (atom-molecular-ion level, around 200 pm), the lattice grid in the D3Q27 can be defined as lattice point $\mathfrak{R}^d$. The relationship between distribution function [31] in the D3Q27 and the structure function in the lattice Green function can be established, and the improved D3Q27 lattice Green function & lattice Boltzmann can be used to exploring the lattice points extended physical-chemical process, and the permeability in the UPPR/M can be calculated.

3. LGF-LBM model for lithosphere-asthenosphere permeability

3.1 Mathematical calculations versus AUTOLAB 200C experience system

Based on the LGF-LBM model, the ultralow permeability of the deep earth under UHPT can be measured. We choose a tight sandstone core samples from the Ordos Basin Triassic formation of China, the rock depth, the diameter, the length, the density, the confining pressure, the PS, the temperature is defined as 862.76~864.36 m, 25.4 mm, 25~30 mm, 2.359~2.426 g/cm³, 0~200 MPa, 0~10 MPa, 35~45 °C respectively, the more detail parameters of the tight sandstone core sample are shown in Tab.1.

Tab.1. Mineral composition, porosity and matrix density of the tight sandstone samples

The diagonal component of permeability tensor obtained as LGF-LBM model and experimental approach are given in the Tab. 2~4. The regular black color values represent the LGF-LBM model results, the value with outside border represent experimental results. From the diagonal component of permeability in x, y (DC-XYZ), we can see that the LGF-LBM model approach results are consisted with fluid pulse experimental results on the Inc AUTOLAB 200C system.

Tab.2. Diagonal component in x-direction as function of effective pressure and temperature

Tab.3. Diagonal component in y-direction as function of effective pressure and temperature

Tab.4. Diagonal component in z-direction as function of effective pressure and temperature

3.2 Upper-crust (0~20 Km)

The P-T conditions are 0~400 °C and 0~1.4 GPa, the water is free-supercritical state, the fluid transport in the multi-grain gap and polycrystalline space, the dehydration-creep-consolidation-enhance process existed, the permeability increases with LS, PS and temperature increasing (LS is most important, PS is second and temperature is least important); the permeability of sandstone and marble are located at 10~100 uD and 0.1~2 uD respectively, and the PA of marble is more apparent than sandstone (Fig.4).

A~C: represnet diagonal element of sandstone in x, y and z direction, respectively; D~F: represent diagonal element of marble in x, y and z direction, respctively.

Sandstone: The diagonal component in x, y and z direction (DC-XYZ) varies 60~155 uD, 30~98 uD and 10~30 uD respectively; The permeability increases slowly at 0~0.65 GPa, increases sharply at 0.65~1 GPa, increases rate slow down again at the range of 1~1.4 GPa; The temperature has little influence and can be neglected at 0~0.6 GPa, has positive influence but still not obvious at 0.6~1 GPa, became obviously when the effective pressure over 1 GPa.

Marble: The DC-XYZ varies 0.5~2 uD, 0.1~1.8 uD and 0.1~1.3 uD respectively; the effect of LS increases fast in x, y and z direction is 0.7~1.1 GPa, 0.5~1.1 GPa and 1.2~1.4 GPa respectively, the most positive influence effective stage in x, y and z direction is 1.0~1.4 GPa, 1.3~1.4 GPa and 1.2~1.4 GPa respectively.

3.3 Lower-crust (20~50 Km)

The P-T conditions are 400~900 °C and 0~1.4 GPa, the water is supercritical state and constitutional state, the fluid flow transport in the polycrystalline space and crystal space, and the dehydration-phase transient process existed, the permeability increases with LS, PS and temperature increasing (LS is most important, PS is second and temperature is least important). The permeability of marble and granite are located at 0.1~2.0 uD and 0.03~1.6 uD respectively, and the PA of granite is greater than marble (Fig.5).

A~C: represnet diagonal element of marble in x, y and z direction, respectively. D~F: represent diagonal element of granite in x, y and z direction, respctively.

Marble: The DC-XYZ varies 0.67~2.3 uD, 0.12~1.9 uD and 0.015~1.4 uD respectively. The effects of temperature on PA are significant at 400 ~ 900 °C, the PA increased sharply when the temperate is higher than 400 °C, which reflects that the water properties over 400 °C is different from below 400 °C.

Granite: The DC-XYZ varies 0.03~1.6 uD, 0.04~1.0 uD and 0.03~0.48 uD respectively. Compare with PA in upper-crust zone, the permeability fast-increasing region of granite move from 0.65~1 GPa to 1~1.4 GPa, which reflected physical properties difference between upper crust and lower-crust. The effect of temperature to the permeability became highly significant when LS is higher than 1.0 GPa, the PA has marked rise at 550°C point, especially for diagonal element in y and z direction.

3.4 Upper mantle-lithosphere (50~150 Km)

The P-T conditions are 400~900 °C and 1~5 GPa, the water is constitutional state, the fluid flow transport in the crystal space and crystal internal clearance. The permeability increases with LS, PS and temperature increasing (temperature is most important, LS is second and PS is least important). The permeability of granite and olivine are 0.27~4.1 uD and 0.04~0.39 uD, respectively (Fig.6).

A~C: represnet diagonal element of granite in x, y and z direction, respectively. D~F: represent diagonal element of olivine in x, y and z direction, respctively.

Granite: The DC-XYZ varies 0.55~4.0 uD, 0.27~1.57 uD and 0.27~1.18 uD respectively. The effect of temperature on the permeability anisotropy (PA) is significant, and increased sharply when the temperate is higher than 400 °C.

Olivine: When temperate is higher than 400 °C varies 0.08~0.38 uD, 0.06~0.24 uD and 0.04~0.26 uD, respectively. There are existed two PA properties dramatic changes region at 400 °C and 600 °C point respectively.

3.5 Lithosphere and asthenosphere (150~660 Km)

The P-T conditions are 900~1500 °C and 5~10 GPa, the water is constitutional state, the fluid flow transport in the crystal space and crystal internal clearance. The permeability decreases with LS and PS increasing, and its increase with temperature increasing, temperature is most important, LS is second and PS is least important (Fig.7).

Fig.7. Permeability of olivine and eclogite as function of lithostatic pressure (5 ~ 10 GPa), pore pressure (400 ~800 MPa) and temperature (900 ~ 1200 °C)
A~C: represnet diagonal element of olivine in x, y and z direction, respectively. D~F: represent diagonal element of eclogite in x, y and z direction, respctively.

Olivine: The DC-XYZ varies 0.022~0.078 uD, 0.017~0.068 uD and 0.011~0.051 uD respectively. There are three-temperature inflection points (1st point is 400 °C, 2nd point is 550~600 °C, 3rd point is 900 °C), the PA increase sharply at the first two points.

Eclogite: The DC-XYZ varies 0.0032~0.048 uD, 0.0029~0.062 uD and 0.0023~0.053 uD respectively. The temperature effects manner on olivine PA is similar on eclogite.

5. Discussion and conclusion

In this paper, we explored the water (free water, supercritical water and constitutional water) transport and dehydration in UPPR/M of the deep earth under UHPT, and analyzed the correlation between permeability under UHPT. Combined with high resolution 3D virtual digital technology, the multi temporal-spatial scale flow driven pore-network dislocation-crack damage theory, the novel new LGF-LBM model for measuring permeability of the deep earth is developed for the first time.

Compared with permeability results of the tight sandstone on the Inc AUTOLAB 200C system under coupled LS (0~200 Mpa), PS (0~60 MPa) and temperature (0~180 ºC), the accuracy, reliability and advantage of LGF-LBM model is validated.

The permeability of the deep earth (0~660 Km) under coupled LS (0~10 GPa) and temperature (0~1500 ºC) is measured (Tab.5), and its variation as function of pressure, temperature and chemical composition is discussed (Fig.8). We provide a new approach to understanding the mechanism of rheology and strength variation of asthenosphere and interaction between lithosphere and asthenosphere.

Tab.5. Permeability anisotropy of the deep earth under UHPT from LGF-LBM model

Fig.8. Permeability distribution as LS, PS and temperature of the deep earth from LGF-LBM model
1A~1C: DC-XYZ of sandston as function of LS (0 ~ 1.4 GPa), PS (0 ~ 200 MPa) and temperature (0 ~ 400 °C), respectively; 1D~1F: DC-XYZ of marble as function of LS (0 ~ 1.4 GPa), PS (0 ~ 200 MPa) and temperature (0 ~ 400 °C), respectively.
2A~2C: DC-XYZ of marble as function of LS (0 ~ 1.4 GPa), PS (0 ~ 200 MPa) and temperature (400 ~ 900 °C), respectively; 2D~2F: DC-XYZ of granite as function of LS (0 ~ 1.4 GPa), PS (0 ~ 200 MPa) and temperature (400 ~ 900 °C), respectively.
3A~3C: DC-XYZ of granite as function of LS (1 ~ 5 GPa), PS (200 ~500 MPa) and temperature (400 ~ 900°C), respectively; 3D~3F: DC-XYZ of olivine as function of LS (1 ~ 5 GPa), PS (200 ~500 MPa) and temperature (400 ~ 900°C), respectively.
4A~4C: DC-XYZ of olivine as function of LS (5 ~ 10 GPa), PS (400 ~ 800 MPa) and temperature (900 ~ 1500°C), respectively; 4D~4F: DC-XYZ of eclogite as function of LS (5 ~ 10 GPa), PS (400 ~800 MPa) and temperature (900 ~ 1500°C), respectively.

Acknowledge

Results described in this paper are obtained on the Deepcomp7000 of Supercomputing Center, Computer Network Information Center of Chinese Academy of Sciences, and Minneapolis Supercomputing Institute (MSI) of University of Minnesota, respectively.
Project SinoProbe-07 of China and the National Natural Science Foundation of China (D0408/4097409) supported this work.
NOTE:
Due to the limit of conference international space, all figures and tables are shared through following link:

https://www.dropbox.com/s/b3aji8vvh86wywb/Full_paper_Bojing%20Zhu.pdf

References

1. D.Connolly, J.A., et al., *Permeability of asthenospheric mantle and melt extraction rates at mid-ocean ridges.* Nature, 2009. **462**: p. 209-212.
2. Crone, T.J., M. Tolstoy, and D.F. Stroup, *Permeability structure of young ocean crust from poroelastically triggered earthquakes.* Geophysical Research Letters. **38**.
3. Hebert, L.B. and L.G.J. Montesi, *Generation of permeability barriers during melt extraction at mid-ocean ridges.* Geochemistry Geophysics Geosystems. **11**.
4. Rapp, R.P., *AMPHIBOLE-OUT PHASE-BOUNDARY IN PARTIALLY MELTED METABASALT, ITS CONTROL OVER LIQUID FRACTION AND COMPOSITION, AND SOURCE PERMEABILITY.* Journal of Geophysical Research-Solid Earth, 1995. **100**(B8): p. 15601-15610.
5. Brace, W.F., J.B. Walsh, and W.T. Frangos, *Permeability of Granite under High Pressure.* Journal of Geophysical Research, 1968. **73**(6): p. 2225-&.
6. C.L.Chu, C.Y.Wang, and W.Lin, *Permeability and frictional properties of San Andreas Fault Gouges.* Geophysical Research Letters, 1981. **8**(6): p. 565~568.
7. Morrow, C., L.Q. Shi, and J. Byerlee, *Permeability and Strength of San-Andreas Fault Gouge under High-Pressure.* Geophysical Research Letters, 1981. **8**(4): p. 325-328.
8. Lockner, D., et al., *Laboratory-Determined Permeability of Cores from the Kola Superdeep Well, Ussr.* Geophysical Research Letters, 1991. **18**(5): p. 881-884.
9. Darot, M., Y. Gueguen, and M.L. Baratin, *Permeability of Thermally Cracked Granite.* Geophysical Research Letters, 1992. **19**(9): p. 869-872.
10. GORBATSEVICH, F.F., S.V. IKORSKY, and A.V. ZHARIKOV, *STRUCTURE AND PERMEABILITY OF DEEP-SEATED ROCKS IN THE KOLA SUPERDEEP BOREHOLE SECTION (SG-3)* Acta Geodynamica et Geomaterialia, 2010. **7**(2(158)): p. 145~152.
11. Gleeson, T., et al., *Mapping permeability over the surface of the Earth.* Geophysical Research Letters, 2011. **38**.
12. Behnsen, J. and D.R. Faulkner, *Water and argon permeability of phyllosilicate powders under medium to high pressure.* Journal of Geophysical Research-Solid Earth, 2011. **116**.
13. Armitage, P.J., et al., *Experimental measurement of, and controls on, permeability and permeability anisotropy of caprocks from the CO2 storage project at the Krechba Field, Algeria.* Journal of Geophysical Research-Solid Earth, 2011. **116**.
14. Boutt, D.F., et al., *Scale dependence of in-situ permeability measurements in the Nankai accretionary prism: The role of fractures.* Geophysical Research Letters, 2012. **39**.
15. Keehm, Y., T. Mukerji, and A. Nur, *Permeability prediction from thin sections: 3D reconstruction and Lattice-Boltzmann flow simulation.* Geophysical Research Letters, 2004. **31**(4).
16. Bosl, W.J., J. Dvorkin, and A. Nur, *A study of porosity and permeability using a lattice Boltzmann simulation.* Geophysical Research Letters, 1998. **25**(9): p. 1475-1478.
17. Sun, W.C., et al., *Connecting microstructural attributes and permeability from 3D tomographic images of in situ shear-enhanced compaction bands using multiscale computations.* Geophysical Research Letters, 2011. **38**.
18. Renard, F., et al., *Synchrotron 3D microtomography of halite aggregates during experimental pressure solution creep and evolution of the permeability.* Geophysical Research Letters, 2004. **31**(7).
19. Bai, L.P., D.R. Baker, and R.J. Hill, *Permeability of vesicular Stromboli basaltic glass: Lattice Boltzmann simulations and laboratory measurements.* Journal of Geophysical Research-Solid Earth, 2010. **115**.
20. Y.H.Liu, *A lattice Boltzmann model for blodd flows.* Applied Mathematical Modelling, 2012. **36**(7): p. 2890-2899.

21. Leclaire, S.b., M. Reggio, and J.-Y. Trépanier, *Numerical evaluation of two recoloing operators for an immiscible two-phase flow lattice Boltzmann model.* Applied Mathematical Modelling, 2012. **36**: p. 2237-2252.

22. B.Yan, j.Y.Zhang, and G.W.Yan, *A lattice Boltzmann model for the eddy-stream equations in two-dimensional incompressible flows.* Applied Mathematical Modelling, 2011. **35**: p. 2358-2365.

23. Zhu, B.J. and Y.L. Shi, *Three-dimensional flow driven pore-crack networks in porous composites: Boltzmann Lattice method and hybrid hypersingular integrals.* Theoretical and Applied Fracture Mechanics, 2010. **53**(1): p. 9-41.

24. Zhu, B.J., et al., *Analysis of 3D fluid driven crack propagation problem in co-seismic slip under P- and S-waves by hybrid hypersingular integral method.* Computer Methods in Applied Mechanics and Engineering, 2009. **198**(30-32): p. 2446-2469.

25. H.H.Cheng, et al., *Extended hybrid pressure and velocity boundary conditions for D3Q27 lattice Boltzmann model.* Applied Mathematical Modelling, 2012. **36**: p. 2031-2055.

26. B.J.Zhu and Y.L.Shi, *Study of tight sandstone permeability from lattice boltzmann & digital rock model.* Chines Journal of Theoretical and Applied Mechancis, 2013. **45**(3): p. 384-394.

27. Bo-Jing, Z., et al., *Application of digital rock technology to saturated rock friction mechanism under high pressure and temperature.* process in Geophysics, 2013. **28**(4): p. 100~111.

28. Poiriet, J.-P., *Creep of crystals: High-temperature deformation processes in metals, ceramics and minerals.* Cambridge university press, 1985.

29. Aboudi, J., *Micromechanical analysis of fully coupled electro-magneto-thermo-elastic multiphase composites.* Smart Materials & Structures, 2001. **10**(5): p. 867-877.

30. Perez-Aparicio, J.L. and H. Sosa, *A continuum three-dimensional, fully coupled, dynamic, non-linear finite element formulation for magnetostrictive materials.* Smart Materials & Structures, 2004. **13**(3): p. 493-502.

31. B.J.Zhu and T.Y.Qin, *3D modeling of crack growth in electro-magneto-thermo-elastic coupled viscoplastic multiphase composites.* Applied Mathematical Modelling, 2009. **33**: p. 1014-1041.

32. B.J.Zhu. and T.Y.Qin, *Application of hypersingular integral equation method to three-dimensional crack in electromagnetothermoelastic multiphase composites.* International Journal of Solids and Structures, 2007. 44(18-19), 5994-6012.

33. B.J.Zhu, et al., *Porosity and permeability evolution and evaluation in anisotropic porosity multiscale-multiphase-multicomponent structure.* Chinese Science Bulletin, 2012. **57**(4): p. 320~327.

34. Dmitriev, N.M., *Surface porosity and permeability of porous media with a periodic microstructure.* fluid Dynamics, 1995. **30**(1): p. 64~69.

35. C.Y, K., F. J.D, and C. J.L.A, *Image analysis determination of stereology based fabric tensors.* Geotechnique, 1998 **48**: p. 515~525.

36. Koo, J. and C. Kleinstreuer, *Liquid flow in microchannels: experimental observations and computational analyses of microfluidics effects.* JOURNAL OF MICROMECHANICS AND MICROENGINEERING, 2003(13): p. 568~579.

Finite Element Analysis of sustainable and deconstructable semi-rigid beam-to-column composite joints

*Abdolreza Ataei[1], Mark A. Bradford[2]

[1,2]Centre for Infrastructure Engineering & Safety, School of Civil & Environmental Engineering, UNSW Australia, Sydney, NSW 2052, Australia

Corresponding author: a.ataei@unsw.edu.au

In this paper, an innovative sustainable semi-rigid beam-to-column composite joint with deconstructable bolted shear connectors is modelled by using the general purpose software ABAQUS. The structural mechanics of this joint considered in the paper requires careful consideration, in order to capture the response accurately using computational techniques as the interactions of the various components is complex. Three laboratory test specimens having sustainable and deconstructable semi-rigid beam-to-column joints have been tested and the results are used for validation of the finite element model. Precast "green concrete" (GC) slabs having reduced CO_2 emissions during their manufacture were attached compositely to the steel beam via pre-tensioned bolted shear connectors, with the composite beam being connected to H-section columns using a flush end plate with two rows of bolts. The experimental testing was full-sized, with all the components being of the same size as would be met in practice. The numerical model simulates the composite beam-to-column connection under hogging moment and includes the non-linear material properties of all constitutive materials of the composite joint. For validation of the computational procedure, the results of the numerical modelling are compared in the paper with the experimental results, with good agreement being demonstrated.

Keywords: FE modelling, Semi-rigid composite joint, deconstructablility, Precast, Connections

Introduction

The traditional flush end plate semi-rigid composite connection is one of the best choices for connecting a composite beam to a column. This kind of connection has several advantages such as its ease of construction as well as being economical compared to a rigid connection. Apart from these benefits, the rigidity in this connection can allow for adequate moment distribution in the frames. These composite connections have higher initial stiffness and moment capacity as well as rotational capacity compared with steel connections, owing to the contribution of the reinforcing bars located in the slab. The induced tensile forces are resisted by the top bolts and the reinforcing bars and compressive forces are resisted by the steel beam. The reinforcing bars contribute significantly to the strength and stiffness of the connection.

Traditional composite systems utilise concrete derived from Portland cement, which is one of the largest global sources of CO_2 emissions. Moreover, the traditional composite floor systems such as a solid reinforced concrete slab or profiled metal decking floor systems are common systems in composite structures. For typical construction practices for these types of systems, concrete casting, profiled steel decking placing and conventional reinforcing detailing are undertaken on-site, which is time consuming and labour intensive, and which can increase the cost of construction, and they can lead to quality reductions in the construction industry.

Combining precast GC slabs having reduced emissions during their manufacture with steel elements by using a deconstructable shear connection may solve these problems and concerns associated with traditional composite structures. Pre-tensioned high strength bolts installed through holes in precast GC slabs into pre-drilled holes in the steel beam produce a composite flooring system that can be deconstructed at the end of the life of the structure (Bradford and Pi 2012a,b, 2013; Rowe and Bradford 2013; Ataei and Bradford 2013; Lee and Bradford 2013). Marshall et al. (1971) appear to be the first researchers to have reported the use of bolted shear connection, but the context of the usage is not clear. Twelve push tests using high strength bolts as shear connectors were carried out and reported by Dallam (1968). In these set of tests, the bolts were embedded in the concrete slab and pre-tensioned by the turn-of-nut method after the concrete had aged 28 days. He pointed out

that high strength bolts displayed a higher capacity and ultimate strength than stud shear connectors. Six full-scale simply supported composite beams with high-strength bolted shear connectors were tested by Dallam and Harpster (1968), but the bolted shear connectors were embedded in the concrete slabs. Based on this, they concluded that pre-tensioned high strength bolts provide a very rigid connection between the steel beam and concrete slab at service loads, and a reserve capacity sufficient to develop the ultimate moment capacity of the fully composite section is attainable. A series of tests was conducted on three types of 22-mm diameter post-installed shear connectors under static and fatigue loading by Kwon et al. (2010). It was concluded that bolted shear connectors exhibited significantly higher fatigue strengths than stud shear connectors. Five full-scale non-composite beams were constructed to investigate the retrofitting of the bridge beams by Kwon et al. (2011). The reinforced concrete slabs were attached compositely to the steel girder via post-installed connectors in four beams. It was concluded that the strength and stiffness of the non-composite bridge girder can be improved significantly. In the tests conducted by Kwon et al. (2010, 2011), the bolts were embedded in the concrete or grout. Lee and Bradford (2013) conducted two series of push-out tests to obtain the behaviour of the post-installed pre-tensioned bolted shear connectors. The first and the second series of this experimental study included five and four push-out specimens, respectively. All specimens in the first series and two specimens in the second series were constructed by using post-installed pre-tensioned bolted shear connectors. The major differences between the first and second series were size of the precast slab, the reinforcement and the number of bolts. These studies did not focus on the testing and modelling of deconstructable and sustainable semi-rigid flush end plate composite joints.

In order to provide a robust and efficient means for modelling sustainable semi-rigid beam-to-column composite connections with deconstructable bolted shear connectors, the present paper presents a three-dimensional modelling using ABAQUS software. Three specimens having sustainable and deconstructable semi-rigid beam-to-column joints have been tested and the results are used for validation of the finite element model. Precast GC slabs are attached compositely to the steel beam via pre-tensioned bolted shear connectors and the composite beam is connected to H-section columns using a flush end plate with two rows of bolts. The model simulates a composite beam-to-column connection under hogging moment and it includes the non-linear material properties of all constitutive materials of the composite joint. Almost all components were modelled as being of the same size as in the experimental tests, including the steel beam, steel column, flush end plate and bolts in the connection region. For validation of the model, the results of the numerical modelling are compared with the experimental test results and good agreement is achieved. The modelling is shown to provide an efficacious technique for conducting parametric studies, so as to develop design guidance in this novel application in composite construction.

Finite Element Model

Material modelling

For the ABAQUS modelling, the actual stress-strain curves for the materials used can be determined from the material tests. Material tests for the bolts, steel beam and column, reinforcing bars, reinforced concrete and bolted shear connectors were conducted and the results were used for the FE model. Von Mises' plasticity was used to model the all the structural steel parts as an elastic-plastic material with hardening in both tension and in compression. The relationship of the strain and stress for all structural steel parts adapted in the model is illustrated in Fig. 1 (a). The load-slip relationship and the points defining the relationship between the load and slip of the bolted shear connectors adopted in the FE modelling are shown in Fig. 1 (b).

Concrete in compression and tension was represented using the damaged plasticity model in ABAQUS. For concrete under uniaxial compression, the formulation of Carreira and Chu (1985) that is commonly used in numerical modelling was adopted (Fig. 2(a)) as

$$\sigma_c = \begin{cases} E\,\varepsilon & \sigma_c \leq 0 \cdot 35 f_c \\ \dfrac{f_c\,\gamma\left(\varepsilon/\varepsilon_c\right)}{\gamma - 1 + \left(\varepsilon/\varepsilon_c\right)^\gamma} & \sigma_c > 0 \cdot 35 f_c, \end{cases} \tag{1}$$

where $\varepsilon_c = 0 \cdot 002$,

$$\gamma = \left(f_c / 32 \cdot 4 \right)^3 + 1 \cdot 55 \qquad (2)$$

is a factor which controls the curvature of the stress-strain relationship and f_c is the mean compressive cylinder strength of the concrete in units of MPa.

In order to model the concrete in tension, the tensile stress was assumed to increase linearly to $0 \cdot 1$ of its compressive strength. After cracking the concrete, the stress declines to zero at strain of about 10 times the failure strain, as can be seen in Fig. 2(b).

(a) (b)
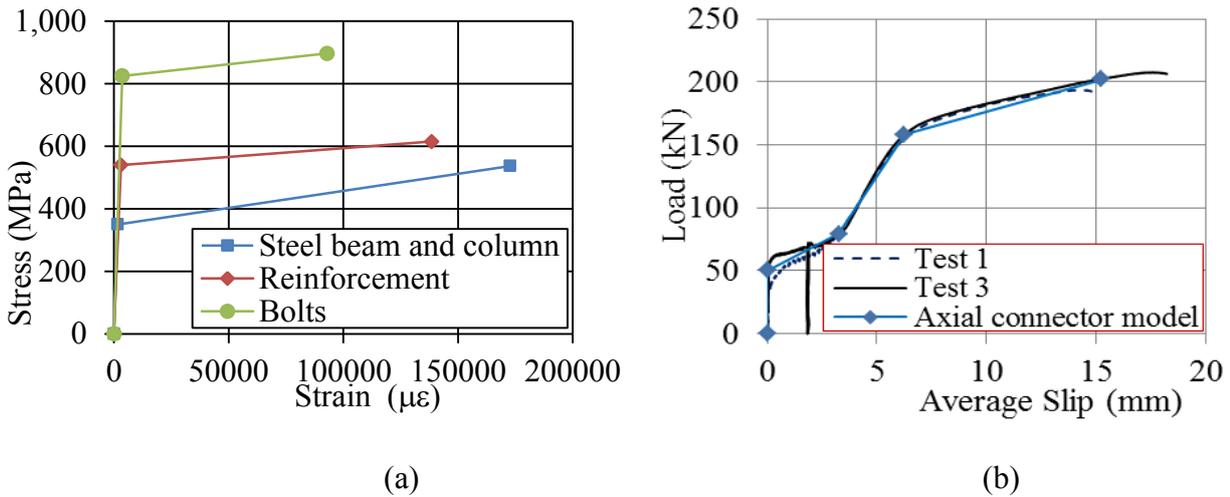
Fig. 1. Stress-strain relationship adapted in FE modelling; (a) Reinforcing bars, bolts and steel beam and column, (b) Axial connectors

(a) (b)
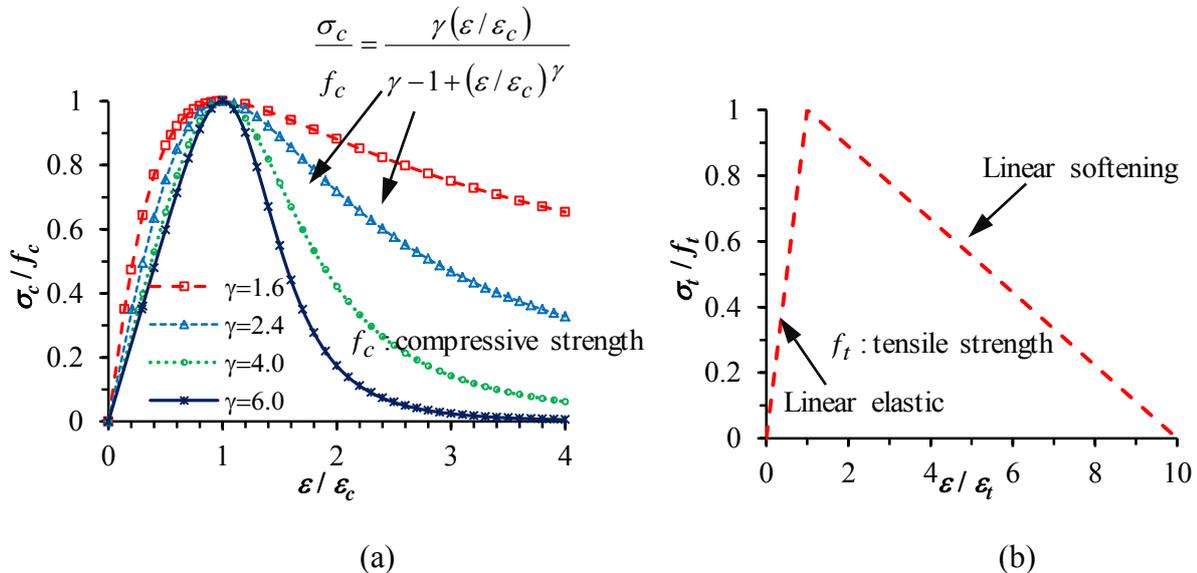
Fig. 2 Outline of the normalised uniaxial stress-strain relationship for concrete under (a) compression (Carreira and Chu, 1985) (b) tension.

Element Type and modelling strategy

The type and size of elements sometimes have a significant effect on the results, and so the determination of the element type and size is one of the important issues in FEM. A finer mesh leads to a better result. However, the finer mesh may leads to computational time problems. Three-dimensional solid elements were used to model the bare and composite connections. Except for the reinforcing bars, all components are modelled by 8-node solid elements (C3D8R) with a reduced integration scheme which prevents shear locking, reduces computational time, and provides the required accuracy (Bursi and Jaspart 1998; Bathe 1996; Cook et al. 2002). For modelling the steel reinforcement, a two-node linear truss element (T3D2) is used. The reinforcement was embedded into the slab, with the slab being the host region and the bars being an embedded region. This technique connects these two different components and prevents slip between them. A typical FE model of a composite joint is shown in Fig. 3.

(a)

(b)
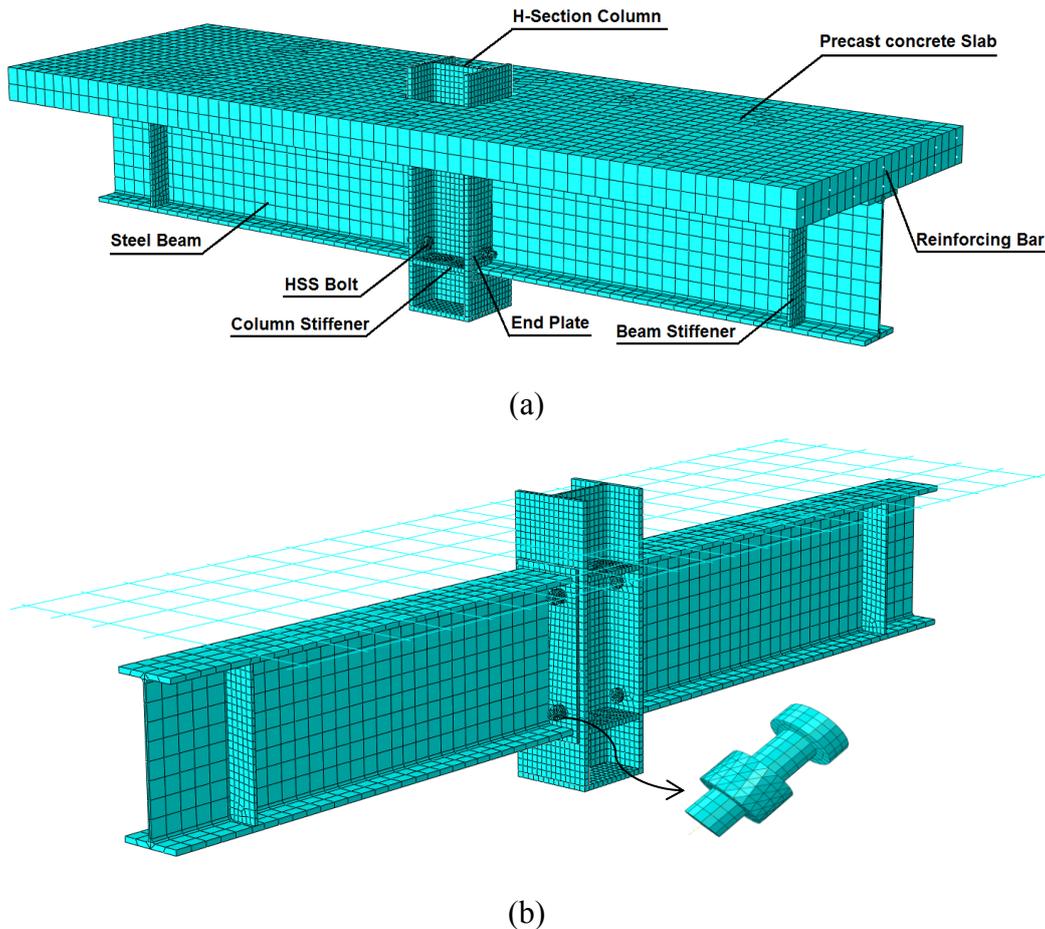
Fig. 3. Finite element model of a composite joint; (a) with showing concrete slab (b) without showing concrete slab

Contact Modelling

There are various components in composite connections that interact with each other, and the results of the FE analysis depend on the accurate modelling of the contact interaction between these components. Experimental results show that there is no separation between the head of the bolt and the flush end-plate, nor between the nut of the bolt and the inner face of the steel column. Therefore, the 'TIE' option was used for connecting these components which provides full interaction between the bolts, nut and bolt head. The 'TIE' option was also used for connecting the steel beam to the flush end-plate, because these two components are welded together. In order to simulate the interaction between the top flange of the steel beam and the lower part of the precast concrete slab, surface-to-surface contact interaction using a penalty method with a coefficient of

friction of 0·25 was adopted, in which the top flange and the concrete slab were considered as master and slave surfaces, respectively.

Bolted shear connector model

Modelling the interface between the concrete slab and shear connectors is one of the main issues in the FE modelling of the composite beam and joint. In this research, the strength and the stiffness characteristics of the pre-tensioned bolted shear connectors determined from the push tests conducted by authors (Fig. 1(b)) are used for modelling of the connection between the concrete slab nodes and steel flange nodes, and so an axial connector model was used to model the interface slip. These connectors were located at the same positions where bolts were placed on the specimen. A schematic diagram of the axial connector model is shown in Fig. 4. The fracture of the bolter shear connectors was assumed to occur at the ultimate slip (15 mm) obtained from the push-out tests.
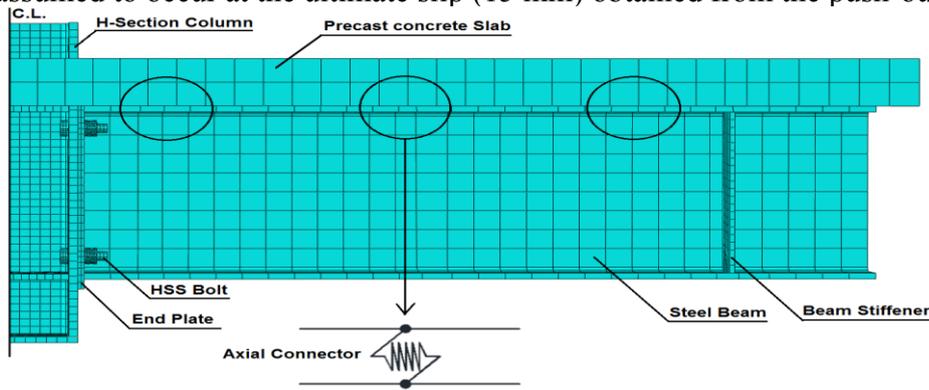
Figure 4: Axial connector model for Composite Joint 1

Load Application and boundary conditions

The loading was applied in two steps. First, the pretension was applied to the bolts located in connection and the joints were then loaded, at which state the bolts were subjected to the pretension as can be seen in Fig. 5. Mirza and Uy (2011) have pointed out that Riks' technique is needed to capture any unloading in the non-linear analysis, and so the GENERAL method and modified RIKS method were used for the first step and second step respectively. A static concentrated load was applied at the centre of the steel column, as was done in the tests. The initial increment plays a vital role in the convergence of the modelling which is why this parameter will be adjusted if a convergence problem occurs during the modelling.

(a)   (b)
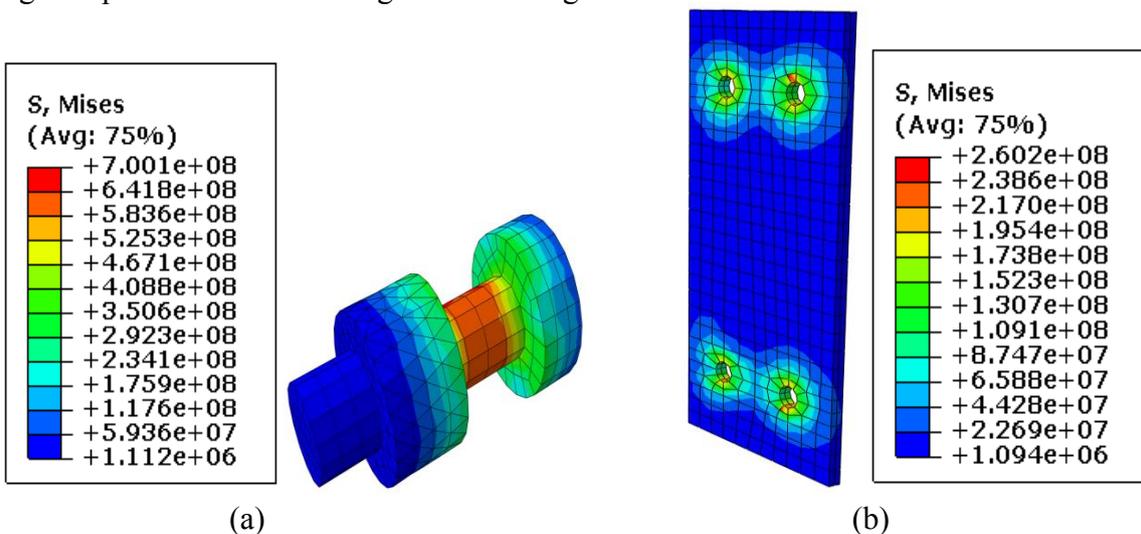
Fig. 5. Stress distribution at the first step of loadings (a) Bolt (b) Flush end plate.

873

In FE modelling, correct representations of the boundary conditions are essential since slightly different boundary conditions can produce significantly different results. For the simulation in this paper, the boundary conditions were taken as being exactly the same as in the tests, with the column being allowed to move in the vertical direction and with the flanking ends of the beams having roller supports.

Experimental study

Three beam-to-column joints were designed and constructed in a cruciform arrangement to simulate the internal joint in a semi rigid frame. The details of beam-to-column joint specimens are summarised in Table 1. A navel methodology of shear connection was adopted by using pre-tensioned bolted shear connectors to attach the precast concrete slabs to the top flange of steel beams. Specimens 1 and 2 were designed as a composite joint (CJ) and Specimen 3, which is a non-composite joint, was designed as a control test specimen to compare against the composite joint tests. All beam-to-column joint specimens consist of a steel beam of 460 UB 82·1 and steel column of 250 UC 89·5. A 12 mm flush end plate welded at the end of steel beam used and connected to the flange of the column by using 4 M24 grade 8·8 bolts. Stiffener plates were welded to the column web at the level of the bottom and top flanges of beam to prevent bending of the column flanges in tension and failure of the column web in compression. The geometric and design details for Composite Joints are presented in Fig. 6.

Table 1: The details of beam-to-column joint specimens.

Specimen	Beam	Column	T_{ep} (mm)	B_d (mm)	R	N_c	B_s (mm)	T_s (Mpa)	N_s
CJ1	460 UB 82.1	250 UC 89.5	12	M24	6N16	6M20	525	120	1
CJ2	460 UB 82.1	250 UC 89.5	12	M24	6N16	6M20	525	120	1
SJ3	460 UB 82.1	250 UC 89.5	12	M24	6N16	6M20	N.A.	N.A.	N.A.

Notes: T_{ep}=end plate thickness; B_d=bolt diameter in connection region; R=Reinforcing; N_c= no. of bolts as shear connectors per beam; B_s= shear connector spacing; T_s=slab thickness; N_s=No. of slab units.

(a)

(b)
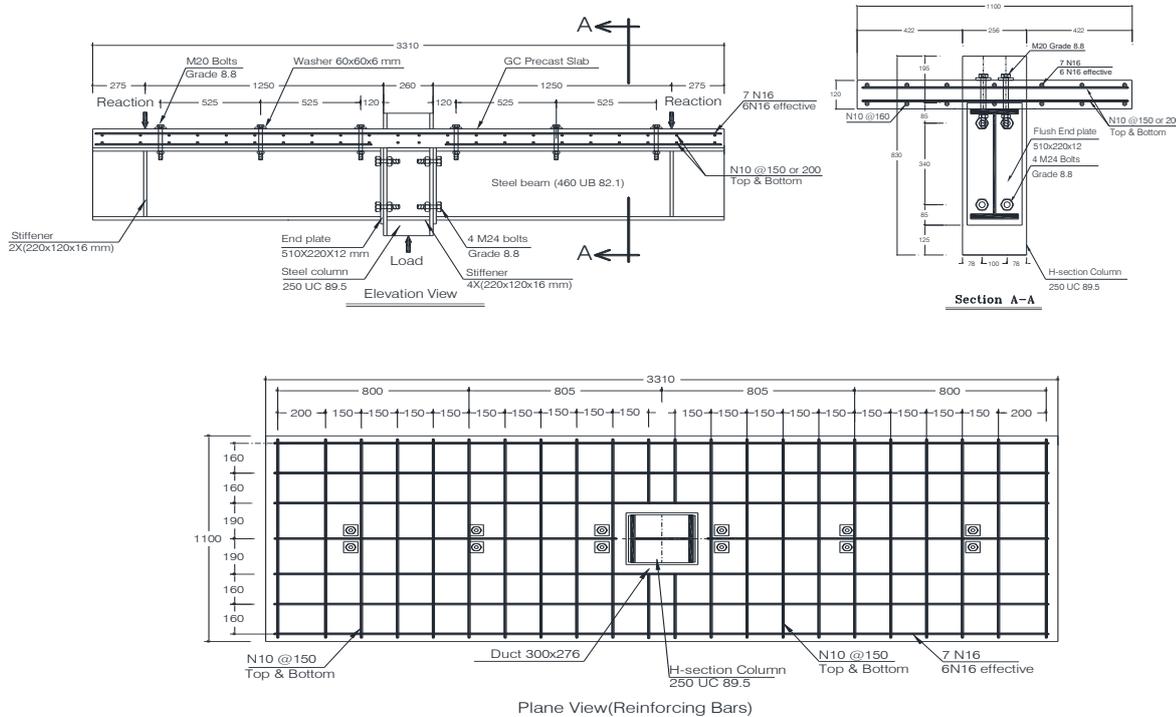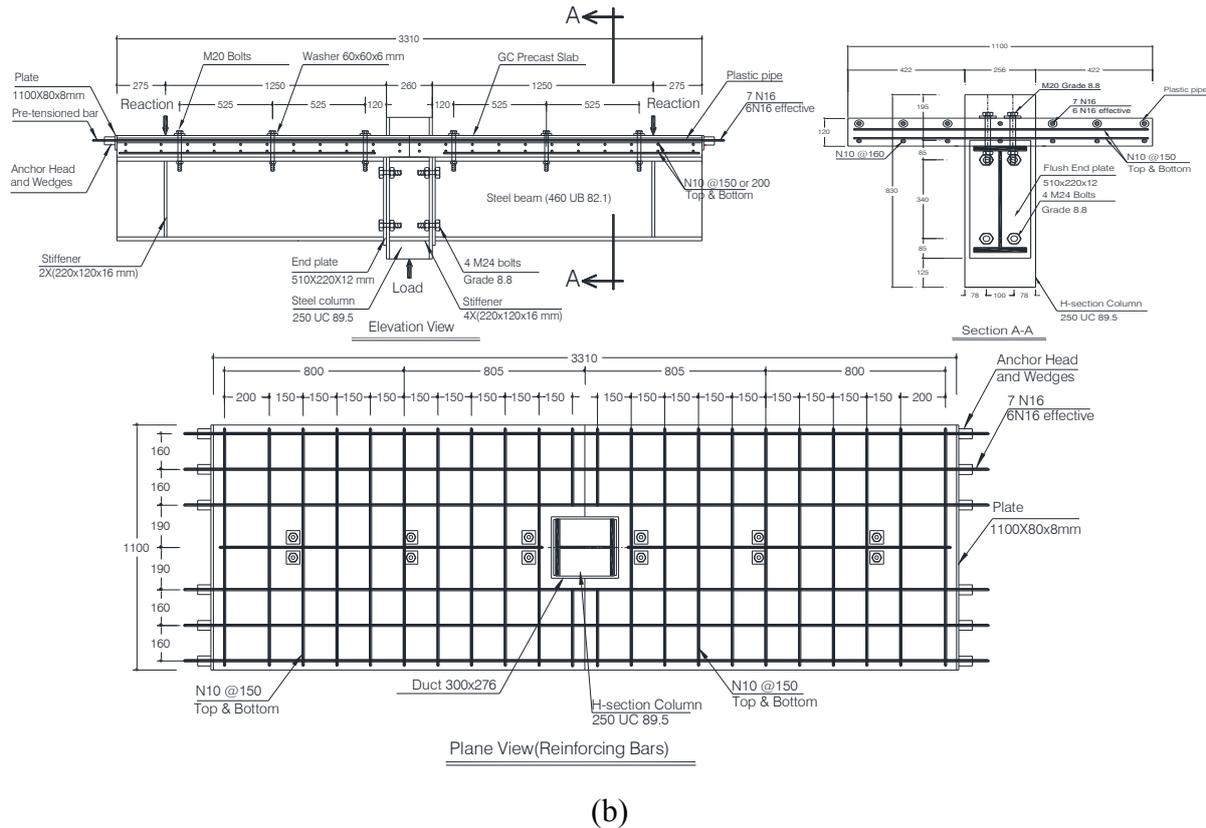
Figure 6: Details of joints: (a) CJ1; (b) CJ2.

Validation of the FEM

The results of the ABAQUS-based FE modelling are compared herein with the experimental results obtained from experimental study to investigate the accuracy and reliability of the FEM, as well as to validate the FEM. Three beam-to-column joint specimens were tested in the Heavy Structures Research Laboratory at the University of New South Wales and the results were used for validation of the FE model. These full-scale semi-rigid flush end plate beam-to-column composite joint tests with deconstructable bolted shear connectors were conducted under symmetrical loading to evaluate the structural characteristics of these new composite joints.

In this comparison, the load versus deflection response was modelled, and the results are given in Figs. 7(a) to 7(b) for the two composite connections tested and in Fig. 7(c) for the connection with a bare steel beam. Table 2 shows the comparison of the FE modelling results and experimental tests. It can be seen that the FE model can predict the ultimate load and deflection of all specimens accurately.

CJ1 consists of one unit of the precast concrete panels attached to the top flanges of steel beam by using pre-tensioned bolted connectors. The longitudinal reinforcement ratio for this slab was 0·9%. Six effective longitudinal N16 reinforcing bars were distributed and placed in the top layer of the precast concrete slab. A maximum load of about 505 kN at the deflection of about 55 mm was recorded for this composite joint and the specimen behaved non-linearly before failure. The load versus deflection response was modelled, and the results are given in Fig. 7(a) for CJ1. It can be seen that the agreement is very good.
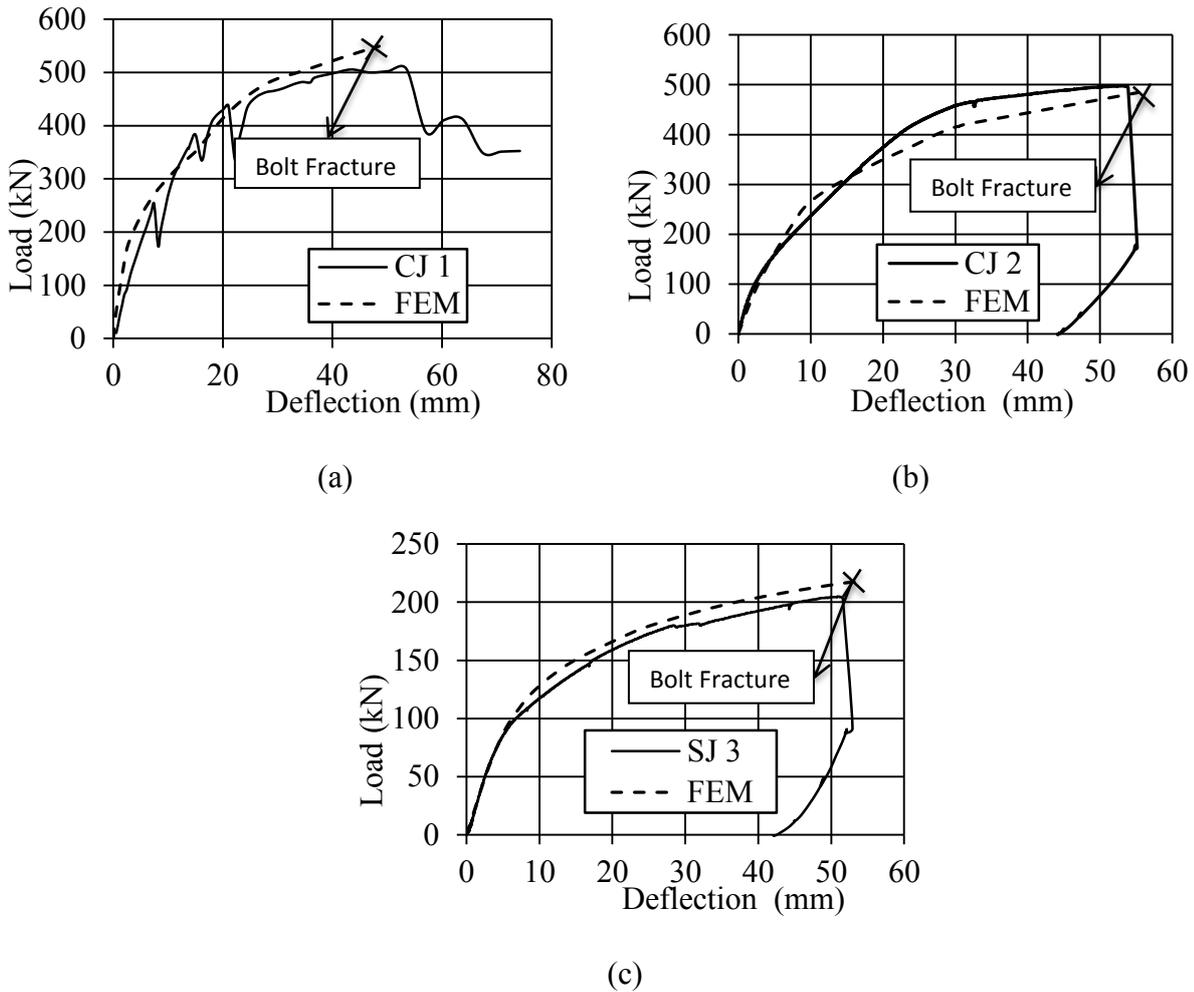
Fig. 7. Comparison of FE model and tests for load-deflection response; (a) CJ1, (b) CJ2, (c) SJ3

Table 2: Comparison of FE model and tests.

	P_{ult}		D_{ult}	
	Test	FEM	Test	FEM
CJ1	505	549	55	48.5
CJ2	498	490	54	56
SJ3	205	217.5	52	53.8

Notes: P_{ult}= ultimate load, D_{ult}= ultimate deflection

CJ2 was similar to CJ1 except that CJ2 consisted of two separated precast concrete panels attached to the top flanges of steel beam. N16 Longitudinal reinforcing bars were not placed in the top layer of the slabs before concrete casting. In order to connect two precast concrete panel together a pre-tensioning procedure was used. The N16 reinforcing bars were placed into the prepared holes by using plastic tube and then stressed to about 10% of their axial load capacity before testing. A maximum load of about 498 kN at the deflection of about 54 mm was recorded for this composite joint and the specimen behaved non-linearly before failure. The load versus deflection response was modelled, and the results are given in Fig. 7(b) for CJ2. It can be seen that the agreement is very good.

SJ3 was designed as the reference and control test specimen. The steel joint (SJ3) is similar to the two composite joints except that SJ3 is non-composite and precast concrete slab was not attached on the top flange of the steel beam. A maximum load of about 205 kN at the deflection of about 52 mm was recorded for this specimen and it behaved non-linearly before the failure. The load versus

deflection response was modelled, and the results are given in Fig. 7(c) for SJ3. It can be seen that the agreement is very good.

Failure modes

All specimens failed by fracture of the M24 bolts due to the resulting tension forces. The FE model can predict accurately the failure mode of the composite connection. Fig. 8 shows the strain distribution for the bolt in the mode of failure for CJ1. As can be seen in Fig. 8, the maximum strains in the shank of the bolt are almost same as the strain capacity of the bolt obtained from the tensile test on the bolts. Table 3 shows a comparison of the failure modes between the FE modelling and experimental testing. It can be seen that the FE model can predict the failure mode of all specimens accurately. The FEM can also predict accurately the plastic deformation and bending of the flush end plate, as illustrated in Fig. 9.

(a) (b)
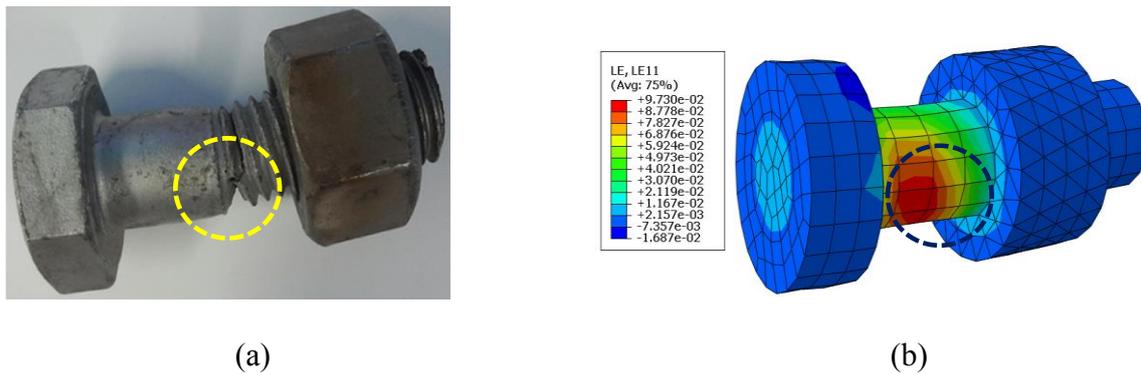
Fig. 8. Bolt failure at the failure mode; (a) Test (CJ1) , (b) FEM

Test (CJ 1) FEM
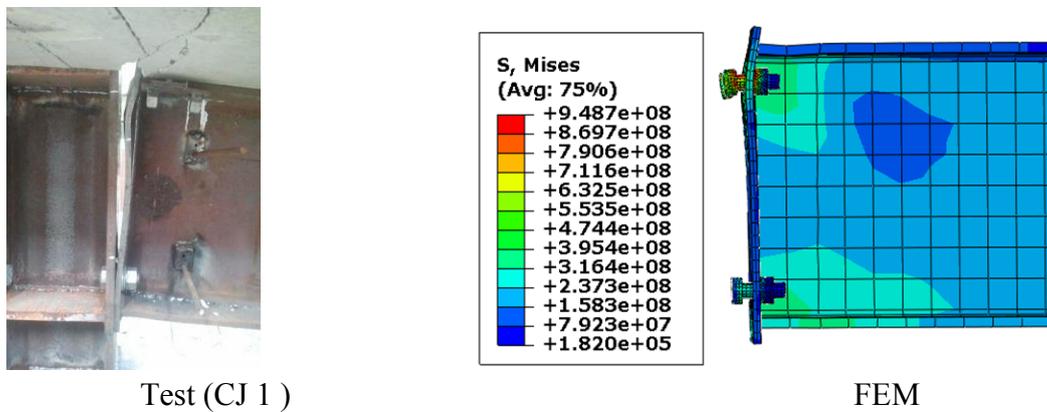
Fig. 9. Deformation and bending of the end plate; (a) Test (CJ1), (b) FEM (in Pa)

Table 3: Comparison of the failure mode.

	CJ1	CJ2	SJ3
Test	BF	BF	BF
FE Model	BF	BF	BF

Note: BF: Bolt failure

Conclusions

This paper has described a numerical modelling of semi-rigid flush end plate beam-to-column composite joint tests with deconstructable bolted shear connectors. Three specimens with sustainable and deconstructable semi-rigid beam-to-column joints have been tested and the results

were used for validation of the finite element model. Precast GC slabs were attached compositely to the steel beam via pre-tensioned bolted shear connectors and the composite beam was connected to H-section columns using flush end plates with two rows of bolts. The model simulates a composite beam-to-column connection under hogging moment and includes non-linear material properties of all constitutive materials of the composite joint. For validation of the model, the results of the numerical modelling were compared with experimental test results and good agreement was achieved. The modelling was shown to provide an efficacious technique for conducting parametric studies, so as to develop design guidance in this novel application in composite construction.

Acknowledgements

References

Ataei A. and Bradford M.A. (2013) "FE Modelling of sustainable semi-rigid flush end plate composite joints with deconstructable bolted shear connectors", International Conference on Composite Construction CCVII, Palm Cove, Queensland.

Bradford M.A. and Pi Y.-L. (2012a) "Numerical modelling of deconstructable composite beams with bolted shear connectors", Conference on Numerical Modeling Strategies for Sustainable Concrete Structures, Aix-en-Provence, France, II-2, 1-8.

Bradford M.A. and Pi Y.-L. (2012b) "Numerical modelling of composite steel-concrete beams for life-cycle deconstructability", 1st International Conference on Performance-Based and Life-Cycle Structural Engineering, Hong Kong, 102-109.

Bradford M.A. and Pi Y.-L. (2013) "Nonlinear elastic-plastic analysis of composite members of high-strength steel and geopolymer concrete", Computer Modeling in Engineering and Sciences 2320: pp. 1-27.

Bursi OS, Jaspart JP. Benchmarks for finite element modelling of bolted steel connections. Journal of Constructional Steel Research 1998;43(1-3);17-42.

Bathe KJ. Finite element procedures. Enwood Cliffs (NJ): Prentice-Hall; 1996.

Cook RD, Malkus DS, Plesha ME, Witt RJ. Concepts and application of finite element analysis. 4thedition John Wiley & Sons Inc;2002.

CEN. "Eurocode 4: Design of Composite Steel and Concrete Structures. Part 1.1: General Rules and Rules for Buildings (prEN 1994-11:2004)".

Dallam L.N. (1968) "Pushout tests with high strength bolt shear connectors", Report for Missouri State Highway Department, Department of Civil Engineering, University of Missouri-Columbia, Missouri, USA.

Dallam L.N. and Harpster J.L. (1968) "Composite beam tests with high-strength bolt shear connectors", Report for Missouri State Highway Department, Department of Civil Engineering, University of Missouri-Columbia, Missouri, USA.

Kwon G., Engelhardt, M.D. and Klinger, R.E. (2010) "Behavior of post-installed shear connectors under static and fatigue loading". Journal of Constructional Steel Research, Vol. 66, pp. 532-541.

Kwon G., Engelhardt, M.D. and Klinger, R.E. (2011) Experimental behavior of bridge beams retrofitted with post-installed shear connectors. Journal of Bridge Engineering, ASCE 16: 536-545.

Lee S.S.M. and Bradford M.A. (2013) "Sustainable composite beams with deconstructable shear connectors", 5th International Conference on Structural Engineering, Mechanics and Computation, Cape Town, South Africa.

Marshall W.T., Nelson H.M. and Banerjee, H.K. (1971) "An experimental study of the use of high-strength friction-grip bolts as shear connectors in composite beams", The Structural Engineer, Vol. 49, pp. 171-178.

Mirza, O. & Uy, B. (2011), Behaviour of composite beam-column flush end-plate connections subjected to low-probability, high-consequence loading. Engineering Structures 2011; 33(2):642-662.

Oehlers, D.J. and Bradford, M.A. (1995) Steel and Concrete Composite Structural Members: Fundamental Behaviour, Pergamon, Oxford.

Rowe M. and Bradford M.A. (2013) "Partial shear interaction in deconstructable composite steel-concrete composite beams with bolted shear connectors", International Conference on Design, Fabrication and Economy of Welded Structures, Miskolc, Hungary, pp. 585-590.

3D In-situ XCT Image Based Meso-scale Fracture Modelling and Validation of Concrete Using Voxel Hexahedron Meshing and Damage Plasticity Model

Yujie Huang[1], Zhenjun Yang[1,2]*, Wenyuan Ren[2], Guohua Liu[1], Chuanzeng Zhang[3]

[1]College of Civil Engineering and Architecture, Zhejiang University, 310058, China.
[2]School of Mechanical, Aerospace and Civil Engineering, the University of Manchester, M13 9PL, UK.
[3]Department of Civil Engineering, University of Siegen, Paul-Bonatz-Str. 9-11, D-57076 Siegen, Germany

*Corresponding author: zhjyang@zju.edu.cn

Abstract

Three-dimensional (3D) meso-scale finite element models of concrete in compression based on in-situ X-ray Computed Tomography (XCT) images are developed and validated in this study. The micro-scale images from a Brazilian-like in-situ XCT test are first compressed and then transformed into manageable meso-scale 3D meshes using a voxel hexahedron meshing technique with a stacking algorithm. The concrete damaged plasticity model in ABAQUS is used to simulate complicated damage and fracture behaviour of concrete. Excellent qualitative agreement is found between the simulations and the XCT test in terms of damage evolution and fracture process on both the surface and interior of the specimen. The effects of internal heterogeneous meso-structures on the macro-scale loading-carrying capacities and failure patterns are quantitatively and qualitatively evaluated by modelling different uniaxial loading directions.

Keywords: Concrete in compression, In-situ X-ray computed tomography, Voxel hexahedron meshing, 3D Image based modelling, Concrete damage plasticity model, Finite element method

1. Introduction

Traditional concrete fracture models assuming homogeneous material properties often predict unrealistically smooth or wrong crack paths and load-carrying capacity of unknown reliability due to its multi-phase, heterogeneous internal structures at micro/meso-scales [Yang and Xu (2008)]. It is highly necessary to conduct micro/meso-scale modelling for accurate understanding of complex damage initiation and evolution until failure, and the relationships between physical properties of multi-phases and the macro-scale mechanical responses [López et al. (2008)].

As to meso-scale modelling, the meso-structures can be directly represented by different phases artificially generated and randomly distributed in space [Caballero et al. (2006); López et al. (2008) Yin et al. (2013)], or be indirectly modelled by random fields satisfying certain correlation functions describing heterogeneous material properties [Yang and Xu (2008); Yang et al. (2009); Su et al. (2010)]. However, most of these studies use assumed meso-scale morphologies or random fields that are not the same as the real internal structures so that the numerical models cannot be directly validated. In addition, most of the existing studies are in 2D and cannot predict non-planar 3D fracture surfaces in reality. Consequently, the simulated results may be neither representative nor fully verifiable. This has led to development of numerical models that are converted from images captured by digital cameras and microscopes etc [Young et al. (2008)]. In this way, more accurate micro/meso-structures can be directly simulated.

Recently, the X-ray Computed Tomography (XCT) technique with non-destructive multi-length scale capabilities becomes increasingly popular, mostly to acquire micro/meso-scale internal structures of concrete [Wang et al. (2003)], and occasionally, to observe damage evolution and fracture process using in-situ XCT tests [Yang et al. (2013)]. The XCT-images have been converted

using commercial packages AVIZO and Simpleware to finite element (FE) meshes by the second author's group, for 2D and 3D cohesive fracture modelling with limited success [Ren et al. (2013; 2014)] and 3D homogenization of elastic properties [Sharma et al. (2014)]. However, in the commercial packages, 3D surface contours are extracted from image datasets and then discretised in mesh generation [Canton and Gilchrist (2010)], which often results in many distorted FE elements when large element sizes are used, or otherwise millions of elements that are beyond the power of conventional computers.

In this study we develop a novel 3D XCT-image based meso-scale FE fracture modelling method for concrete under compression, and attempt to validate 3D damage initiation and evolution until failure predicted by the models for the first time. The voxel hexahedron meshing method [Keyak et al. (1993); Hollister et al. (1994); Crawford et al. (2003)] is augmented with image compression and slice stacking algorithms to efficiently generate 3D FE meshes. It avoids the problem of commercial packages and is able to control the mesh density while maintaining the original 3D morphology. The concrete damage plasticity (CDP) model in ABAQUS is used to simulate complicated damage initiation and evolution in concrete under compression. The in-situ XCT test of a concrete cube under Brazilian-like compression [Yang et al. (2013)] is modelled to validate the developed method, followed by detailed investigation of traditional uniaxial compressive tests.

2. XCT-image based hexahedron mesh generation

Most of the existing image-based 3D FE models based on voxel hexahedron meshing are constructed by direct conversion of voxels in digital images to the same-sized cubic finite elements [Hollister et al. (1994)]. They cannot readily adjust element size while faithfully maintaining the original morphology. In this study, a bottom-up algorithm with the following steps is proposed and fully automated in a MATLAB code:

2.1 2D image processing

For each slice of images from the in-situ XCT test [Yang et al. (2013)], there are 372 pixels of 0.1mm in both directions (Figure 1a). The grey value of pixels ranges from 0 to 255 and drastically fluctuates near the phase interfaces. Segmentation is conducted on each slice using proper thresholds, resulting in ternary images with 1 for aggregates, 2 for mortar and 3 for voids (Figure 1a, refer to [Ren et al. (2013)]). To build lower resolution models, the segmented images are compressed by re-building connectivity of aggregate pixels. Adjustment of grey value of a small number of pixels is then carried out to maintain the morphological details. Figure 1b and 1c show the compressed image with 0.2mm and 0.4 mm pixels, respectively.
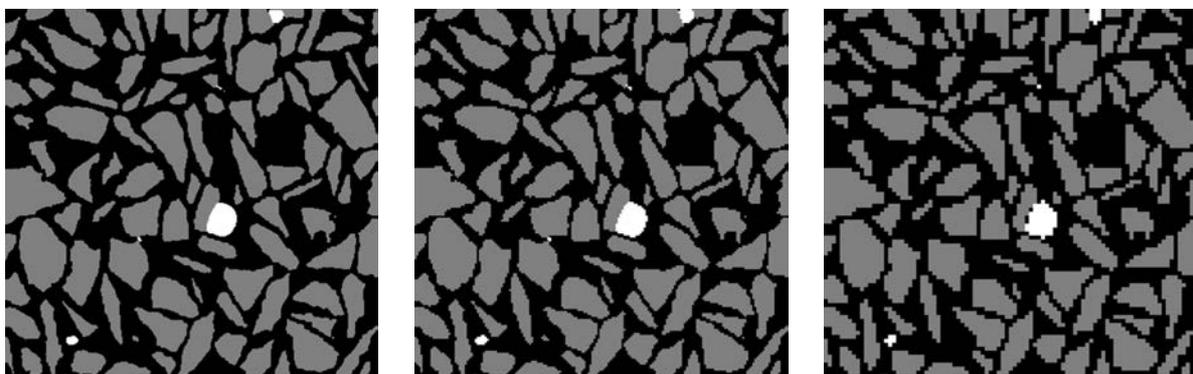
(a) 372×372 pixels (0.1mm)   **(b) 186×186 pixels (0.2mm)**   **(c) 93×93 pixels (0.4mm)**
Figure 1. Image compression

2.2 Slice stacking

The 2D image slices are then stacked to generate voxels. The stacking algorithm widely used [Terada et al. (1997); Huang and Li (2013)] is adopted here and illustrated in Figure 2. The resultant voxels can be converted into eight-noded hexahedral elements and directly used for modelling inclusions of simple shapes, as in the above existing studies, but not for modelling complicated aggregates and voids in concrete of this study. To maintain the true internal morphology, further operations on the voxels are carried out to avoid cases such as contact between aggregates, mortar inside aggregates etc. Figure 3a and 3b show the resultant morphology of aggregates and mortar, respectively. The initial cracks and voids are shown in Figure 3c.
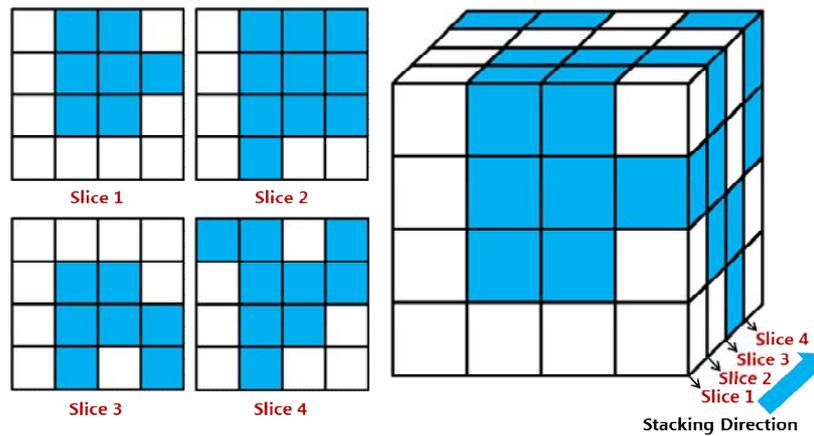
Figure 2. Slice stacking
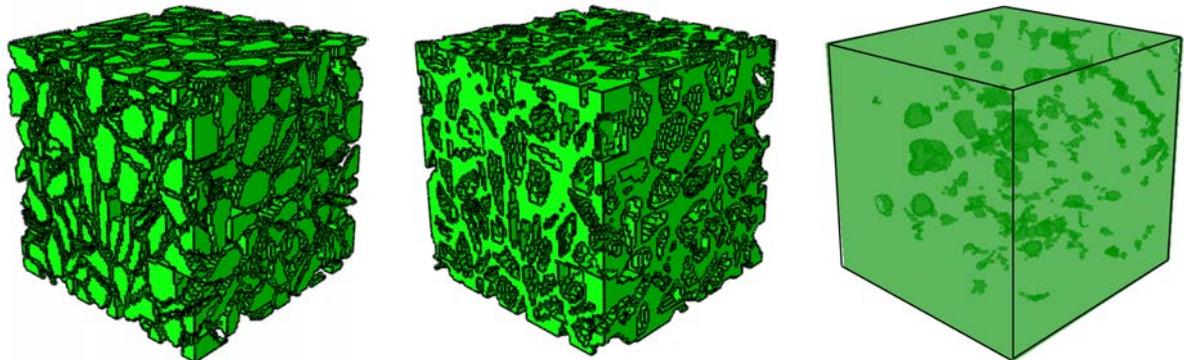
(a) Aggregate   **(b) Mortar**   **(c) Initial voids and cracks**
Figure 3. 3D meso-scale morphology after slice stacking

2.3 Identification of interfacial transition zone (ITZ)

In this step, the mortar voxels connected with aggregate voxels are identified and used to model the weaker aggregate-mortar interfaces, namely the interfacial transition zone (ITZ). Figure 4 shows a small part of the full model with the ITZs displayed in red as an example. The ITZ thickness is 0.4 mm for the coarsest model and 0.1mm for the finest model, respectively.
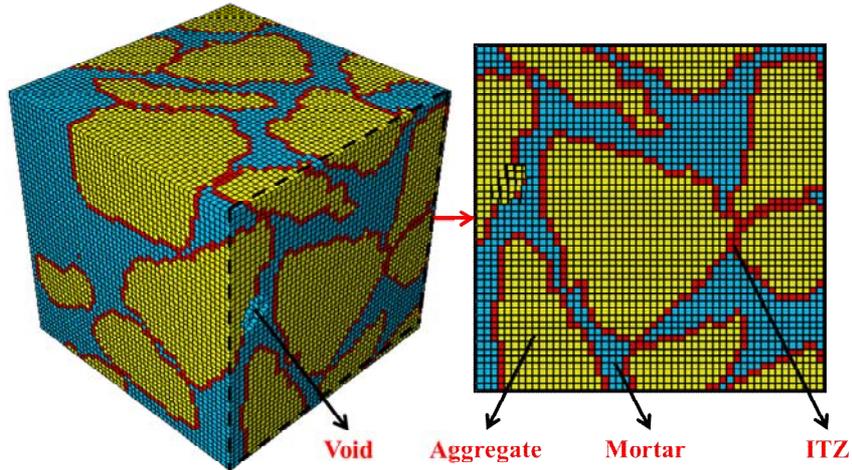
Figure 4. Identification of materials with ITZs

2.4 Generation of 3D FE meshes

After the material labels of all the voxels are determined, the FE meshes are generated by converting each voxel into an eight-noded hexahedral cubic element. Figure 5b shows the final model with ITZ highlighted (the mesh is too dense to be seen clearly), compared with the original XCT image in Figure 5a.

To investigate the effect of image compression and slice stacking, the volume fractions of different phases are calculated. It is 45.755% for aggregates and 1.068% for voids respectively for 0.4mm size model, compared closely with 48.212% and 0.912% respectively for the original 0.1mm size specimen. This suggests that the proposed method is able to maintain the original internal morphology even with 64-time compression of voxel number and thus a nearly 64-time reduction in 3D cubic finite elements.
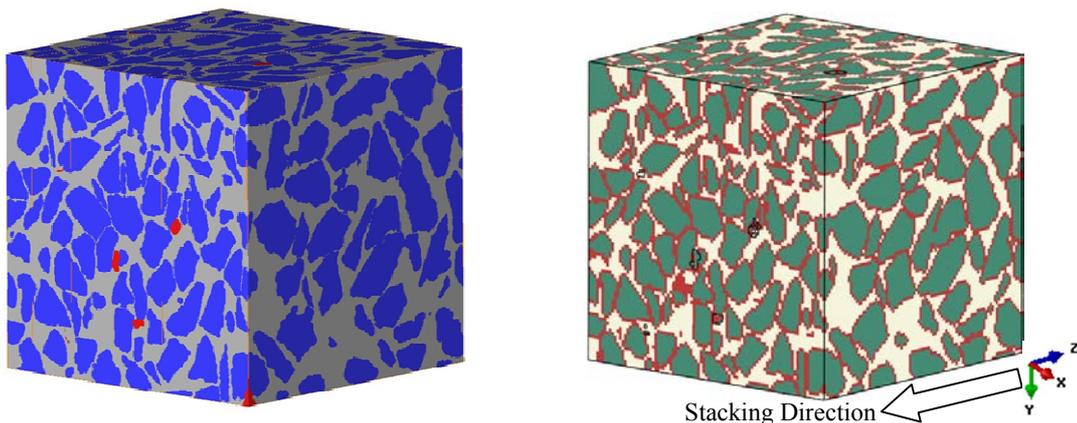
(a) XCT specimen   (b) Numerical model

Figure 5. Comparison of morphology

3. Numerical Simulations and Validation

3.1 In-situ XCT Test

The in-situ XCT test [Yang et al. (2013)] of a concrete cube is modelled first. The tested concrete cube is 40mm (see Figure 6 below). The compressive loading is applied on a central area of 17.5

mm×17.5 mm on the top face; the bottom face is fixed in the same area. This loading condition is similar to typical Brazilian tests. The voxel resolution in the XCT scans is 37.2 μm.

The uneven surfaces of the 40mm cube are removed to build the 37.2 mm FE model shown in Figure 5b with uniform element size 0.4mm. The model consists of 795,764 cube elements and 837,371 nodes. The loading is applied by uniformly distributed displacements on the loaded area with the maximum displacement 0.186mm or strain 0.005. The ABAQUS/Explicit solver is employed with total time 0.01s to ensure the quasi-static loading condition.

3.2 Material model and parameters

The CDP model in ABAQUS, which has proved very powerful for modelling concrete damage and fracture [Lubliner et al. (1989); Lee and Fenves (1998); Chen et al. (2012); Mahmud et al. (2013)], is used to model mortar and ITZs, and the aggregates are assumed elastic. The compressive strength is 35MPa and 27MPa for mortar and ITZs, respectively. The corresponding tensile strength is 4.5MPa and 3.5MPa, respectively. The stress-strain curves in the 2010 Chinese Code for Design of Concrete Structures are used to model compression and tension. The tensile behaviour is defined by an equivalent stress-displacement curve to ensure mesh independence of results. The Young's modulus is 50GPa, 20GPa and 15GPa for aggregates, mortar and ITZs, respectively. Default values in ABAQUS are used for other parameters of the CDP model.

3.3 Validation of final crack pattern

As the material parameters are not from the in-situ XCT test, only the predicted final crack pattern is validated against the XCT test. Figure 6 compares the failed specimen in the test with predicted macroscopic crack pattern represented by the maximum principal strain on the surface. The crack pattern resembles typical ones in Brazilian tests [Lopez et al. (2008)]. Figure 7 compares the crack paths with respect to the aggregates on the surface in a different view from Figure 6. Figure 8 compares the internal voids and cracks under zero and peak loading from the XCT test and the simulation. A very high level of similarity can be seen in the crack patterns, both on the surface and the interior, indicating the capability of the developed image-based model in predicting the phenomenological damage and fracture processes qualitatively.
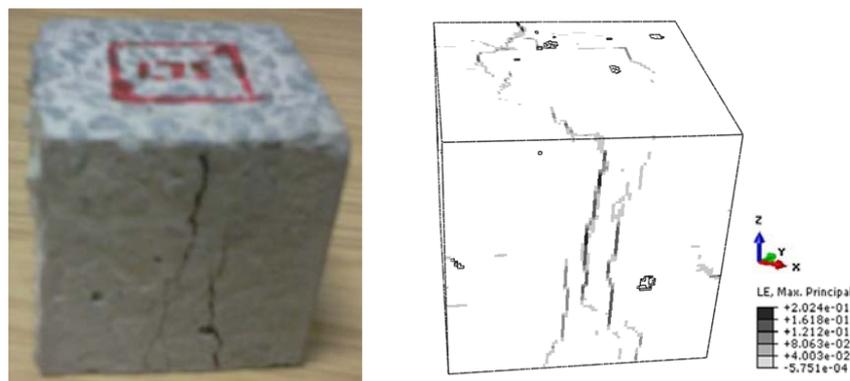
Figure 6. Crack patterns: XCT test (left) and numerical results (right)
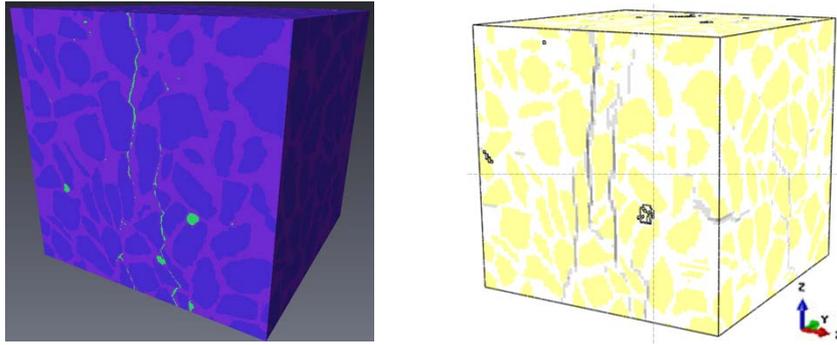
Figure 7. Crack paths: XCT test (left) and numerical results (right)
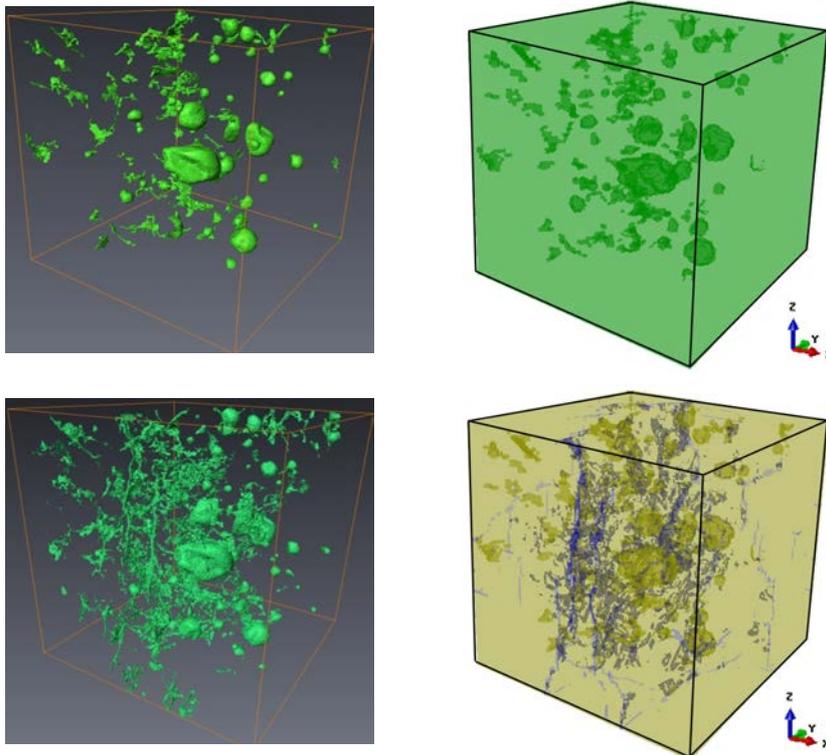
Figure 8. Voids and cracks at zero load stage (top) and peak load (bottom): XCT test (left) and numerical results (right)

More detailed comparison can be made for each slice/cross-section inside the specimen. An example of the middle slice vertical to the loading direction (z) at peak load is given in Figure 9. Figure 9a shows the original XCT image and Figure 9b highlights the crack pattern identified by comparing Figure 9a with the image at zero load. Figure 9c and 9d show the predicted maximum principal strain contours from this study and a digital volume correlation (DVC) analysis [Yang et al. (2013)]. A sound resemblance in the crack patterns can be observed.

(a) XCT image　　　　　　**(b) Highlighted cracks**

(c) Present simulation　　　**(d) DVC results [Yang et al, (2013)]**
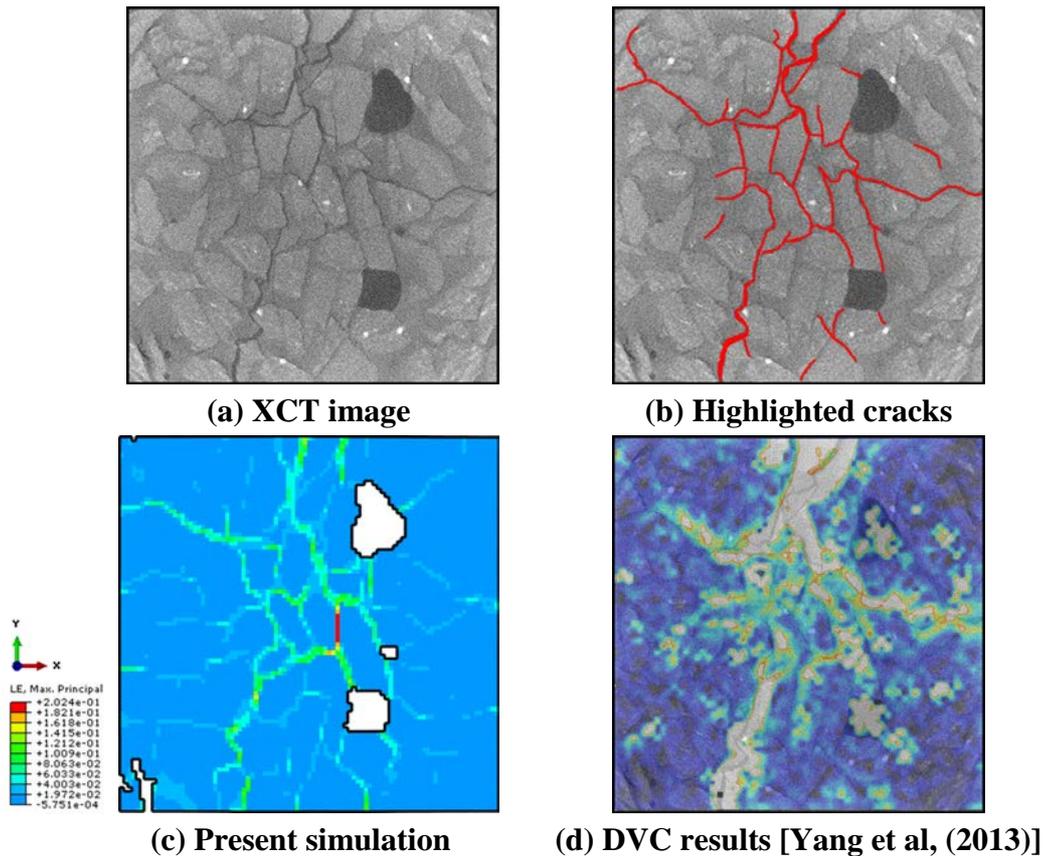
Figure 9. Crack patterns on a slice vertical to the loading at peak load: XCT images (top), numerical results (bottom left) and DVC results (bottom right)

3.4 Further uniaxial compression tests

The same FE model in Figure 5b is used to simulate the uniaxial compression tests, with one entire surface loaded and the opposite fixed vertically without lateral friction. Loading in three directions (x, y and z) is modelled to investigate the effects of multiphase distribution. The predicted stress-strain curves are shown in Figure 10, with key points (A-E) marked. The stress-volumetric strain curves are presented in Figure 11. The different stages in these curves agree well with the typical behaviour of concrete under uniaxial compression [Van Mier (2012)]. The predicted strength is 31.0MPa, 29.0MPa and 28.2MPa for x-load, y-load and z-load respectively, with 9% maximum difference. This together with different post-peak curves indicates the multi-phase distribution at meso-scale affects the macro-scale structural responses. The volumetric strain (Figure 11) decreases first and then increases, reflecting that the specimen becomes more compact under compression first and then expanded after cracks initiate and propagate at late stage. This is also confirmed by the variation of void and crack volume fractions in the XCT test [Yang et al. (2013)].
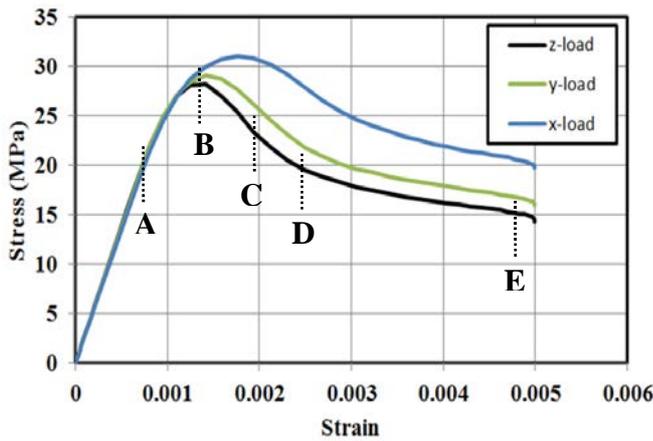
Figure 10. Simulated stress-strain curves for uniaxial compression
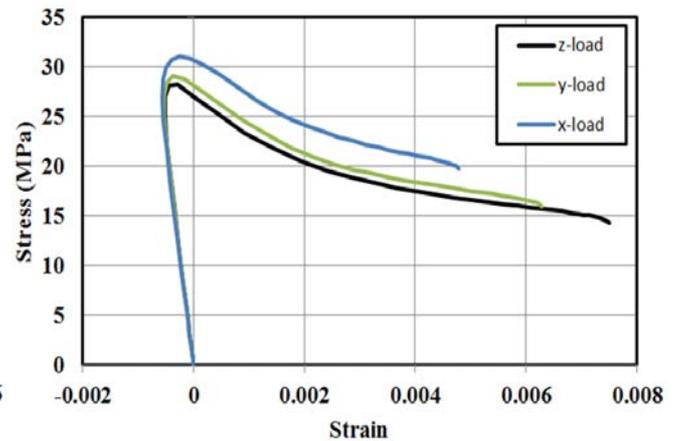
Figure 11. Simulated stress-volumetric strain curves for uniaxial compression

Figure 12 shows the initiation and evolution of compressive damage index (DAMAGEC in ABAQUS) on the surfaces in the front and rear views. The simulated failure pattern with inclined cracks (displayed in red with high damage index) shows similar features to 2D numerical results [Song and Lu (2012)] and typical tests with low friction [Van Mier and Vonk (1991)]. The very different pictures on the front and rear views reflect the heterogeneous mechanical properties caused by random distribution of phases.
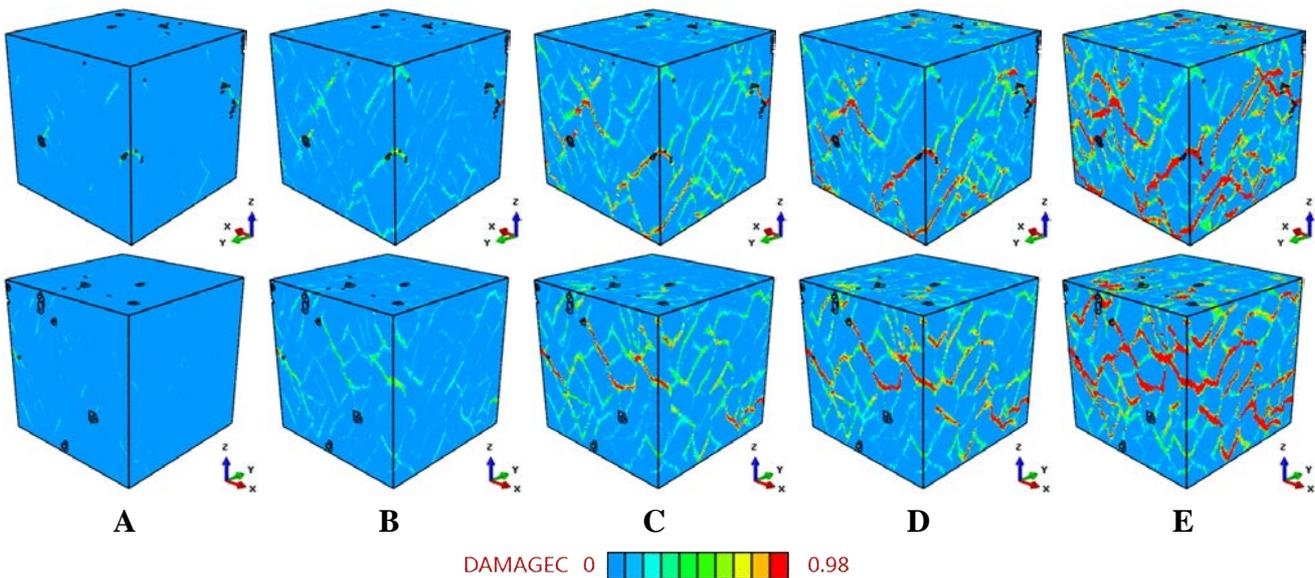
Figure 12. Damage initiation and evolution under z-load at loading points A-E: front view (top) and rear view (bottom)

The internal damage initiation and evolution in a 3D cut-off view (Figure 13) of the model is shown in Figure 14. The view cut is selected to show how the voids and aggregates affect initiation and propagation of damage. From Figure 12 and Figure 14, it can be seen that the damage initiates mostly around the voids. The damage bands tend to propagate towards nearby voids to form a connected 3D damage network with complicated crack bridging and branching. In some cases, the propagation of damage bands is slowed down or hindered by big aggregates (refer to the red arrow in Figure 13) on the way. It can also be noted that the damage and fracture appears to propagate from the surfaces to the interior.
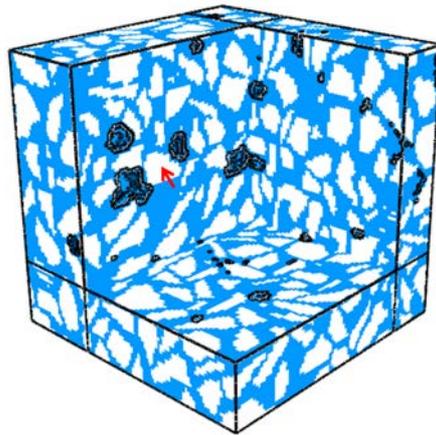
Figure 13. A cut-off view showing main voids

A   **B**   **C**   **D**   **E**

DAMAGEC 0 [color scale] 0.98
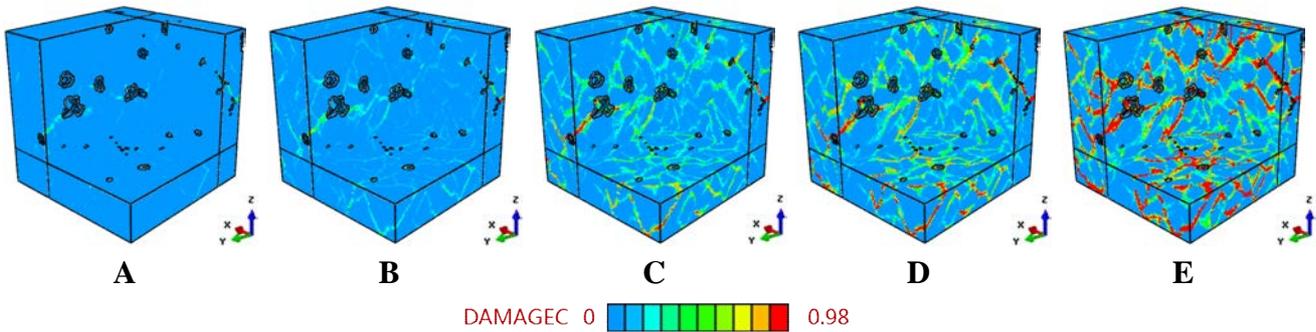
Figure 14. Damage initiation and evolution under z-load at loading points A-E

Figure 15 compares the damage patterns externally and internally at loading point E (Figure 10) from different loading directions. The very different pictures demonstrate that the loading direction, or equivalently, the distribution of phases, can make significant differences to the failure pattern as well as the load-carrying capacities.
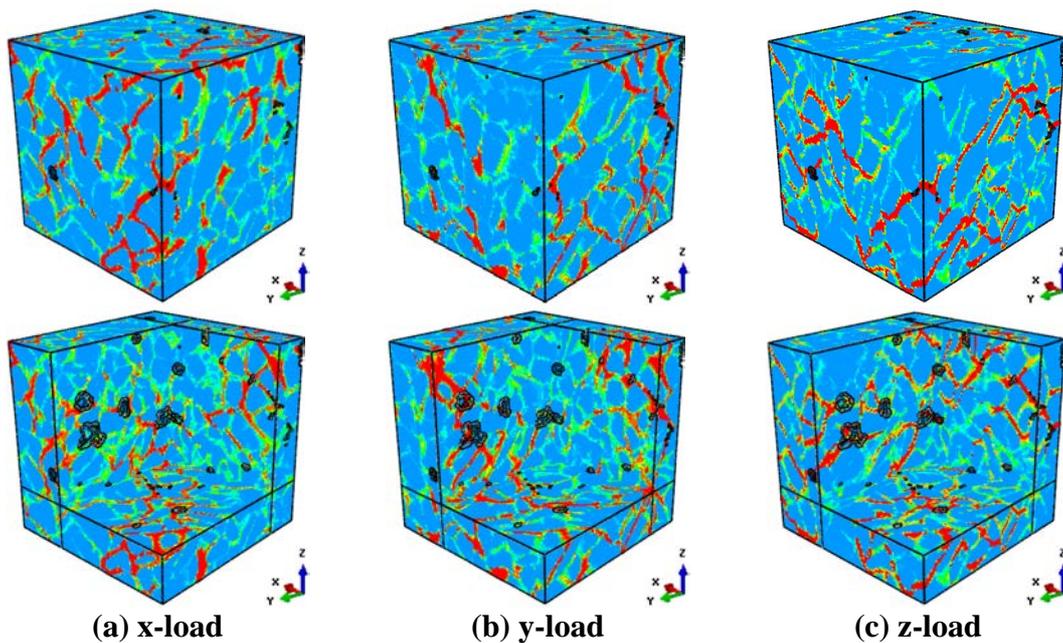
(a) x-load   **(b) y-load**   **(c) z-load**
Figure 15. Failure patterns under different loading directions

4. Conclusions

3D meso-scale FE models based on XCT images have been developed for accurate understanding of damage initiation and evolution until structural failure of concrete under compression, using the voxel hexahedron meshing method and the concrete damage plasticity model. An attempt to validate the developed models has been made against an in-situ XCT test for the first time, although still phenomenologically and qualitatively in terms of 3D damage and fracture processes.

The numerical results clearly show that the intrinsic heterogeneity of meso-structures caused by random spatial distribution of multi-phases can significantly affect macroscopic responses of concrete, e.g., crack patterns and load-carrying capacities. Moreover, the image-based 3D models are very powerful and promising in elucidating the fundamental mechanism of very complicated damage initiation and propagation behaviour that 2D studies are incapable of modelling.

Acknowledgements

References

Caballero, A., López, C. M., and Carol, I. (2006) 3D meso-structural analysis of concrete specimens under uniaxial tension. *Computer Methods in Applied Mechanics and Engineering*, 195(52), 7182-7195.

Canton, B. and Gilchrist, M. D. (2010) Automated hexahedral mesh generation of complex biological objects: dedicated to the memory of Professor Bertram Broberg, colleague and friend. *Strength, Fracture and Complexity*, 6(1), 51-68.

Chen, G. M., Chen, J. F. and Teng, J. G. (2012) On the finite element modelling of RC beams shear strengthened with FRP. *Construction and Building Materials* 2012;32:13–26.

Crawford, R. P., Keaveny, T. M. and Rosenberg, W. S. (2003) Quantitative computed tomography-based finite element models of the human lumbar vertebral body: effect of element size on stiffness, damage, and fracture strength predictions. *Journal of biomechanical engineering*, 125(4), 434-438.

Hollister, S. J. and Kikuchi, N. (1994) Homogenization theory and digital imaging: a basis for studying the mechanics and design principles of bone tissue. *Biotechnology and Bioengineering*, 43(7), 586-596.

Huang, M., and Li, Y. M. (2013) X-ray tomography image-based reconstruction of microstructural finite element mesh models for heterogeneous materials. *Computational Materials Science*, 67, 63-72.

Keyak, J. H., Fourkas, M. G., Meagher, J. M. and Skinner, H. B. (1993) Validation of an automated method of three-dimensional finite element modelling of bone. *Journal of Biomedical Engineering*, 15(6), 505-509.

Lee, J., and Fenves, G. L. (1998) Plastic-damage model for cyclic loading of concrete structures. *Journal of Engineering Mechanics*, 124(8), 892-900.

López, C. M., Carol, I., and Aguado, A. (2008) Meso-structural study of concrete fracture using interface elements. II: compression, biaxial and Brazilian test. *Materials and Structures*, 41(3), 601-620.

Lubliner, J., Oliver, J., Oller, S., and Oñate, E. (1989) A plastic-damage model for concrete. *International Journal of Solids and Structures*, 25(3), 299-326.

Mahmud, G. H., Yang, Z. J., Hassan, A. (2013) Experimental and numerical studies of size effects of Ultra High Performance Steel Fibre Reinforced Concrete (UHPFRC) beams, *Construction and Building Materials*, Vol. 48, 1027-1034.

Mishnaevsky Jr, L. L. (2005) Automatic voxel-based generation of 3D microstructural FE models and its application to the damage analysis of composites. *Materials Science and Engineering*: A, 407(1), 11-23.

Ren, W. Y., Yang, Z. J. and Sharma, R. (2014) 3D meso-scale image-based fracture modelling of concrete using cohesive elements. *22nd UK National Conference on Computational Mechanics in Engineering*. Exeter, April 2014.

Ren, W. Y., Yang, Z. J. and Withers, P. (2013) Meso-scale fracture modelling of concrete based on X-ray computed tomography images. *The 5th Asia-Pacific Congress on Computational Mechanics (APCOM),* 11-14th December, 2013, Singapore.

Sharma, R., Ren, W. Y. and Yang, Z. J. (2014) Homogeneous response of random structured concrete with realistic microstructure. *22nd UK National Conference on Computational Mechanics in Engineering*. Exeter, April 2014.

Song, Z., and Lu, Y. (2012) Mesoscopic analysis of concrete under excessively high strain rate compression

Su, X. T., Yang, Z. J. and Liu, G. H. (2010) Monte Carlo simulation of complex cohesive fracture in random heterogeneous quasi-brittle materials: A 3D study. *International Journal of Solids and Structures*, 47:2336-2345.

Terada, K., Miura T. and Kikuchi N. (1997) Digital image-based modelling applied to the homogenization analysis of composite materials. *Computational Mechanics,* 20(4): 331-346.

Van Mier, J. G. M (2012). Concrete Fracture: A Multiscale Approach. *CRC Press.*

Van Mier, J. G. M. and Vonk, R. A. (1991) Fracture of concrete under multiaxial stress-recent developments. *Materials and Structures*, 24(1), 61-65.

Wang, L. B., Frost, J. D. and Voyjadjis, G. Z. (2003) Quantification of damage parameters using X-ray tomography images. *Mechanics of Materials,* 35: 777-790.

Yang, Z. J., and Frank Xu, X. (2008) A heterogeneous cohesive model for quasi-brittle materials considering spatially varying random fracture properties. *Computer methods in applied mechanics and engineering*, 197(45), 4027-4039.

Yang, Z. J., Ren, W. Y., Mostafavi, M., Mcdonald, S. A. and Marrow, T. J. (2013) Characterisation of 3D fracture evolution in concrete using in-situ x-ray computed tomography testing and digital volume correlation. *VIII International Conference on Fracture Mechanics of Concrete and Concrete Structures,* Toledo, Spain, CIMNE.

Yang, Z. J., Su, X. T., Chen, J. F. and Liu, G. H. (2009) Monte Carlo simulation of complex cohesive fracture in random heterogeneous quasi-brittle materials. *International Journal of Solids and Structures,* 46(17): 3222-3234.

Young, P. G., Beresford-West, T. B. H., et al. (2008) An efficient approach to converting three-dimensional image data into highly accurate computational models. Philosophical Transactions of the Royal Society A: *Mathematical, Physical and Engineering Sciences*, 366(1878), 3155-3173.

Active Control and Potential Exploitation of Parametrically Excited Systems

*M. Ghandchi Tehrani[1] and M. Kalkawski[,1]

[1]Institute of Sound and Vibration Research, University of Southampton, Southampton, UK.

*Corresponding author: m.ghandchi-tehrani@soton.ac.uk

Abstract

This paper presents active control of parametrically excited systems. Parametrically excited systems can exhibit complex dynamic behavior such as inherent instability. Active control can be used to first stabilize and then increase the stability regions of such systems, using velocity feedback. A beam subject to an axial load is considered, representing a parametrically excited system with periodic time-varying stiffness. For amplitudes that are well below the critical buckling load and for axial load excitation at twice the first natural frequency of the beam, the system becomes parametrically unstable. It is demonstrated, how the system can be stabilized using active control. Alternatively, parametric excitation can be exploited for energy harvesting. When the system is close to the transition curves or instability regions, due to the high amplitude level of vibrations, parametrically excited systems can harvest much more energy compared to the time-invariant systems.

Keywords: Parametrically excited systems, Active control

Introduction

Parametrically excited systems include a parameter(s) in their dynamic equations, which varies periodically in time. Parametric resonance is a dynamic instability associated with such systems. It involves interaction between the parametric excitation frequency and the natural frequency of the system, leading to negatively damped modes and unstable oscillations. When the parametric excitation frequency is at about twice the natural frequency, the system exhibits instability, leading to large oscillations and potentially fatigue or failure. One example of parametrically excited systems is cable-stayed bridges, in which the tension of the cables can vary periodically due to the vibration of the deck, resulting in parametric resonance and instability [Reynolds et al. (2006)]. Other examples of such systems are flexible risers or ships, in which the wave motion can be the source of parametric excitation [Ahmed et al. (2010)].

Parametrically excited systems have been the subject of research investigation for decades [Cartmell (1990), Nayfeh and Mook (1995)]. The most simple and widely used equation of parametric excitation is the well-known Mathieu equation with linear, periodic, time varying stiffness coefficient. This rather simple single degree-of-freedom (DOF) mechanical system exhibits complex unstable behaviour but also interesting stability regions, as known from the inverted pendulum problem, depending on the amplitude and frequency parameters of the time periodic (harmonic) term. Parametrically excited systems exhibit combination resonances of summed or difference types. When subjected to an external forcing frequency, a periodically time-varying system will be resonant when the external frequency equals the combination of natural frequency and parametric excitation frequency.

Recent research emphasizes the potential for deliberately introducing parametric excitation to increase the capability of a system to suppress vibrations [Ecker (2010), Dohnal and Mace (2008)]. Parametric excitation can also be exploited for energy harvesting. Daqaq et al. [Daqaq et al. (2009)] investigated the problem of energy harvesting using a parametrically excited cantilever beam. The cantilever beam was excited vertically, perpendicular to the direction of the oscillatory displacement, at twice of its fundamental frequency. In 2011, the same parametrically excited

harvesting beam configuration was considered including higher modes and nonlinear effects of the piezoelectric patch [Abdelkefi et al. (2012)].

This paper describes both active control of parametrically excited systems and their potential exploitation for energy harvesting. In active control, the problem is to suppress large amplitude of vibration or to extend the stability boundaries using velocity feedback control. A beam under axial load is considered, representing a parametrically excited system. For certain amplitudes and frequencies of the axial load, the system can exhibit parametric instability. Active control is used to stabilise the system. In addition, an investigation is carried out to determine the amount of energy that can be harvested from the beam with and without parametric excitation. It is shown that when the system is parametrically excited, it can harvest much more energy.

Parametrically Excited Systems

A vertical cantilever steel beam subjected to an axial time-harmonic load $P(t)$ associated with the base acceleration is shown in Figure 1. A static compressive load is expected to reduce the first natural frequency. When the amplitude of the axial load reaches the critical buckling load, the beam can experience buckling instability and zero stiffness. If the axial load is harmonic, i.e. $P(t) = P\cos\Omega t$, the bending stiffness of the beam varies periodically. For a specific amplitude of parametric excitation P_0 at a frequency Ω almost twice the first bending natural frequency ω_1, the system becomes parametrically unstable.

force distribution
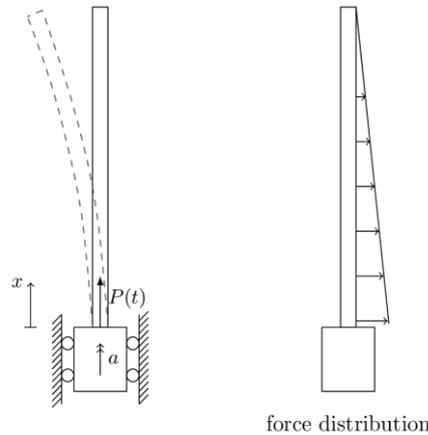
Figure 1: A parametrically excited system- a beam subject to a harmonic axial load

Vibration of an axially loaded cantilever at its first mode can be described using a single degree-of-freedom equation with the time-varying natural frequency. We add a damping term to obtain:

$$\ddot{q} + 2\varsigma\omega_0\dot{q} + \omega_0^2\left(1 - \frac{P(t)}{P_{cr}}\right)q = 0 \qquad (1)$$

This is the well-known damped Mathieu Equation with a periodic time-varying stiffness.

The cantilever beam considered in this paper is mounted vertically, therefore, accounting for the effect of gravity, the forcing term takes the following form:

$$P(x,t) = P_0(x) + P_1(x)\cos\Omega t = m(x)g + m(x)a\cos\Omega t \qquad (1)$$

where P_0 represents the weight of the beam and $P_1(t)$ is the force resulting from the acceleration of the base. We note that the load is not uniform over the length of the beam since the mass is distributed, thus the space average of the mass distribution is used:

$$\frac{1}{L}\int_0^L \rho A x \, dx = \frac{\rho A L}{2} \tag{3}$$

The final form of the governing equation can be written as,

$$\ddot{q} + 2\varsigma\omega_P \dot{q} + \omega_P^2\left(1 - \frac{P_1 \cos\Omega t}{2P_{cr'}}\right)q = 0 \tag{4}$$

with $P_{cr'} = P_{cr} - \dfrac{P_0}{2}$ and $\omega_P^2 = \omega_0^2\left(1 - \dfrac{P_0}{2P_{cr}}\right)$.

Stability Analysis

We approximate the response with three terms in the Fourier series:

$$q(t) = \sum_{n=0}^{\infty} A_n \cos\left(\frac{n\Omega}{2}t\right) + \sum_{n=1}^{\infty} B_m \sin\left(\frac{n\Omega}{2}t\right) \tag{5}$$

Taking derivatives, substituting into Eq. (1), and partitioning the $\sin\left(\dfrac{n\Omega t}{2}\right)$ and $\cos\left(\dfrac{n\Omega t}{2}\right)$ terms,

leads to a set of equations in terms of the coefficients A_n and B_n. The transition curves (stability curves) are obtained from solving the determinant of the coefficient matrices.

$$\begin{vmatrix} \omega_P^2\left(1 - \dfrac{P_1}{4P_{cr'}}\right) - \dfrac{\Omega^2}{4} & \zeta_1\omega_P\Omega \\[4mm] -\zeta_1\omega_P\Omega & \omega_P^2\left(1 + \dfrac{P_1}{4P_{cr'}}\right) - \dfrac{\Omega^2}{4} \end{vmatrix} = 0 \tag{6}$$

which results in the two transition curves separating the stable from the unstable regions,

$$\left(\frac{\omega_P}{\Omega}\right)^4\left(1 - \frac{P_1^2}{16P_{cr'}^2}\right) + \left(\frac{\omega_P}{\Omega}\right)^2\left(\zeta_1^2 - \frac{1}{2}\right) + \frac{1}{16} = 0 \tag{7}$$

Transition curves were computed for a steel cantilever beam which dimensions and properties are listed in

Table 1. The natural frequency associated with the first bending mode accounting for the weight of the beam was found to be $f_P = \omega_P/2\pi = 3.18$ Hz, whereas the critical buckling load diminished by the average weight of the beam was calculated as $P_{cr'} = 2.91$ N.

Table 1: Beam dimensions and material properties

property	b, m	d, m	L, m	E, GPa	ρ, kg/m^3	ς_1
value	0.0105	0.00144	0.583	186	7850	0.002

The two transition curves are plotted using the analytical approximation in Figure 2. It can be seen that parametric instability occurs when the parametric excitation frequency is twice the first natural

frequency $(\Omega = 2\omega_P)$. The effect of damping is to move the transition curves upwards, hence increasing the stable region, when comparing a damped system (black solid line) with an undamped system (grey dashed line). Considering higher order terms in the Fourier series, other transition curves can be obtained.
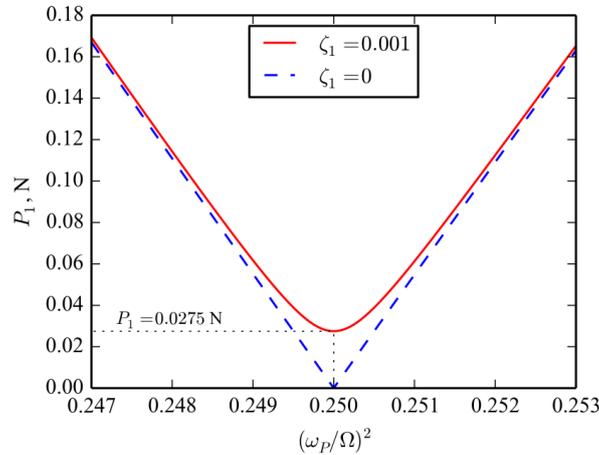
Figure 2: Transition curves for the beam subject to axial load, where the area above the curve is unstable. The transition curves are shown with no damping and for a damping ratio of 0.1%.

For the beam example, the parametric amplitude is found analytically to be 0.0275 N, which is well below the critical buckling load 2.91N. For lower level of damping, a small axial perturbation can make the system unstable. Numerical simulation using ode45 in Matlab also validates the analytical results. The time response for different amplitudes P_1 is shown in Figure 3 for $\varsigma_1 = 0.1\%$. The system is stable for amplitudes below 0.0275 N, Figure 3(a), and unstable above this amplitude, Figure 3(c). The response of the system exhibits a limit cycle oscillation, Figure 3 (b), when excited at 0.0275N. In this case, the system is on the transition curve.

(a) $P_1 = 0.01$ N (b) $P_1 = 0.0275$ N (c) $P_1 = 0.1$ N
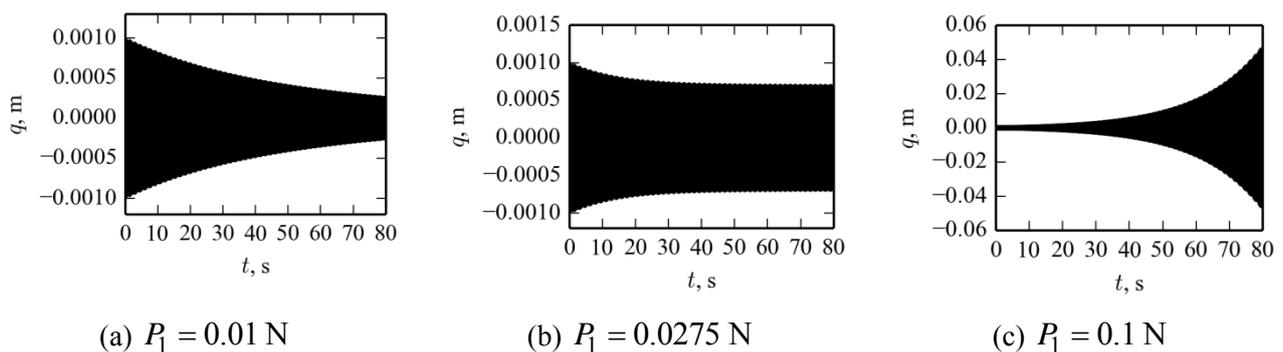
Figure 3: Time-response of the parametrically excited beam with 0.1% damping when excited at twice its natural frequency with different parametric amplitudes.

Experiments

In order to illustrate the described phenomenon a simple physical model was build. A steel cantilever beam was vertically mounted on the shaker table in a way that emulates a clamped

boundary condition. The shaker provided an axial excitation to the beam which was indirectly measured by placing an accelerometer on the shaker table. Transverse vibrations of the cantilever were recorded using a laser vibrometer.
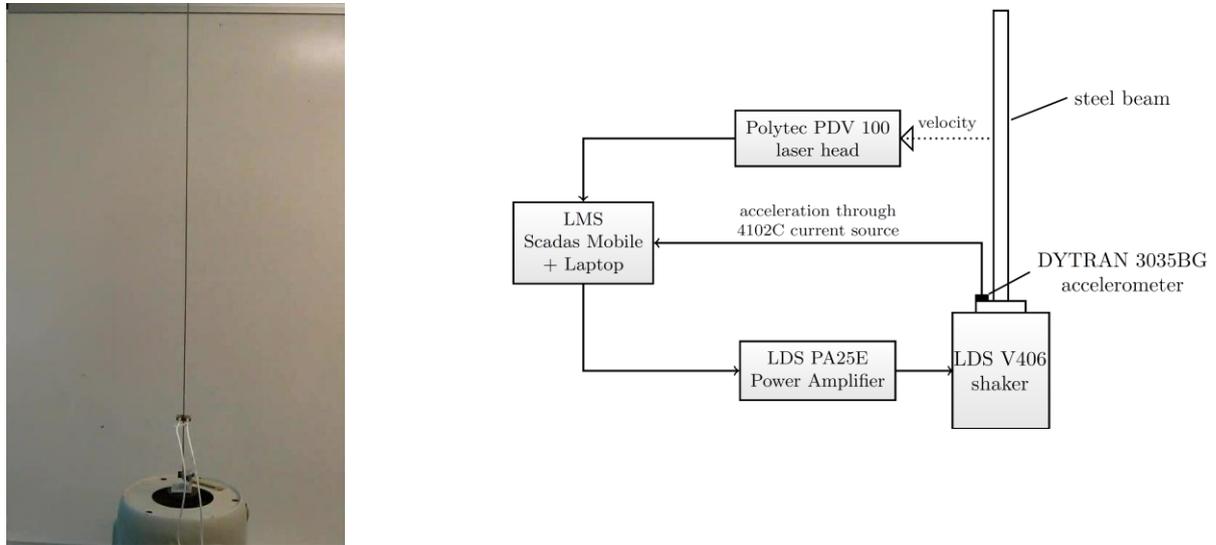
Figure 4: Schematic diagram of the experimental set-up

Firstly, the natural frequency associated with the first bending mode of the vertically mounted cantilever was determined via an impact hammer test. Based on this result, the parameters of the beam were identified (Table 1).

The transition curves were experimentally determined by sweeping through excitation frequencies close to the natural frequency associated with the first bending mode at different levels of the axial load and analysing the shape of the time histories of transverse vibrations. The system was qualified as stable if its response decayed exponentially with time and as unstable if the response was growing and experiencing a limit cycle oscillation. In order to ensure that all useful information is captured a very long time window was used (5 min) given the structure being very lightly damped.

The experimental results compared to theoretical transition curves as derived in the previous section are presented in Figure 5. Both are in a good agreement confirming the validity of the theoretical approach adopted. In Figure 6 we present chosen velocity response time histories and their Fourier transforms for the stable and the unstable case. The influence of nonlinearities is evident in the response at the unstable state since the response growth is limited.
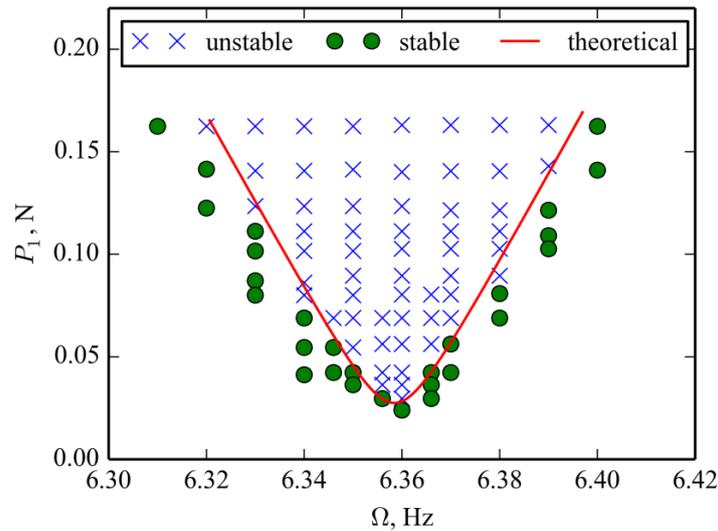
Figure 5: Transition curves – comparison between the experiment and the theory
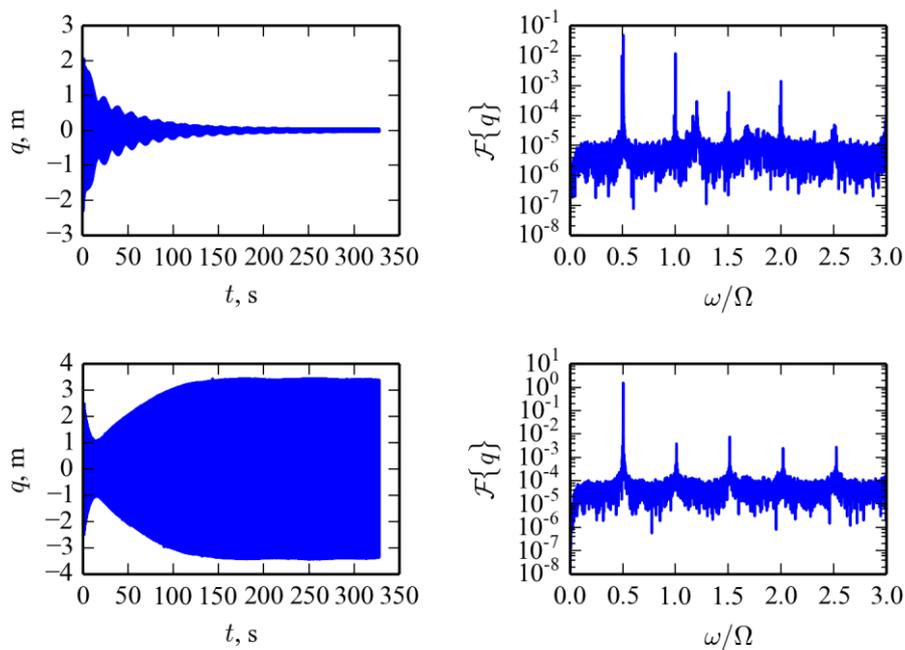
Figure 6: Response time histories and their Fourier transforms: (a) stable; (b) unstable

Active control

To stabilize the parametric instability as well as to increase the stability region, a velocity feedback and pole placement is considered. The system with control has the following form in state-space,

$$\dot{\mathbf{x}} = \begin{bmatrix} 0 & 1 \\ -\omega_P^2 \left(1 - \gamma \dfrac{P_1 \cos \Omega t}{P_{cr'}}\right) & -2\zeta_1 \omega_P \end{bmatrix} \mathbf{x} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u(t) = \mathbf{A}(t)\mathbf{x} + \mathbf{B}u(t) \tag{8}$$

where $u(t)$ is the control force and is a single-input.

Simulation: Velocity feedback

The amplitude of axial load is chosen to be $P_1 = 0.075\,\text{N}$ and $\Omega = 2\omega_P$. The open-loop response is 7unstable as shown with the blue-dashed line in Figure 5. The control law is considered to be:

$$u(t) = -h\dot{x}(t) \tag{9}$$

where h is the velocity feedback gain. The system is stabilized using the velocity feedback gain of $h = 1$ as shown with the red solid line in Figure .
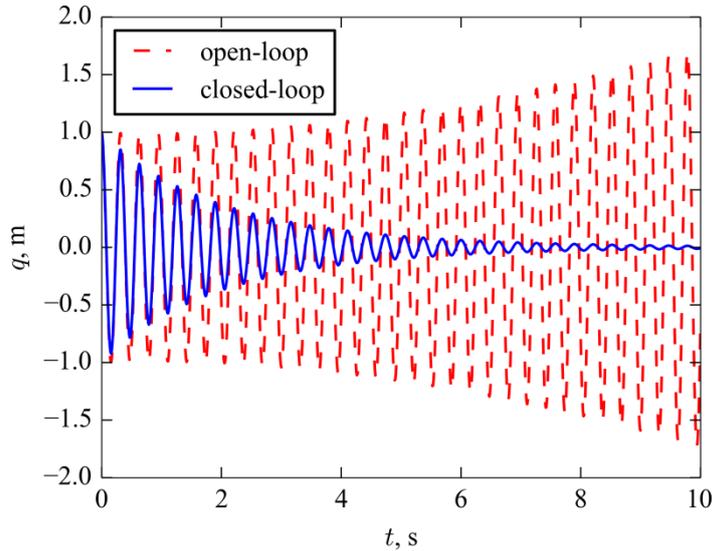
Figure 7 – Simulation: Time response - velocity feedback control; open-loop: grey dashed line, closed-loop: black solid line

Experiment: Velocity feedback

Experiment is carried out to implement the velocity feedback control and stabilize the system. The parameters of the axial load are chosen so that the open-loop system has parametric instability. A piezoelectric actuator is attached to the beam. The velocity of the beam at the other end is measured using LMS data acquisition and the velocity is fed back to the amplifier of the piezo with a feedback gain of 100.

The open-loop and closed-loop time responses are measured as shown in Figure 7. When the controller is switched on, the response decays due to the increase of the damping. This clearly demonstrates that using active control, the system is stabilized.
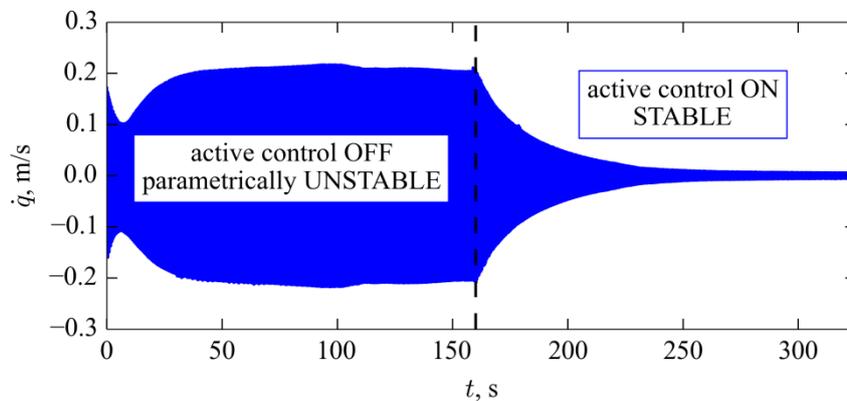
Figure 7 : Experiments: Active control of a parametrically excited beam

896

Exploitation: Energy Harvesting

Parametric resonance can be exploited in parametrically excited systems to harvest more vibration energy due to large amplitudes of vibrations compared to linear time-invariant systems. A piezoelectric is attached to the parametrically excited beam, as shown in Figure 8, which is shunted to a resistor for energy harvesting. The "electrical damping", used to harvest energy is assumed to be $0.1 \, \text{Ns/m}$.
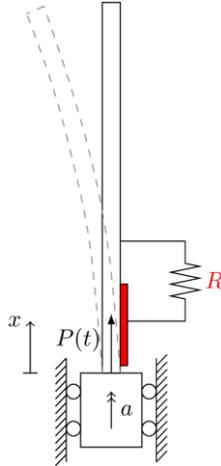
Figure 8 : Energy harvesting from a parametrically excited beam

Numerical simulation is carried out to investigate the amount of energy that can be harvested when the beam is subjected to the axial load with an amplitude, close to the parametric instability, for example $P_1 = 0.0272 \, \text{N}$ and $\Omega = 2\omega_1$. The harvested energy is calculated using Eq. (10) for the duration of 80 s with an initial displacement of 0.01m.

$$E = \int_0^{80} 0.1 \dot{q}^2(t) dt = 3.4 \, \text{J} \qquad (10)$$

Figure 9(a) shows that the harvested energy is maximum when the excitation frequency is almost twice the first bending frequency. In addition, the amount of energy increases when the excitation amplitude increases as shown in Figure 8(b) when $\Omega = 2\omega_1$. The amount of harvested energy from the parametrically excited beam with $P_1 = 0.0272 \, \text{N}$ is $E = 3.4 \, \text{J}$, while the harvested energy from the time-invariant system is $E = 1.78 \, \text{J}$. This dynamic behaviour can be exploited for the design of energy harvesters.

(a) (b)
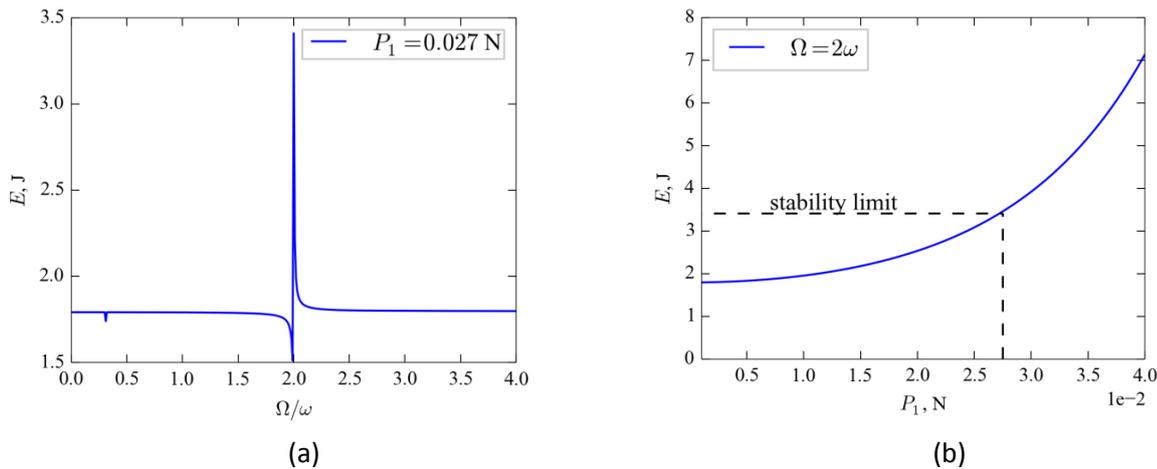
Figure 9 : Harvested energy-(a) for different frequency ratio when $P_1 = 0.0272$ N (b) for different parametric amplitude when $\Omega = 2\omega_1$

Acknowledgement

Conclusions

In this paper, control and exploitation of parametrically excited systems was presented. A beam subject to axial load was considered as a parametrically excited system. For an amplitude well below the critical buckling load, the beam experienced parametric resonance. Parametric instability was controlled using active vibration control. Velocity feedback control was considered to stabilize the system and assign stable poles. In addition, it was demonstrated that parametric excitation could be exploited to increase the amount of harvested energy. Practical implementation of the energy harvesting will be considered as part of future work.

References

Reynolds, P., Pavic, A. (2006) Vibration Performance of a Large Cantilever Grandstand during an International Football Match, *ASCE Journal of Performance of Constructed Facilities*, 20(3), 202-212.

Ahmed, T.M., Hudson, D.A. and Temarel, P.(2010) An investigation into parametric roll resonance in regular waves using a partly non-linear numerical model, *Ocean Engineering*, 37, 1307-1320.

Cartmell, M. (1990) *Introduction to Linear, Parametric and Nonlinear Vibrations*, Chapman and Hall, London.

Nayfeh, A. Mook, D. (1995) *Nonlinear Oscillations*, Wiley-VCH.

Ecker, H. (2010) Exploiting the use of parametric excitation, *10th International Conference RASD2010*, Southampton, July 2010.

Dohnal, F. Mace, B.R. (2008), Amplification of damping of a cantilever beam by parametric excitation, Proceedings-CD MOVIC 2008. *9th International Conference on Motion and Vibration Control (MOVIC 2008)*, September 2008, Munich, Germany, Institute of Applied Mechanics.

M.F. Daqaq, C. Stabler, Y. Qaroush, T. Seuaciuc-Osório, Investigation of power harvesting via parametric excitations, *Journal of Intelligent Material Systems and Structures*, 20, 545-557.

Abdelkefi, Nayfeh, N and Hajj,M. (2012) Global nonlinear distributed-parameter model of parametrically excited piezoelectric energy harvesters, *Nonlinear Dynamics*, 67, 1147-1160.

An industrial application of

thermal convection analysis

*H. Kanayama

Department of Mathematical and Physical Sciences, Faculty of Science,
Japan Women's University,
2-8-1, Mejirodai, Bunkyo-ku, Tokyo 112-8681, Japan

*Corresponding author: kanayamah@fc.jwu.ac.jp

Abstract

A coupling analysis of thermal convection problems is performed in this work. By approximating the material derivative along the trajectory of fluid particle, the characteristic curve (CC) method can be considered. The most attractive advantage of this method is the symmetry of the linear system, which enables some classic symmetric linear iterative solvers, like the conjugate gradient (CG) method or the minimal residual method (MINRES), to be used to solve the interface problem of the domain decomposition system. Applications to industrial problems are demonstrated to show the effectiveness of our approach.

Keywords: The characteristic curve method, Symmetric linear solvers, Finite element method, Domain decomposition method, Vending machines

1   Introduction

It is well known that the Galerkin approach for Navier-Stokes equations faces the problem caused by the nonlinear convective term, which yields the difficulty to obtain a numerical solution, and that the difficulty even increases with the Reynolds number. A lot of researchers contributed to solve this; the stiffness matrix was generally non-symmetric and some product-type methods[1] such as GPBi-CG, Bi-CGSTAB, Bi-CGSTAB2 were utilized as the iterative solver for non-symmetric linear systems, like ADV_sFlow 0.5, which was one of our previous works.[2]-[4] In this research, based on the approximation of the material derivative along the trajectory of fluid particle, a characteristic curve (CC) method[5],[6] is employed to approximate the material derivate terms. The method is natural from the viewpoint of the simulation of physical phenomena; it is also advantageous as it renders the matrix of the linear system symmetric. When solving the interface problem of the Schur complement system, which is generated by the domain decomposition method (DDM), the symmetry enables the conjugate gradient (CG) method or the minimal residual method (MINRES) to be employed instead of product-type iteration solvers. Despite the fact that there is no sufficient theoretical proof to assure this convergence currently, the CG or MINRES method does show good convergence even in cooperation with several kinds of preconditioners[7] including the balancing domain decomposition (BDD) preconditioner[2],[5] in this research.

For thermal convection problems, some research[8] has been shown in references. A number of researches[9]-[12] about the CC method were done in the case of a single processor. However, rare research is done on the implementation of this scheme in the domain decomposition system, in which better computation results can be expected as the computation capability is extended. The purpose of our work is to apply the CC method to the Navier-Stokes equations and the convection-diffusion equation, and to enable the coupling analysis of these two kinds of applications.

A new parallel coupling thermal convection solver[6] has been developed. Based on a CC method, the scheme can provide solvability for non-stationary thermal convection problems. The new solver

can reduce memory consumption compared with solvers of product-types. The computation speed is also improved, as is expected. The reliability and accuracy of numerical results have successfully been validated by comparing with the exact solution.[6] Comparisons of our numerical results with results of other recognized solvers or available benchmarks also convince us that the application of the characteristic curve method to thermal convection problems has been a success. Accompanied with several new features, the new solver is speedy and worthy to be expected. As one member of the ADVENTURE system (http://adventure.sys.t.u-tokyo.ac.jp/), it will be published as an updated version of ADV_sFlow on the homepage after the in-house testing.

Finally, some practical applications are demonstrated for a vending machine[13]. Non-stationary thermal convection problems are solved to show the effectiveness of the above approach.

In recent years, energy conservation has become an important topic in Japan. One focus of current research is the use of numerical analysis techniques to control cooling and heating systems in vending machines to improve their efficiencies and reduce their electric power consumptions. In this study, we report results of analyses using the ADVENTURE_sFlow parallel solver to study problems of thermal convection in mechanical components inside a vending machine.

The rest of this paper is organized into several sections. In Section 2, formulations related to thermal convection problems are introduced; the characteristic curve method, as well as the related finite element scheme is also demonstrated in this section. Models and various settings including boundary settings are described in Section 3 and Section 4, respectively. Section 5 shows numerical results obtained by using various models and boundary settings. Conclusions that can be drawn from current results are presented in Section 6.

2   Formulations

2.1 The thermal convection problem

Let Ω be a three-dimensional polyhedral domain with the boundary $\partial\Omega$. Using the Boussinesq approximation to couple the Navier–Stokes equations to the convection–diffusion equation, the conservation equations of momentum and mass are as follows;

$$\begin{cases} \dfrac{\partial u}{\partial t} + (u \cdot \nabla)u - 2v\nabla \cdot D(u) + \nabla p = g\beta(T_r - T) & in\ \Omega \times (0, \bar{t}), \\ \nabla \cdot u = 0 & in\ \Omega \times (0, \bar{t}). \end{cases} \tag{1}$$

Also, the thermal convection equation is given as

$$\frac{\partial T}{\partial t} + u \cdot \nabla T - a\Delta T = S \quad in\ \Omega \times (0, \bar{t}). \tag{2}$$

In Eqs. (1) and (2), $u = (u_1, u_2, u_3)^T$ is the velocity [m/s]; t is time [s]; v is the kinematic viscosity coefficient [m^2/s]; p is the gauge pressure normalized by the density [m^2/s^2] (hereafter p is shortly called pressure); $g = (g_1, g_2, g_3)^T$ is the gravity [m/s^2]; β is the thermal expansion coefficient [1/K]; T is the temperature [K]; a is the thermal diffusion coefficient [m^2/s]; S is the source term [K/s]; and D_{ij} is the rate of strain tensor [1/s] defined by

$$D_{ij}(u) \equiv \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) \quad i, j = 1, 2, 3.$$

Though initial and boundary settings are described in details for different problems in Section 4, they are here written as follows;

$$u = \hat{u} \qquad\qquad on\ \Gamma_u \times (0, \bar{t}), \tag{3}$$

$$T = \hat{T} \qquad\qquad on\ \Gamma_T \times (0, \bar{t}), \tag{4}$$

$$a\frac{\partial T}{\partial n} = 0 \qquad\qquad \text{on } (\partial\Omega - \Gamma_T)\times(0,\bar{t}),\qquad\qquad (5)$$

$$\sum_{j=1}^{3}\sigma_{ij}n_j = 0 \qquad\qquad \text{on } (\partial\Omega - \Gamma_u)\times(0,\bar{t}),\qquad\qquad (6)$$

$$u = u^0,\, T = T^0 \qquad\qquad \text{in } \Omega \text{ at } t = 0, \qquad\qquad (7)$$

where
$$\sigma_{ij} = -p\delta_{ij} + 2vD_{ij}(u) \qquad\qquad i,j = 1,2,3$$

with the Kronecker delta δ_{ij} and n is the unit normal vector. $\bar{t}$ is the total time [s]; u^0 is the initial velocity [m/s]; T^0 is the initial temperature [K]; $\hat{u}$ is the boundary velocity [m/s]; and $\hat{T}$ is the boundary temperature [K]. Γ_u and Γ_T are the velocity specified boundary and the temperature specified boundary, respectively.

As the weak form, the following system is considered;

$$\left(\frac{\partial u}{\partial t}, v\right) + \left((u\cdot\nabla)u, v\right) + \left(2vD(u), D(v)\right) - \left(p, \nabla\cdot v\right) = (g\beta(T_r - T), v) \quad \text{for } v\in V, \qquad (8)$$

$$-(\nabla\cdot u, q) = 0 \qquad\qquad \text{for } q\in Q, \qquad\qquad (9)$$

$$\left(\frac{\partial T}{\partial t}, \theta\right) + (u\cdot\nabla T, \theta) + (a\nabla T, \nabla\theta) = (S, \theta) \qquad\qquad \text{for } \theta\in\Theta. \qquad (10)$$

Here, $L^2(\Omega)$ denotes the space of square summable functions in, Ω and $H^1(\Omega)$ is the space of functions in $L^2(\Omega)$ with derivatives up to the first order.

$$V(g_1) \equiv \left\{v\in(H^1(\Omega))^3; v = g_1 \text{ on } \Gamma_u\right\}, V \equiv V(0), \qquad\qquad (11)$$

$$Q \equiv \left\{q\in L^2(\Omega)\right\}, \qquad\qquad (12)$$

$$\Theta(g_2) \equiv \left\{\theta\in H^1(\Omega); \theta = g_2 \text{ on } \Gamma_T\right\}, \Theta \equiv \Theta(0), \qquad\qquad (13)$$

where $(\cdot,\cdot)$ denotes the L^2- inner product over Ω.

2.2 The characteristic curve method
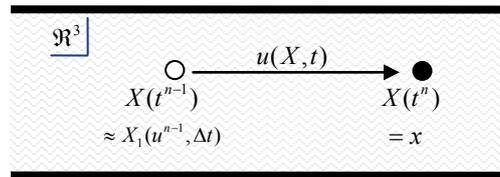
Figure1. Trajectory of a fluid particle

Let $X(t)$ be a function of position and Δt be the time increment. With the definitions of $t^n \equiv n\Delta t, N_T \equiv [\bar{t}/\Delta t]$, it is assumed that $X(t^n) = x$ and the fluid field velocity $u(X,t) = X'(t)$. A fluid particle's position at t^{n-1} can be approximated by

$$X(t^{n-1}) = X_1(u^{n-1}, \Delta t) + O(\Delta t^2), \qquad\qquad (14)$$

where $X_1(u^{n-1}, \Delta t)(x) \equiv x - u^{n-1}(x)\Delta t$ is an approximation of the position function $X(t)$ ($t=t^{n-1}$) used by the Euler scheme[12] (see Fig. 1). With this approximation, the material derivative term can be written as

$$\frac{\partial u}{\partial t} + (u\cdot\nabla)u = \frac{u^n - u^{n-1}\circ X_1(u^{n-1}, \Delta t)}{\Delta t}(x) + O(\Delta t). \qquad\qquad (15)$$

Here, the notation $\circ$ designates the composition of functions, u^n denotes the velocity at t^n, and u^{n-1} is a known value, which denotes the velocity at t^{n-1}.

2.3 Finite element scheme

Let Θ_h and $V_h \times Q_h$ be the piecewise linear finite element spaces for test functions of the convection-diffusion equation and Navier-Stokes equations, respectively, and $\Im_h$ represents a triangulation (h is the diameter). With the following bilinear forms;

$$a_0(u,v) \equiv \int_\Omega 2\nu D(u) : D(v)dx, \;\; b(v,q) \equiv -\int_\Omega q\nabla \cdot vdx, \tag{16}$$

the scheme for coupling analysis can be written as

for $n = 1, \cdots N_T$, $\forall \theta_h \in \Theta_h$ and $\forall(v_h, q_h) \in V_h \times Q_h$,

   STEP 1: Find T_h^n by

$$\left(\frac{T_h^n - T_h^{n-1} \circ X_1(u_h^{n-1}, \Delta t)}{\Delta t}, \theta_h \right) + \left(a\nabla T_h^n, \nabla \theta_h \right) = (S, \theta_h). \tag{17.a}$$

   STEP 2: Find (u_h^n, p_h^n) by

$$\left(\frac{u_h^n - u_h^{n-1} \circ X_1(u_h^{n-1}, \Delta t)}{\Delta t}, v_h \right) + a_0(u_h^n, v_h) + b(v_h, p_h^n) + b(u_h^n, q_h)$$

$$+ \sum_{K \in \Im_h} \tau_K (\nabla p_h^n, -\nabla q_h)_K = \left(g\beta(T_r - T_h^n), v_h \right). \tag{17.b}$$

Here, τ_K is a parameter of the stabilization term, which is used to smooth away the potential oscillation caused by P1/P1 elements. $(\cdot, \cdot)_K$ denotes the element wise inner product. In each non-stationary loop, STEP 1 and STEP 2 are performed. It keeps running until the maximum non-stationary loop number is reached.

3   Computational Models

In this study, we consider a simplified model for a vending machine depicted in from Fig. 2 to Fig. 4. Here, we have omitted the portions containing the cooling and heating systems and have retained only the columnar racks of canned beverages, the thermal barrier walls surrounding these racks, and the tilted ramp along which the cans slide inside the machine. The left panel of Fig. 2 depicts a side view of the model, while the right panel depicts a front view of the cross-sectional area indicated by the red line in the left panel. Below, we refer to components 1, 2, and 3 as x, y, and z, respectively.

The model used in the present analysis is a trapezoidal box with a height of 0.9 [m], a depth of 0.5676 [m], and a width of 0.1461 [m]. Air flows in from the slanted lower surface of the box and flows out through the upper surface. The cans that constitute the merchandise are 0.064 [m] in diameter and 0.114 [m] high. The separation between the cans and the thermal barrier walls is 0.01 [m] in the x direction. The cans and the intermediate board are separated from the thermal barrier walls by 0.014 [m] in the y direction. However, the number of cans depends on each model, and Fig. 2 and Fig. 4 are representative examples. Fig. 4(a), Fig. 4(b), and Fig.4(c) are models that have different number of cans.
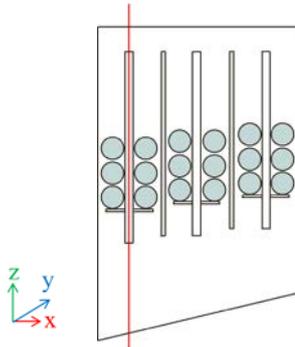
Figure 2. A side view of the model
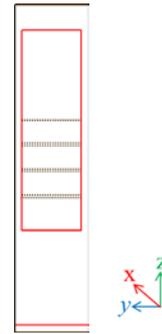
Figure 3. A front view of the model
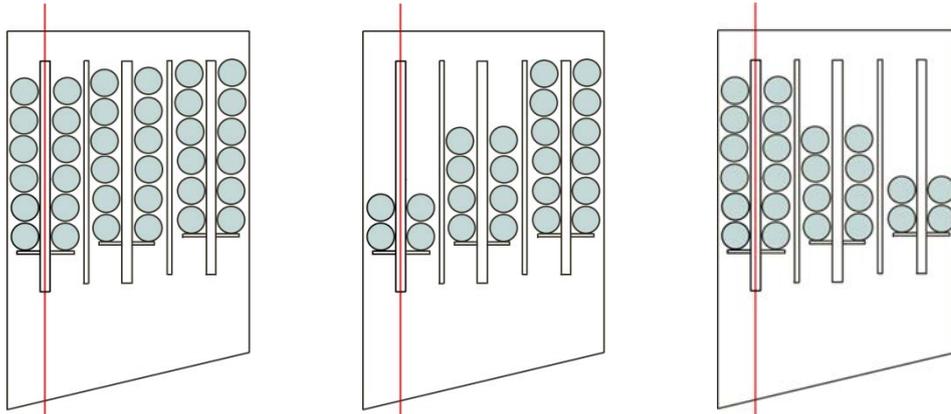
Figure 4. Side views of models which have different number of cans (a)full; (b)stairs; (c)stairs

4   Computational Conditions

4.1 Material properties

Table 1 lists the material properties used in the present analysis. In the winter case (Win.), an eddy viscosity constant is used for the first trial.

Table 1. Material properties

S[K/s]	v[m^2/s]	T_r [K]	β[1/K]	a [m^2/s]
0	1.05×10^{-3} (Win.) 1.583×10^{-5} (Sum.)	300.15(Win.) 278.15(Sum.)	0. 0034	2. 207×10^{-5}

4.2 Initial conditions and boundary conditions

The initial conditions are a uniform velocity of 0 [m/s] and a uniform temperature of 5[°C]. The boundary conditions have two types, which are Winter Condition 1 and Summer Condition 2. In Winter Condition 1, the airflow entry surface (the lower surface) is set as follows;

$$u_x = u_y = 0, u_z = 0. 1 \text{ [m/s]},$$

$$T = 55 + 273.15 \text{ [K]}.$$

At the airflow exit,

$$\sum_{j=1}^{3} \sigma_{ij} n_j = 0 \ [m^2/s^2],$$

$$a\frac{\partial T}{\partial n} = 0 \ [\text{m·K/s}].$$

At wall surfaces, the cans and the partitions within the machine,

$$u_x = u_y = u_z = 0 \ [\text{m/s}],$$

$$a\frac{\partial T}{\partial n} = 0 \ [\text{m·K/s}].$$

On the other hand, Summer Condition 2 are different from Winter Condition 1 in terms of temperature profile of entering air flow;

$$T = -2\sim4.5 \ [^\circ\text{C}].$$

In Summer Condition 2, we vary the temperature of the air flowing into the machine over the range –2 to 4.5 [°C]. This is because, in actual vending machines, the cooling system is calibrated over time to reduce power consumption. The temperature profile used in this analysis is plotted in Fig. 5.
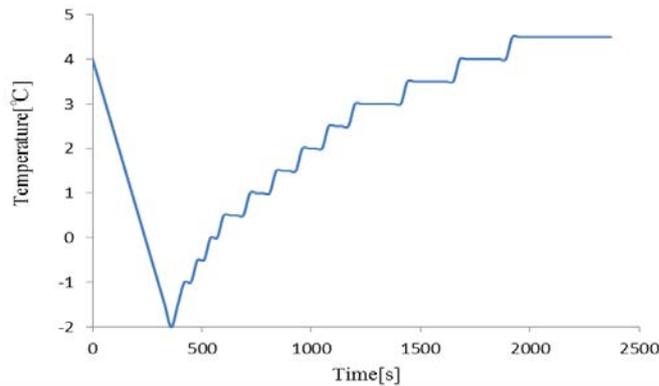
Figure 5. Temperature profile of entering airflow

4.3 Other parameters

Numbers of elements, numbers of nodal points and degrees of freedom are 598,916, 148,757 and 743, 785, respectively in the winter case. On the hand, they are, respectively, 744,585, 176,105 and 1,114,655 in the summer case. The time increment, numbers of time steps and the total time are 0.1 [s], 2,000 and 200 [s] in the winter case. They are, respectively, 0.1 [s], 24,000 and 2,400 [s] in the summer case. As the solver for (17.b), the CG method with the BDD preconditioner was used for the interface problem in the summer case, while the CG method with the Jacobi preconditioner was used for (17.a) and for (17.b) in the winter case. Using Core i7 920 (2.66 [GHz]) with 4 cores, it took about 5 hours by 5 PCs in the winter case and about 88 hours by 3 PCs in the summer case.

5   Results

5.1 Results of Winter Condition 1

Figures 6(a) and 6(b) depict the temperature distributions after 10 and 100 [s] predicted by the model in Fig.2. Figures 7(a) and 7(b) plot the corresponding velocity vectors.
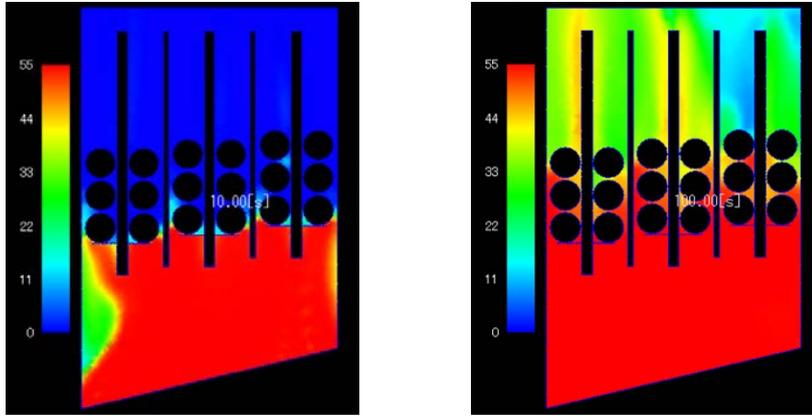
(a) after 10 [s];   (b) after 100 [s];

Figure 6. Temperature distribution
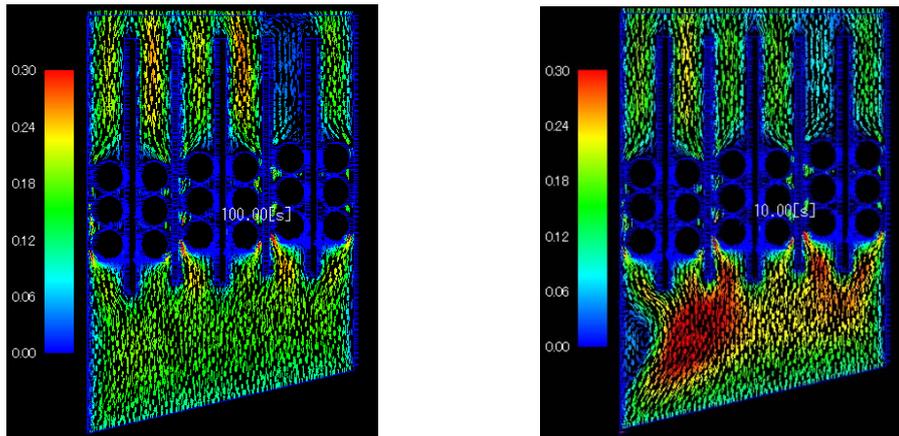
(a) after 10 [s];   (b) after 100 [s];

Figure 7. Velocity vectors

Figure 8 plots the temporal evolution of the temperature at the three points indicated in Fig. 8(a). Table 2 lists the coordinates of these three points.
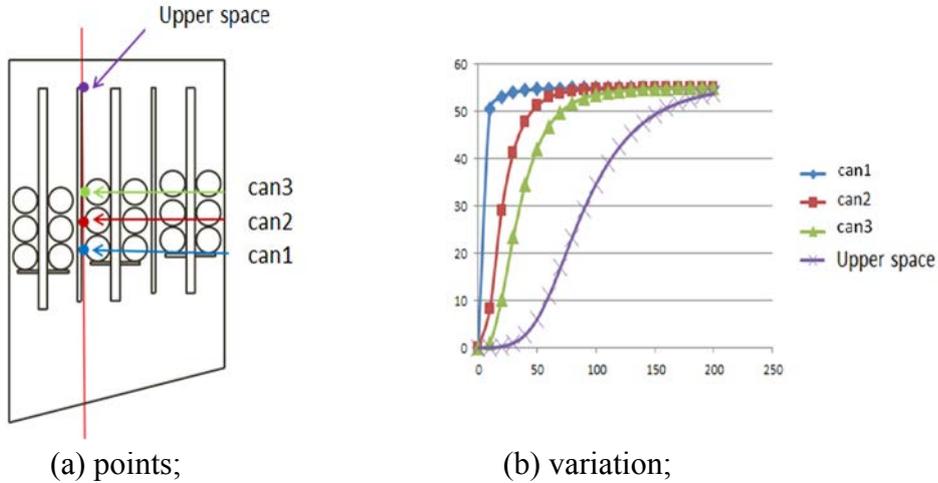
(a) points;   (b) variation;
Figure 8. Temporal temperature variation

Table 2. Coordinates of points plotted in Figure 8

	x[m]	y[m]	z[m]
Can 1	0.1960	0.0738	0.4158
Can 2	0.1960	0.0738	0.4801
Can 3	0.1960	0.0738	0.5397
Upper space	0.1960	0.0738	0.8580

The computational results reveal that the temperature in the lower portion of the machine begins to rise first with the temperature in the upper portion of the machine gradually following. The flow achieves steady-state conditions at a surprisingly early time of 25 [s] because of the eddy viscosity constant, for which further study is required.

5.2 Results of Summer Condition 2
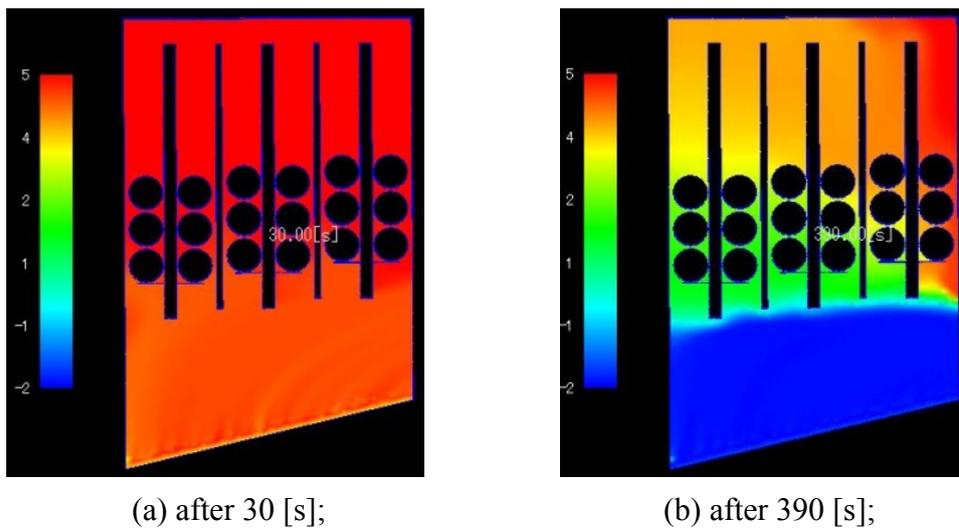
(a) after 30 [s];   (b) after 390 [s];
Figure 9. Temperature distribution

Table 3. Coordinates of three points for the temporal temperature variation

	x [m]	y [m]	z [m]
Point 1	0.136	0.073	0.321
Point 2	0.368	0.073	0.341
Point 3	0.524	0.073	0.361
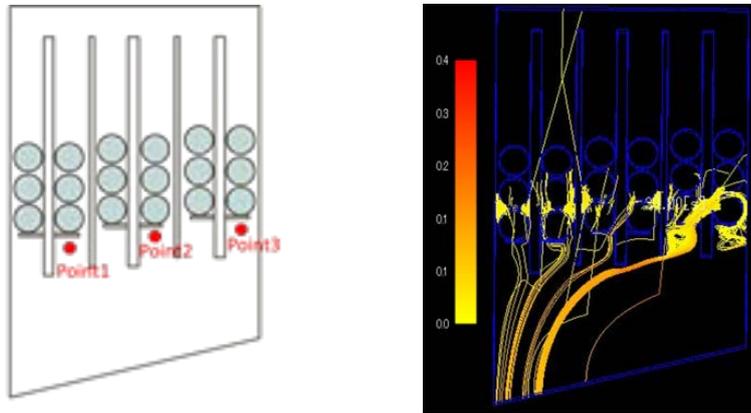
    (a) monitored;   (b) streamline diagram;

Figure 10. Three points for the temporal temperature variation
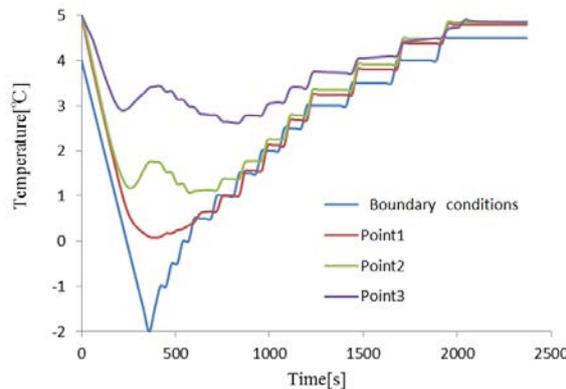
Figure 11. Temporal temperature variation at the three points depicted in Figure10(a)

Figure 11 indicates that the temperature variation inside the vending machine follows that of the airflow into the machine. The airflow appears to be poor only in the region on the right side of the machine. To investigate this, we consider the streamline diagram of Fig. 10 (b), which indicates a counter flow along the surfaces of the cans in this region. The question of why such a counter flow arises entails many uncertainties and requires further study.

5.3 Summer case results of another models
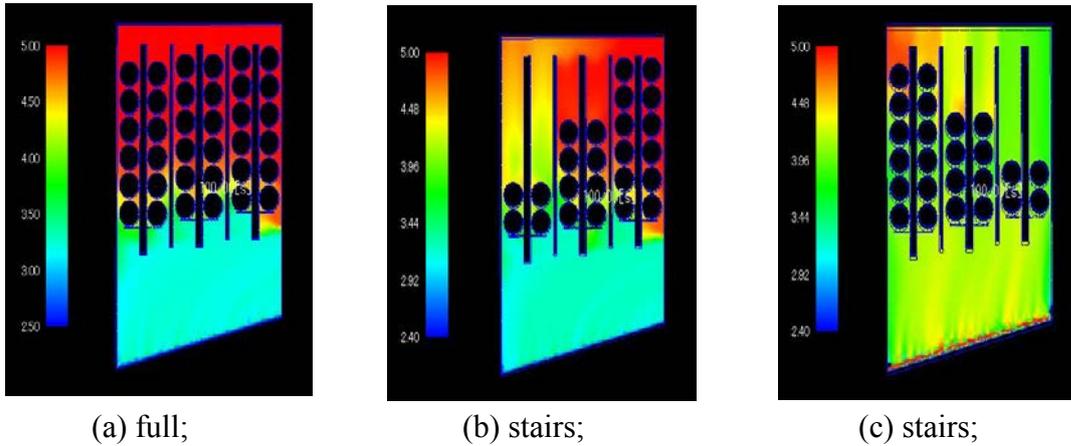
(a) full; (b) stairs; (c) stairs;
Figure 12. Temperature distribution after 100 [s]
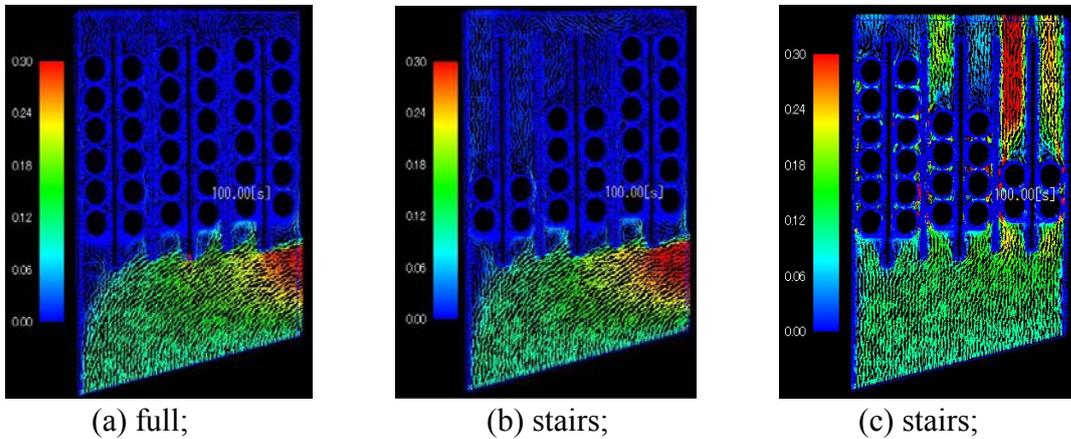
(a) full; (b) stairs; (c) stairs;
Figure 13. The temporal velocity variation of stream diagram after 100 [s]

Here, we consider another summer cases. Namely, different number of cans are considered for each column. Case (a) is a fully occupied case and Case (b) and Case (c) consider stairs. In all cases, many cans produce high temperature because cans become obstacles of the air flow. It is specially noted that Case (b) relatively shows high temperature, compared with Case (c).

6 Conclusions

A coupling analysis of thermal convection problems is performed in this work. By approximating the material derivative along the trajectory of fluid particle, the characteristic curve (CC) method can be considered. The most attractive advantage of this method is the symmetry of the linear system, which enables some classic symmetric linear iterative solvers, like the conjugate gradient (CG) method or the minimal residual method (MINRES), to be used to solve the interface problem of the domain decomposition system. Applications to industrial problems are demonstrated to show the effectiveness of our approach.

References

[1] Vorst, H.A. van der. Iterative Krylov method for large linear systems. Cambridge University Press, (2003).

[2] Kanayama, H., Ogino, M., Takesue, N. and Mukaddes, A. M. M. Finite element analysis for stationary incompressible viscous flows using balancing domain decomposition. Theoretical and Applied Mechanics (2005) Vol. 54:211-219.

[3] Kanayama, H., Komori, K., Sato D. and Sakuragi, O. Development of a thermal convection solver with hierarchical domain decomposition method. Proceedings of Computational Engineering Conference of JSCES (2008) Vol. 13, No. 2: 721-722. (in Japanese)

[4] Kanayama, H., Tsukikawa, H. and Ismail, I. Simulation of hydrogen dispersion by the domain decomposition method. *Japan Journal of Industrial and Applied Mathematics* (2011) Vol.28, No.1:43-53.

[5] Yao, Q., Kanayama, H., Notsu, H. and Ogino, M. Balancing domain decomposition for non-stationary incompressible flow problems using a characteristic-curve method. *Journal of Computational Science and Technology* (2010) Vol. 4, No. 2:121-135.

[6] Yao, Q. and Kanayama, H. A coupling analysis of thermal convection problems based on a characteristic curve method. *Theoretical and Applied Mechanics Japan* (2011) **59**:257-264.

[7] Yao, Q., Kanayama, H., Notsu, H. and Ogino, M. Incomplete balancing domain decomposition for large scale 3-d non-stationary incompressible flow problems. *IOP Conference Series: Materials Science and Engineering* (2010) Vol. 10, No. 1: 012029.

[8] Tabata, M. Finite element approximation to infinite Prandtl number Boussinesq equations with temperature dependent coefficients - thermal convection problems in a spherical shell. *Future Generation Computer Systems* (2006) Vol. 22:521-531.

[9] Tabata, M. and Fujima, S. Robustness of a characteristic finite element scheme of second order in time increment. *Computational Fluid Dynamics 2004(eds. Groth, C. and Zingg, D. W.)* Springer-Verlag (2006):177-182.

[10] Notsu, H. Numerical computations of cavity flow problems by a pressure stabilized characteristic-curve finite element scheme. *Transactions of the Japan Society for Computational Engineering and Science* (2008) Vol. 2008, No. 20080032, ONLINE ISSN : 1347-8826.

[11] Notsu, H. and Tabata, M. A combined finite element scheme with a pressure stabilization and a characteristic-curve method for the Navier-Stokes equations. *Transactions of the Japan Society for Industrial and Applied Mathematics* (2008) Vol. 18, No. 3:427-445.

[12] Pironneau, O. *Finite element methods for fluids*. John Wiley & Sons (1989).

[13] http://www.fujielectric.co.jp/frs/story/index.html

Relations between the Poynting and axial force-twist effects

Dong Wang, M.S. Wu*

School of Mechanical and Aerospace Engineering, Nanyang Technological University, 50 Nanyang Avenue, Singapore 639798

*Corresponding author: mmswu@ntu.edu.sg

Abstract

The relations between the Poynting effect, in which a cylinder elongates or contracts axially under torsion, and the axial force-twist effect, in which the twist of a torsionally loaded cylinder is affected by the axial loading, are investigated using second-order elasticity for an elastic homogeneous cylinder. The explicit expressions for the two effects and their relations are presented. The relations show that under tension: (a) negative Poynting effect implies negative axial force-twist effect, (b) positive axial force-twist effect implies positive Poynting effect, whereas (c) the converse statements are not true. Further results show that (a) the Poisson ratio captures the difference between the two effects, and (b) reduced elastic coefficients, which uniquely characterize the effects, lead to universal relations between the effects and the applied loading. Both effects also exhibit a strong inverse power law dependence on the radius.

Keywords: Axial force-twist effect, Poynting effect, torsion-axial loading, second-order elasticity

Introduction

Soft materials may exhibit complex nonlinear behavior such as the Poynting effect, in which a cylinder elongates or contracts axially under torsion. Poynting (1909) experimentally found that some metals exhibited the positive effect, i.e., they elongated axially under torsion. Recently, Janmey et al. (2007) found that networks of semiflexible biopolymers such as actin, collagen, fibrin and neurofilaments, exhibited the negative Poynting effect.

Wang and Wu (2014) showed that in contrast to the Poynting effect, an axial force-twist effect may also exist. It refers to their theoretical result that the twist of a cylinder under combined torsion and axial loading can be affected by the axial loading. The axial force-twist effect can also be positive or negative. The former means that both the twists produced by the axial loading and torsion are in the same direction, while the latter means that the twists produced by them are in the opposite directions. Though Wang and Wu (2014) presented the solutions for the Poynting and axial force-twist effect, the relations between them were not investigated.

This paper focuses on these relations, from which some fundamental conclusions can be drawn. The dependence of the two effects on the linear and nonlinear elastic constants is also studied. The organization of the paper is as follows. The derivation of the relations is first presented, followed by numerical results, a further discussion, and a set of conclusions.

Relations between the Poynting effect and the axial force-twist effect

Figure 1 shows a cylinder of length L and radius R under combined axial loading P and torsion T. Here P represents either a tensile or compressive stress. The materials are nonlinear elastic, isotropic and homogeneous. The initial coordinates of a particle of the cylinder are chosen as (r, θ, z). The strain energy density of Murnaghan (1951) is adopted, i.e.:

$$W = \frac{\lambda + 2\mu}{2} J_1^2 - 2\mu J_2 + \frac{l+2m}{3} J_1^3 - 2mJ_1 J_2 + nJ_3, \qquad (1)$$

where λ and μ are the second-order and l, m, n the third-order elastic constants, respectively, and J_1, J_2, and J_3 are the strain invariants of the Lagrangian strain **E**. The detailed solutions of the stress and displacement fields are given in Wang and Wu (2014). For the purpose of deriving the relations between the effects, the results on the axial and circumferential displacements from the earlier paper are given below.

The axial displacement u_z under pure torsion loading can be written as:

$$u_z = Dz, \qquad (2)$$

where D is the Poynting effect coefficient given by:

Figure 1. A homogeneous elastic cylinder with radius R and length L under combined torsion T and axial loading P.

$$D = -C_D \frac{T^2}{4\pi^2 R^6}, \qquad C_D = \frac{n\lambda + 4\mu m + 4\lambda\mu + 8\mu^2}{\mu^3 (3\lambda + 2\mu)}. \qquad (3)$$

Note that a change in the direction of T does not change the sign of D. The parameter C_D is a reduced coefficient of the four elastic constants. It uniquely characterizes the quadratic relation between the Poynting effect and T. If a modified Poynting effect coefficient $\bar{D} = D / C_D$ is defined, then a universal relation between $\bar{D}$ and T can be obtained:

$$\bar{D} = -\frac{T^2}{4\pi^2 R^6}. \qquad (4)$$

Furthermore, the circumferential displacement under combined axial loading P and torsion T is:

$$u_\theta = u_\theta^L + u_\theta^{NL}, \qquad (5)$$

where u_θ^L represents the linear twist due to torsion T:

$$u_\theta^L = \frac{2T}{\pi R^4 \mu} rz, \qquad (6)$$

and u_θ^{NL} represents the nonlinear twist associated with the axial force-twist effect:

$$u_\theta^{NL} = -\frac{PT(n\lambda + 4\mu m + 6\lambda\mu + 8\mu^2)}{2\pi R^4 \mu^3 (3\lambda + 2\mu)} rz. \qquad (7)$$

The axial force-twist effect coefficient can be defined as:

$$H = \frac{u_\theta^{NL}}{u_\theta^L} = -C_H \frac{P}{4}, \qquad C_H = \frac{n\lambda + 4\mu m + 6\lambda\mu + 8\mu^2}{\mu^2 (3\lambda + 2\mu)}. \qquad (8)$$

C_H is a reduced coefficient which characterizes the relation between H and P. It is similar in form to C_D. By defining the modified axial force-twist effect coefficient $\bar{H} = H / C_H$, a universal linear relation between $\bar{H}$ and P can be obtained:

$$\bar{H} = -\frac{P}{4}. \tag{9}$$

It can be seen from Eq. (7) that the axial force-twist effect only exists under combined axial loading and torsion, i.e., $P \neq 0$ and $T \neq 0$. Eq. (8) implies that if H is positive, u_θ^{NL} has the same direction as u_θ^L and the axial force twist effect is positive; otherwise, it's negative. Because H depends on P and not T, two further observations can be made from Eq. (8):

(1) Change of the direction of T does not change the sign of H.
(2) Change of the sign of P changes the sign of H.

Eqs. (3) and (8) show that materials with different elastic constants can have the same Poynting effect or the axial force-twist effect, provided the reduced coefficients of these materials are the same. Another observation of Eq. (3) is that for a particular μ, if m and n are chosen in a way that makes $(n + 4\mu)/3 = (4m + 8\mu)/2$, or $8\mu + 6m - n = 0$, then λ has no influence on the Poynting effect. A similar conclusion can be made for H on the basis of Eq. (8). If m and n are chosen such that $(n + 6\mu)/3 = (4m + 8\mu)/2$, or $6\mu + 6m - n = 0$, then λ has no influence on the axial force-twist effect.

The relation between H and D in dimensionless form can be obtained easily from Eqs. (3) and (8):

$$\frac{H}{P/4\mu} = \frac{D}{T^2/4\pi^2 R^6 \mu^2} - \frac{2\lambda}{3\lambda + 2\mu}. \tag{10}$$

Since $\nu = \lambda/(2\lambda + 2\mu)$, the above equation can be rewritten as:

$$\frac{H}{P/4\mu} = \frac{D}{T^2/4\pi^2 R^6 \mu^2} - \frac{2\nu}{1+\nu}. \tag{11}$$

The term on the left-hand side represents the axial force-twist effect coefficient normalized by the axial loading, while the first term on the right-hand side represents the Poynting effect normalized by the torsion. An explicit relationship between the axial force-twist effect and the Poynting effect is thus established.

Since ν is positive generally, several conclusions can be drawn from Eq. (11), assuming that the axial loading P is tensile:

(a) If $D < 0$, then necessarily $H < 0$,

(b) If $H > 0$, then necessarily $D > 0$.

(c) If H and D have different signs, then necessarily $H < 0$ and $D > 0$.

It should be emphasized that the converses of (a) and (b) are not true, i.e., $H < 0$ does not necessarily imply $D < 0$, and $D > 0$ does not necessarily imply $H > 0$. A further observation is that the sign of H will change if the sign of P changes. Thus for the case of compressive axial loading, the above three conclusions should be changed to:

(d) If $D < 0$, then necessarily $H > 0$,

(e) If $H < 0$, then necessarily $D > 0$.

(f) If H and D have the same sign, then necessarily $H > 0$ and $D > 0$.

The Poisson ratio plays a key role since the difference between the normalized H and the normalized D is the term $2\nu/(1+\nu)$. This difference reaches its maximum when $\nu = 0.5$, i.e., the material is incompressible.

Finally, the size dependence of the Poynting effect can be judged from Eq. (3) to be inversely proportional to the sixth power of the cylinder radius. For the axial force-twist effect, Eq. (7) shows that the maximum circumferential displacement ($r = R$) is inversely proportional to the third power of the cylinder radius. Hence, the Poynting effect is relatively more important than the axial force-twist effect for small cylinders, and the reverse holds for large cylinders.

Numerical results

This section focuses on the influence of the elastic constants on the Poynting effect and the axial force-twist effect. The elastic constants of the soft materials were adapted from Wang and Wu (2013, 2014) for poly(acrylic acid) (PAA) gels and capillary muscles, respectively, and Catheline et al. (2003) for an agar-gelatin. The geometry of the cylinder is fixed as $R = 0.002$ m and $L = 0.01$ m. The applied axial loading and torsion may vary for different figures.

Fig. 2 plots the $H = 0$ and $D = 0$ contours in the $\mu-\nu$ space, for $m = -2420$ kPa and $n = -2350$ kPa. The axial loading P is chosen as positive. It can be seen that the $\mu-\nu$ space is partitioned into three regions: Region I with $H > 0$ and $D > 0$, Region II with $H < 0$ and $D < 0$ and Region III with $H < 0$ and $D > 0$.

Several interesting phenomena can be observed, in agreement with the conclusions (a) to (c) stated above. First, negative Poynting effect ($D < 0$) implies negative axial force-twist effect ($H < 0$) as shown in Region II. However, the converse is not true, i.e., negative axial force-twist effect ($H < 0$) does not imply negative Poynting effect ($D < 0$) necessarily, as shown in the small Region III. Secondly, positive axial force-twist effect implies positive Poynting effect (i.e., $H > 0$ means $D > 0$, as shown

Figure 2. Contours of the Poynting effect coefficient $D = 0$ (dashed line) and axial force-twist effect coefficient $H = 0$ (solid line) in $\mu-\nu$ space for a homogeneous elastic cylinder. The contours partition the space into three regions.

in Region I). However, the converse is not true. Positive Poynting effect does not imply positive axial force-twist effect (i.e., $D > 0$ does not necessarily imply $H > 0$ as shown in region III). Moreover, when the two effects differ in sign, the Poynting effect must be positive and the axial force-twist effect must be negative (as shown in Region III). Region III, where the two effects have different signs, is generally small, suggesting that only careful choices in the material parameters can lead to different signs for the two effects.

Fig. 3 plots H and D against the Poisson ratio ν. The material parameters are based on those of polymers with $\mu = 10.3$ kPa, $m = -24.2$ kPa and $n = -23.5$ kPa. The loadings are $P = 10$ kPa and $T = 300$ kPa·m^3. It can be seen that when ν increases, both H and D decrease from positive to negative monotonically. Thus, the Poisson ratio can be an important parameter in controlling the two effects. Secondly, the magnitudes of H and D are of the order of 10^{-1}, suggesting that the nonlinear effects can be significant. Note that ν_1 and ν_2 are the particular Poisson ratios which make $H = 0$ and $D = 0$, respectively. This figure further shows that (a) if $H > 0$, then $D > 0$, as shown when $\nu < \nu_1$, (b) if $D < 0$, then $H < 0$, as shown when $\nu > \nu_2$, and (c) if H and D have different signs, then $H < 0$ and $D > 0$, as shown when $\nu_1 < \nu < \nu_2$.

Fig. 4 shows how the linear elastic constants μ and ν affect the Poynting effect and the axial force-twist effect. The parameters are $m = -360$ kPa, $n = 20$ kPa, $P = 10$ kPa and $T = 1000$ kPa·m^3. It can be seen that there exists a μ_1 for which H is independent of ν. Similarly, there exists a μ_2 for which D is independent of ν. As mentioned above,

Figure 3. Dependence of H and D on the Poisson ratio v, with $\mu = 10.3$ kPa, $m = -24.2$ kPa, and $n = -23.5$ kPa.

Figure 4. Dependence of H and D on the shear modulus μ for different Poisson ratios $v = 0.1$, 0.3, 0.4 and 0.49, with $m = -360$ kPa and $n = 20$ kPa.

μ_1 and μ_2 can be determined from the equations $(n+6\mu_1)/3 = (4m+8\mu_1)/2$ and

$(n+4\mu_2)/3 = (4m+8\mu_2)/2$, respectively, yielding $\mu_1 = 363.3$ kPa and $\mu_2 = 272.5$

kPa. A further observation is that the negative H and D values appear to have upper bounds, while the positive values are unbounded. More generally, however, D or H may either have a positive or negative bound, depending on the values of m and n.

Fig. 5 shows how the nonlinear elastic constant m can significantly influence both

effects. Here H and D are plotted against μ for $m = \pm 2 \times 10^6$, $\pm 10^6$ and 0 kPa. The

other elastic parameters are $\lambda = 60$ kPa and $n = -23.5$ kPa. The loadings are $P = 0.01$

kPa and $T = 10$ kPa·m^3. For this set of parameters, increasing μ will decrease the magnitudes of the coefficients. Secondly, both effects are positive for negative m and

negative for positive m. Changing the sign of m will change the sign of both H and D. Thirdly, decreasing the magnitude of m will also decrease the magnitudes of H and D. The magnitudes can reach the order of 10^{-2} to 10^{-1} when μ is small; thus the nonlinear behavior can be significant when the material is very soft with a small μ.

Fig. 6 plots H and D versus μ for the same sets of m, with $\lambda = 35700$ kPa and $n = -23500$ kPa. The loadings are $P = 0.01$ kPa and $T = 10$ kPa·m^3. The nonlinear effects are different from those shown in Fig. 5. For m positive, both H and D decrease to a negative maximum and subsequently decrease slowly to zero with increasing μ. However, for m negative, they decrease monotonically to zero with μ.

Figure 5. Dependence of H and D on the shear modulus μ for $m = -2\times10^6$, -10^6, 0, 10^6, 2×10^6 kPa, with $\lambda = 60$ kPa and $n = -23.5$ kPa. The loadings $P = 0.01$ kPa and $T = 10$ kPa·m^3.

Figure 6. Dependence of H and D on the shear modulus μ for $m = -2\times10^6$, -10^6, 0, 10^6, 2×10^6 kPa, with $\lambda = 35700$ kPa and $n = -23500$ kPa. The loadings $P = 0.01$ kPa and $T = 10$ kPa·m^3.

Discussion

Many biological materials, from soft to hard, are subjected to complex loading in their physiological environment. A few examples are described here. Arterial walls associated with human brain aneurysms were subjected to combined extension, torsion and inflation in finite element studies, in order to mimic the real physiological conditions (Tóth et al., 2005). The behavior of lumbar spinal units under torsion, compression and flexion/extension were also experimentally studied (Haberl et al., 2004). It is also well-known that articular cartilage is subjected to combined compression and shear during normal activities (Mansour, 2003). Fatigue tests were conducted on cylindrical bovine cortical bone specimens under axial, torsional and combined axial-torsional loadings (Vashishth et al., 2001). Finite extension and torsion were applied on capillary muscles in order to characterize their behavior under physiological conditions (Criscione et al., 1999).

Because of the prevalence of combined loadings, the Poynting effect and the axial force-twist effect may be highly relevant. In particular, large stresses may be generated by both effects if the specimen is confined in one way or another, i.e., the additional axial and rotational displacements are restrained. These large stresses can, for instance, alter the overall force balance and the cytoskeleton structure of cells, or the movement of a human red blood cell through narrow capillaries. The diameter of a human red blood cell is 7.0-8.5 μm, while that of narrow capillaries is smaller than 3 μm (Bao and Suresh, 2003).

The effects can also be utilized in the design of devices such as actuators and sensors. One can imagine a bio-inspired polymer actuator based on the axial force-twist effect, i.e., a torsionally loaded cylinder may generate an additional output twist, if subjected to an input axial force. By carefully selecting the elastic parameters of the materials and the structural dimensions, the amount of twist can be increased significantly and the desired output can be achieved.

Conclusions

Explicitly expressions for the Poynting effect, the axial force-twist effect and their relation are presented in this paper. The dependence of the relation on elastic constants is investigated.

The results show that under a tensile stress P, (a) negative Poynting effect implies negative axial force-twist effect, (b) positive axial force-twist effect implies positive Poynting effect, and (c) if the two effects differ in sign, the Poynting effect must be positive and the axial force-twist effect negative. The loadings P and T are such that (d) changing the direction of T will not change the sign of both effects, and (e) changing the direction of P will change the direction of the axial force-twist effect. Moreover, the Poynting and axial force-twist effects exhibit a very significant size

dependence, respectively of the inverse sixth and third power of the cylinder radius.

Reduced elastic coefficients characterize universal relations between the effects and the applied loadings. The elastic constants μ, ν and m have significant influence on the magnitude and direction of the Poynting and axial force-twist effects. For certain combinations of elastic constants, changing the sign of m can directly change the sign of the two effects. The two effects may have a positive or negative bound, depending on the elastic constants. From the perspective of material design, the elastic constants are thus of vital importance.

References

Bao, G. and Suresh, S. (2003) Cell and molecular mechanics of biological materials. *Nature Materials* **2**, 715-725.

Catheline, S., Gennisson, J.L. and Fink, M. (2003) Measurement of elastic nonlinearity of soft solid with transient elastography. *Journal of the Acoustical Society of America* **114**, 3087-3091.

Criscione, J.C., Lorenzen-Schmidt, I., Humphrey, J.D. and Hunter, W.C. (1999) Mechanical contribution of endocardium during finite extension and torsion experiments on papillary muscles. *Annals of Biomedical Engineering* **27**, 123-130.

Haberl, H., Cripton, P.A., Orr, T.E., Beutler, T., Frei, H., Lanksch, W.R. and Nolte, L.P. (2004) Kinematic response of lumbar functional spinal units to axial torsion with and without superimposed compression and flexion/extension. *European Spine Journal* **13**, 560-566.

Janmey, P.A., Mccormick, M.E., Rammensee, S., Leight, J.L., Georges, P.C. and Mackintosh, F.C. (2007) Negative normal stress in semiflexible biopolymer gels. *Nature Materials* **6**, 48-51.

Mansour, J.M. (2003) Biomechanics of Cartilage, in *Applied Kinesiology*, Lippincott Williams and Wilkins, Philadelphia, pp. 66-79.

Murnaghan, F.D. (1951) Finite Deformation of an Elastic Solid. John Wiley, New York.

Poynting, J.H. (1909) On pressure perpendicular to the shear planes in finite pure shears, and on the lengthening of loaded wires when twisted**.** *Proceedings of the Royal Society of London Series A-Containing Papers of a Mathematical and Physical Character* **82**, 546-559.

Tóth, B.K., Raffai, G. and Bojtar, I. (2005) Analysis of the mechanical parameters of human brain aneurysm. *Acta of Bioengineering and Biomechanics* **7**, 3-23.

Wang, D. and Wu, M.S. (2013) Second-order elasticity of soft multilayer capsules: Universal relations and parametric studies. *International Journal of Engineering Science* **73**, 17-32.

Wang, D. and Wu, M.S. (2014) Poynting and axial force-twist effects in nonlinear elastic mono- and bi-layered cylinders: Torsion, axial and combined loadings. *International Journal of Solids and Structures* **51**, 1003-1019.

Vashishth, D., Tanner, K.E. and Bonfield, W. (2001) Fatigue of cortical bone under combined axial-torsional loading. *Journal of Orthopaedic Research* **19**, 414-420.

AUTHORS INDEX